UNIVERSITÀ
DEGLI STUDI
FIRENZE

DOTTORATO DI RICERCA IN
SCIENZE BIOMEDICHE
CURRICULUM: Oncologia Sperimentale e Clinica
CICLO XXVI

COORDINATORE Prof. Persio Dello Sbarba

Development of a system to study the onset of chromosomal translocations in B-cells

Settore Scientifico Disciplinare MED/04

Tutore:
Prof. Persio Dello Sbarba

Dottorando:
Dr. Giorgio Mattiuz

Supervisore:
Dr. Silvestro Conticello

Anni 2011/2014

UNIVERSITÀ
DEGLI STUDI
FIRENZE

DIPARTIMENTO DI
SCIENZE BIOMEDICHE
SPERIMENTALI E CLINICHE

Dottorato di Ricerca in
SCIENZE BIOMEDICHE
sede amministrativa: Dipartimento di
Scienze Biomediche Sperimentali e Cliniche
coordinatore: Prof. Persio Dello Sbarba

Dottorato di Ricerca in Scienze Biomediche

Presentazione del candidato: Giorgio Mattiuz
Curriculum: Oncologia Sperimentale e Clinica
Ciclo: XXVI
Titolo della tesi: Development of a system to study the onset of chromosomal translocation in B-cells

A conclusione del corso triennale del XXVI Ciclo del Dottorato di Ricerca in Scienze Biomediche (*curriculum* Oncologia Sperimentale e Clinica), il Collegio dei Docenti, facendo propria la relazione presentata dal Prof: Persio Dello Sbarba in qualità di *tutor*, circa l'attività di ricerca, l'operosità e l'assiduità del candidato, rilascia con parere unanime il seguente giudizio da presentare alla Commissione Giudicatrice ai fini dell'espletamento dell'esame finale.

Il Dott. Giorgio Mattiuz, nato a Firenze il 19/02/1981, laureato in Scienze Biologiche 15/07/2009, discutendo una tesi dal titolo "Il ruolo della duplicazione genica nell' evoluzione delle vie metaboliche: il caso del gene *hisA*" con la votazione di 102/110, è stato/a ammesso, a partire dal 01/01/2011, al Dottorato di Ricerca in Scienze Biomediche *curriculum* Oncologia Sperimentale e Clinica (XXVI Ciclo), svolgendo la propria attività di ricerca presso l'Istituto Toscano Tumori, sotto il tutoraggio del Prof. Persio Dello Sbarba e la supervisione del Dott. Silvestro Conticello.

Descrizione dell'attività di ricerca/Risultati ottenuti:

Lo scopo della ricerca di dottorato del Dott. Mattiuz è stato la messa a punto di un sistema che permettesse di individuare e studiare i fattori coinvolti nell'insorgenza della traslocazione cromosomica *c-myc/IgH* durante i processi di diversificazione degli anticorpi. Tale traslocazione fa si che il gene *c-myc* sia posto sotto il controllo del enhancer intronico delle immunoglobuline, causandone così l'overespressione. Studi precedenti hanno dimostrato come le traslocazioni cromosomiche nelle malattie linfoproliferative coinvolgano spesso il locus genico delle Immunoglobulone (Ig) e siano direttamente connesse agli specifici meccanismi molecolari in atto nei linfociti B durante la diversificazione secondaria del repertorio anticorpale.
Il Dott. Mattiuz ha utilizzato come modello sperimentale le cellule CH12-F3, una linea cellulare di linfoma di topo in cui è possibile indurre il cambio di isotipo. In questa linea cellulare ha sostituito mediante gene targeting uno degli alleli del gene *c-myc* con la sequenza codificante un reporter fluorescente (EGFP, Enhanced Green Fluorescent Protein). Nel momento in cui ha luogo la traslocazione cromosomica, l'EGFP si attiva in quanto viene traslocata sotto il controllo degli elementi regolatori del gene IgH. Questo permette la visualizzazione mediante FACS delle cellule in cui è avvenuta tale traslocazione. Sviluppi futuri di questo sistema permetterà di analizzare sia l'effetto di fattori specifici che facilitino l'insorgenza di queste alterazioni genetiche, sia l'individuazione di fattori ancora non noti.

UNIVERSITÀ
DEGLI STUDI
FIRENZE

DIPARTIMENTO DI
SCIENZE BIOMEDICHE
SPERIMENTALI E CLINICHE

Pubblicazioni/Presentazione dati a congressi nazionali e internazionali:

Saraconi G*, Severi F*, Sala C, **Mattiuz G**, Conticello SG "The RNA editing enzyme APOBEC1 induces somatic mutations and a compatible mutational signature is present in esophageal adenocarcinomas." Genome Biol. 2014 Jul 31;15(7):417

FISV 2014-XIII Congress in Pisa. Poster presentation "AID, DNA repair and cell cycle"

Seminari:

Dottorato Anno I - Luglio 2011 "Analisys of factors involved in the onset of chromosomal translocation"
Dottorato Anno II - Settembre 2012 "Analisys of factors involved in the onset of chromosomal translocation"
Dottorato Anno III - Novembre 2013 "Bioinformatic approaches for mutational analysis"

Durante il corso di dottorato, il candidato ha seguito con il massimo impegno il programma didattico stabilito dal Collegio dei Docenti ed ha portato avanti con entusiasmo e determinazione le sue ricerche, dando prova di grande inventiva ed intraprendenza, nonché di notevole elasticità nella elaborazione dei dati sperimentali. Nel corso del triennio, il candidato ha inoltre maturato una buona cultura di base ed una vasta esperienza diretta in metodiche sperimentali.

Per quanto sopra, il Collegio dei Docenti unanime ritiene che IL Dr. Giorgio Mattiuz possa meritatamente aspirare a conseguire il titolo di Dottore di Ricerca.

Firenze, 12/02/2015

Il Coordinatore del Corso

Prof. Persio Dello Sbarba

Mendelian laws of inheritance

In the 1870s, Walther Flemming (1843-1905) exploiting the new microscopic technologies observed a phenomenon that he described as "metamorphosis" of cells. Flemming observed the onset of filamentous structures, able to absorb the basophilic dyes (coloured body), that appeared in the nucleus during the cellular division: these threadlike structures were the chromosomes. Flemming performed different studies about the cellular division and observed an equal distribution of chromosomes between the mother and daughter cell during the cell division. Initially he called this process "karyomitosis" (from greek μιτο namely thread).
The discovery of chromosomes during the cellular division (mitosis) had become one of primary focus in cytological science.

Subsequently, in the 20th century, the rediscovery of Mendelian Genetic combined with the new discoveries about the cellular division, brought into the spotlight the concept of units of heredity.
Theodor H. Boveri (1862-1915) and Walter S. Sutton (1877-1916) formulated their own theory of "Mendelian laws of inheritance" (1902-1903) in which they clarified the role of chromosome during mitosis: chromosome could be the vehicles of the genetic material.
Boveri observed that the entire chromosomal set was necessary to a correct embryonic development; on the other hand, Sutton formulated the hypothesis that chromosomes could follow the Mendelian laws.

Initially, the "Mendelian laws of inheritance" was not widely accepted by the scientific community; in the 1915s, Thomas H. Morgan (1866–1945), studying Drosophila Melanogaster, demonstrated that chromosomes carry the genes and play a crucial role in heredity. This discovery allowed him to win, in 1933, the Nobel Prize in Physiology or Medicine.

Morgan, with his studies, confirmed another important observation from Theodor Boveri; Boveri hypothesized that uncontrolled cell grown of
could be directly connected with an incorrect chromosomal composition, so this could be one of the causes of carcinogenesis.

Table of abbreviation V

Chapter I 1

Introduction 1

Lymphoproliferative diseases (LPDs) 1

B-cells receptor development 3

Antigen - independent: primary diversification 6

Antigen - dependent: secondary diversification 8

Immunoglobulin gene conversion 8

Somatic Hypermutation (SHM) 9

Class Switch Recombination (CSR) 9

Activation Induced Deaminase (AID) 11

Discovery 11

Structure and function 12

Alternative splicing of AID 14

AID in secondary-immune response 16

AID in Somatic Hypermutation (SHM) 16

AID in Class Switch Recombination (CSR) 18

Recruitment of AID 20

Homologous Recombination and Non-homologous end-joining: two pathways to repair DBSs 21

The repair of DSBs in CSR 23

Chromosomal Translocations 24

AID and B-cell tumors 25

The translocation t(8;14) in Burkitt's lymphoma 27

Aid outside B-cell 30

Aim of the work 31

Chapter II 33

Solution and buffers 33

Bacteria 34

Bacteria Techniques 35

Preparation of chemically competent bacteria 35

Transformation of chemically competent bacteria 35
Storage of bacterial strains 36
Cracking System: a massively screening for bacteria transformation 36
Eukaryotic Cells 37
Cellular techniques 38
Maintenance of cells 38
Freezing of cells 38
Thawing of the cells 38
Transfection of eukaryotic cells 38
Analysis of Class Switch Recombination 39
Analysis of the expression levels 39
CRE-Lox system 40
Molecular Biology techniques 41
PCR 41
Polymerase Chain Reaction (PCR) protocols 41
PCR programs 43
Reverse Transcription (RT-PCR) 44
Plasmid DNA extraction 44
DNA extraction from eukaryotic cells 44
Agarose gel electrophoresis 44
DNA fragment purification from gel 45
TOPO cloning kit 45
DNA sequencing 45
Statistical analysis 45
Southern Blot 46
Genome editing 47
TALENs 47
CRISPR/Cas9 47
Plasmids construction 48
Chapter III 50
Results and Discussion 50

Designing a vector for Knock-in CH12-F3 51
Building the Knock-in vector 53
Transfection of the Knock-in construct and first screens 54
TALEs and TALENs 55
Visualizing MYC/IgH translocation 58
Expression of AID mutants to increase the translocation frequency 58
Hyperactive AID 60
Caffeine treatment to promote translocation 60
CRISPR-Cas system 62
Inducible promoters 63
Work in Progress 64
LITE system 64
CRISPR KNOCK-OUT P53 in CH12-F3 myc+/EGFP 66
Conclusions 67
Splice variants of Activation Induced Deaminase (AID) do not affect the efficiency of Class Switch Recombination in murine CH12-F3 cells 68
Introduction 68
Results and Discussion 69
Conclusions 73
Appendices 74
A - Oligonucleotides 74
B - Plasmids and vectors 77

Table of abbreviation

Abbreviation	Full name
A	Adenine
aa	Amino acid
Ab	Antibody
AID	Activation Induced Deaminase
A-EJ	Alternative End Jonig
APE	Apurinic/apyrimidinic Endonuclease
APOBEC1	APOlipoprotein B mRNA Editing enzyme Catalytic polypeptide
Amp	Ampicillin
BCL2	B-cell lymphoma 2 gene
BCL6	B-cell lymphoma 6 gene
BCR	B cell receptor
BSR	Blasticidin S resistance gene
BER	Base Excision Repair
bp	Base pair
C-terminal	Carboxyl-terminal
cDNA	Complementary DNA
C	Cytosine
cds	coding sequence
CDR	Complementary determining region
CSR	Class Switch Recombination
DNA	Deoxyribonucleic Acid
FBS	Fetal Bovine Serum
CRISPR	Clustered Regulatory Interspaced Short Palindromic Repeat

sgRNA	single guide RNA
dNTP	DeoxyNucleoside TriPhosphate
DBS	Double Strand Break
DLBCL	Diffuse large B-cell lymphoma
EGFP	Enhanced Green Fluorescence Protein
FACS	Fluorescence-activated cell sorting
G	Guanine
GC	Gene Conversion
GC	Germinal Center
HR	Homologous Recombination
IRES	Internal Ribosome Entry Site
HRS	Hodgkin and multinuclear Reed Sternberg cells
Ig	Immunoglobulin
Ig C-region	Immunoglobulin constant region
IgH	Immunoglobulin heavy chain
IgV	Immunoglobulin variable region
IL-4	Interleukin-4
mRNA	messenger RNA
MMR	Mismatch Repair
PAM	Protospacer-Adjacent Motif
MSH2	MutS Homolog 2
MSH6	MutS Homolog 6
Puro	Puromycin
NES	Nuclear Export Signal
NGS	Next Generation Sequencing
RNA	Ribonucleic Acid
NHEJ	Non-Homologous Recombination
NHL	non Hodgkin Lymphoma

NLS	Nuclear Localisation SIgnal
PCR	Polymerase Chain Reaction
PI	Propidium Iodide
RAG	Recombination Activating Gene
RPA	Replication Protein A
RSS	Recombination Signal Sequences
RT-PCR	Reverte Transcription PCR
S region	Switch region
SHM	Somatic Hyper-Mutation
T	Thymine
TGF-β	Transforming Growth Factor-β
ts	Transition
tv	Transversion
UNG	Uracil-N-Glicosylase
V, D and J	Variable, diversity and junction

Chapter I

Introduction

Lymphoproliferative diseases (LPDs)

Lymphoproliferative diseases (LPDs) are tumors that derive from white-blood cells of the immune system and are characterized by specific genetic features strictly correlated with the developmental stage of the cells of origin, be them B, T or NK cells. LPDs present different features and can be divided in two groups: leukemias and lymphomas. They are characterized by a number of genetic alterations (point mutations, insertions and deletions) and are classified according to these alterations, such the presence of common chromosomal rearrangements {Look, 1997; Raimondi, 1999}. Overall a single genetic alteration is not be sufficient to triggers the disease {Hanahan and Weinberg, 2000}, yet many LPDs are characterized by a very small number of lesions, compared to other types of cancer, and these alterations play a crucial role in the phenotype and in the development of the tumor.

Leukemias comprise of a large number of blood cancers that take origin from immature white cells in the bone marrow and are characterized by an uncontrolled clonal proliferation; 90% of leukemias affect adults, and they represent the most common type of neoplasm in children (about 70% of cases). Depending on the nature of their precursor, leukemias are divided in myeloid and lymphoid, and, depending on their aggressiveness, are either acute or chronic.

Myeloid leukemias derive from myeloid blood cells and are characterized by a rapid growth of white blood cells able to accumulate on bone marrow; the leukemic cells replace the normal bone marrow cells and interfere with the physiological production of cells in this compartment. This substitution induces a sharp decrease in the number of red blood cells, platelets and healthy white blood cells.
Lymphoid leukemia, either acute or chronic, result from the exponential growth of circulating neoplastic lymphocytes.
Acute lymphoid leukemia (ALL) is a highly proliferative and invasive neoplasm that originates in the bone marrow and can metastasizes to different organs such as spleen, liver, nervous system, lymph nodes, and gastrointestinal tract. On the other hand, chronic lymphoid leukemia (CLL), is less proliferative in nature. Diagnosis of CLL can be very difficult as, especially in the early stages, the cells involved are low in number and difficult to detect, due to the possible lack of cytological features in the neoplastic cells.
While leukemias principally involve blood and bone marrow, on the other hand lymphomas originate from mature cells in lymph nodes and can be detected through the observation of the mass of lymphoid organs (lymphadenopathy).

Lymphomas are a heterogeneous group composed by two categories: Hodgkin lymphomas (HL) and non-Hodgkin lymphomas (NHL).
Hodgkin lymphoma is one of the most common lymphoma and originates from mature B-cells; the neoplastic clones are a mixtures of mononuclear Hodgkin and multinuclear Reed Sternberg cells (HRS). Hodgkin cells represent the proliferative part of the disorder and derive directly from mature B-cells; Reed-Sternberg cells are normally unable to proliferate and take origin from Hodgkin cells due to errors during mitosis by which the nuclear division is not followed by the cell division {Newcom, 1988; Ikeda, 2010;}. Origin of HRS cells has been clarified by experiments demonstrating the presence of clonal and somatic mutations on immunoglobulin heavy and light chain gene, and rearrangements such as those happening during B-cell development {Küppers, 1994}.
HL cells present an unusual phenotype due to co-expression of various hematopoietic markers and due to a global down-regulation of gene expression with respects to healthy B-cells.

Non-Hodgkin lymphomas (NHL) are a heterogeneous group formed by neoplastic T and B-cells that arise initially in secondary lymphoid tissues.
NHL represents about 70% of lymphomas, and 80% of NHLs take origin from B-cells.

Studies on the origin of genetic alterations in B-cell lymphomas have revealed how many of the most characteristic genetic lesions arise from failure of the unique molecular processes that occur in B-cells during their development. Thus, point mutations on non-immunoglobulin genes that resemble those observed during Somatic Hypermutation process can be found {Gaidano, 2003; Montesinos-Rongen, 2004; Liso, 2006; Deutsch, 2009}. For example somatic mutations on the BCL6 oncogene present a mutational pattern similar to that observed on the immunoglobulin loci {Pasqualucci, 1998; Capello., 2000; Pasqualucci, 2000}. Moreover, reciprocal translocations between a proto-oncogene and the regulatory element of the B-cell receptor (BCR) are common; this translocation results in a high constitutive expression of the proto-oncogene, which can promote the cancer development {Küppers and Dalla Favera, 2001; Lossos, 2002; Jardin, 2003; Küppers, 2005}.
These recurrent alterations are used as hallmark to describe and classify LPDs and show a direct correlation with the expression of AID, the effector of the antigen-driven antibody diversification processes {Bödör, 2005; Reiniger, 2006}.

B-cells receptor development

B-lymphocytes are important elements involved in the immune response as they play a crucial role in the humoral immunity of the adaptive immune system through the production of antibodies (Ab). The main role of the antibodies depends on their particular ability to recognize the antigens and present them to the effectors of immunity. Antigens (Ag) are structural molecules originate from the body (self) or from external environment, which are the targets of T- and B-cells receptors. This interaction triggers the uniques molecular processes of immune system that allow an increase the affinity and specificity against antigens.

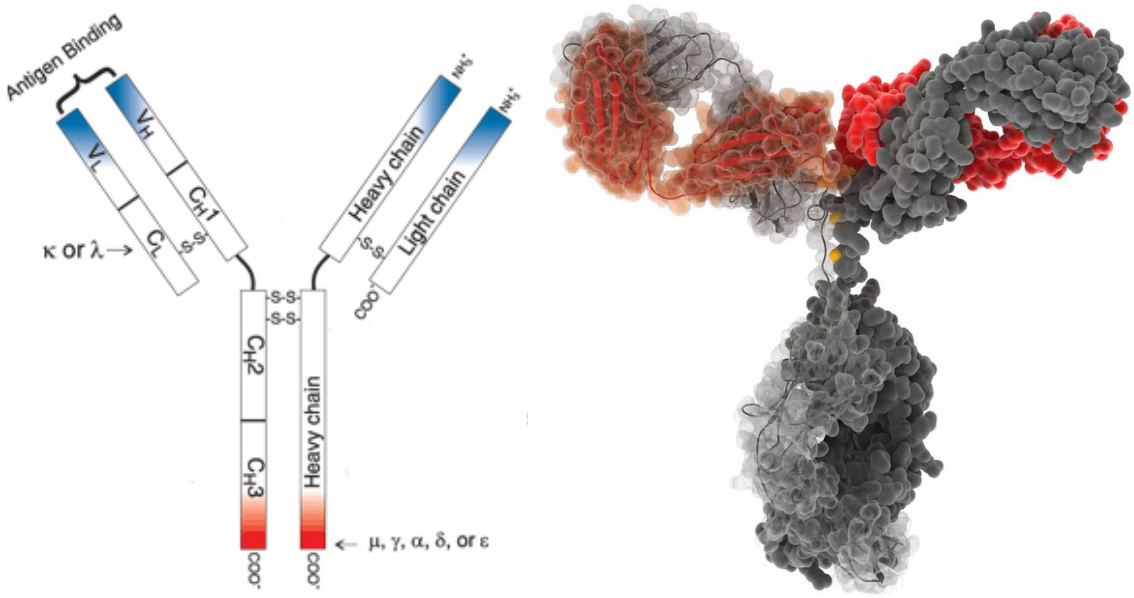

Figure 1.1 - Structure of B-cell receptor. A) The basic structure of mammalian immunoglobulins is composed by four polypeptide chains—two light chains and two heavy chains, which are connected by disulfide bonds. The N-terminal region (blue) is involved in the recognition of the antigens. **B)** Chemical structure of an immunoglobulin G (IgG, antibody) molecule.

Also know as immunoglobulins (Ig), antibodies can be either transmembranal or soluble glycoproteins, and are composed by two heavy chains (IgH) covalently bound to two light chain (IgL). Both chains are formed by a constant (C) region in the carboxy-terminal part and a variable (V) region in the amino-terminal portion (**Figure 1.1**).

C-regions of the heavy chains determine the isotype of the immunoglobulins. This domain specifies the particular function of any given antibody in relation to the immune system. In mammals there are five types of IgH, called with greek letters: μ, δ, γ, ε that encode for different classes of antibodies: IgA, IgD, IgE, IgG and IgM., respectively. On the other hand, the light chains present only two forms: the kappa (κ) and the lambda (λ), In humans immunoglobulin heavy chain genes locate on chromosome 14, immunoglobulin kappa genes on chromosome 2 and lambda genes chromosome 22 (**Figure 1.2**).
Recognition and binding to the antigen (Ag) depend on the complementarity determining regions (CDR), which are hypervariable stretches within V-regions in both heavy and light chain genes.

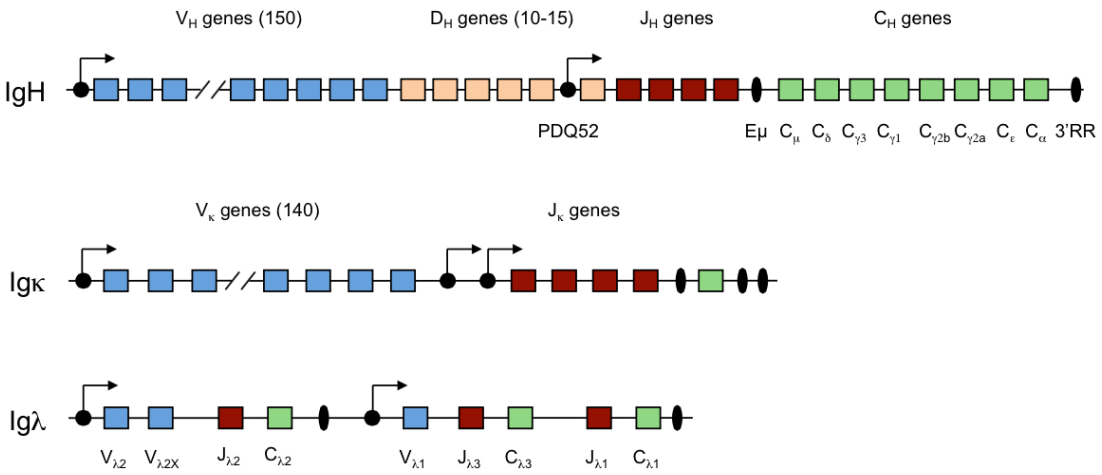

Figure 1.2 - Structure of mouse immunoglobulin loci. IgH locus is composed by about 150 variable (V) genes, 15 diverse (D) genes and 4 joining (J) gene and 8 constant (C) genes; Igκ presents 140 V genes and 4 J genes and one C exon; Igλ locus has a small number of variable exons positioned in distinct cassettes Diagram not drawn in scale; Diagram not drawn in scale; adapted from {Cobb, 2006}.

The estimated number of antigens that can be recognized by the antibody repertoire is higher than 10^{11}. The repertoire is obtained through a two-step diversification process that happen alongside the maturation of the B-cells.
This process involves a primary diversification, happening before birth, and a secondary one, which is triggered by moment the naïve B-cell encounters the antigen.
Different specific processes contribute to the diversification of the immunoglobulin genes: V(D)J recombination, which takes place during the early stages of both B- and T-cells in the bone marrow without encountering antigen is defined independent; the others, classified antigen-dependent, are the Somatic Hypermutation (SHM), the Class Switch Recombination (CSR), and in species such as chicken and rabbit, Immunoglobulin Gene Conversion (IGC). These latter processes are specific to B-cells in the germinal center (GC) of secondary lymphoid organs (**Figure 1.3**).

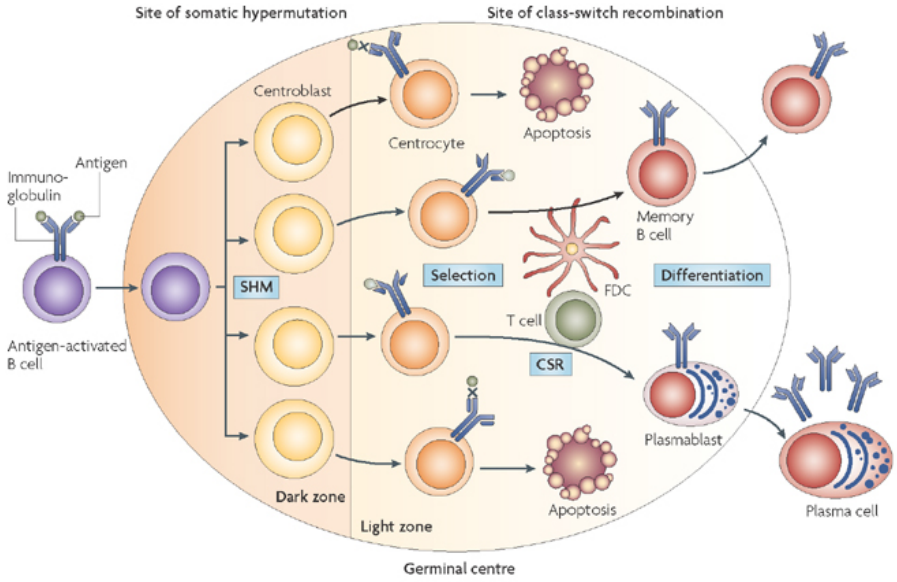

Figure 1.3 - Differentiation of B-cells in the germinal center. Naïve mature B cells activated through 'T-cell help' signal migrate in secondary lymphoid organs establishing germinal centre. The dark zone represents the proliferating B-cells in which the process of secondary diversification of the immunoglobulin genes is activated. This process increases the affinity of the antibody against the antigen introducing mutations on the variable region. Unfavorable mutations lead the cell to apoptosis, however the acquisition of advantageous mutations on the variable region selects positively the B-cells and allow them to migrate to another zone of the GC (light zone). There the mature B-cells differentiate into memory or plasma cells and move out from GC. (Reproduced from Küppers, 2005).

Antigen - independent: primary diversification

The V(D)J recombination occurs during the maturation of T- and B-cells in bone marrow (primary lymphoid organ), and consists in a recombinational process that alters the primary sequence of the B- and T-cell receptor and the Ig genes. The process acts on an array of segments, upstream the constant region of the heavy and light chain, to produce functional T and B cell receptors (TCR, BCR). These segments belong to three types: variable (V), diverse (D) and joining (J). This process directly depends on the activity of Recombination-Activating Genes 1 and 2 (RAG1 and RAG2) {Oettinger, 1990; Agrawal, 1998}.

RAG recombinase enzymes are able to bind specific sequences, flanking the genes V, D and J, called recombination signal sequences (RSSs); RSSs are positioned at the 3' side of the V regions, at the 5' side of the J regions and on both sides of D regions. RSS are composed of seven conserved nucleotide (heptamer), a non-conserved spacer sequence (12 - 23 base pairs) and an AT rich nonamer {Sakano,1980} (**Figure 1.4**). The genetic organization of the region dictates the order by which the recombinatorial reaction will proceed after the cleavage mediated by the RAG complex.

This multistep process starts in pre-pro B-cells, through the expression of the RAG genes; RAG1 bind the nonamer in the RRS sequence {Swanson and Desiderio, 1998}, thus allowing the RAG complex to interact with the heptamer and to introduce two ssDNA nick in the nearby spacer, respectively at positions 12 and 23 bp (also know as 12/23 rule) {Tonegawa, 1983}.

The double nick causes a double strand break (DBS) and it is repaired by non-homologous recombination (NHEJ) {Malu; 2012}; NHEJ merges the DNA fragments through the micro-homology in the conserved heptamer. The result is a rearranged V(D)J region. Pro-pre B-cells that successfully conclude

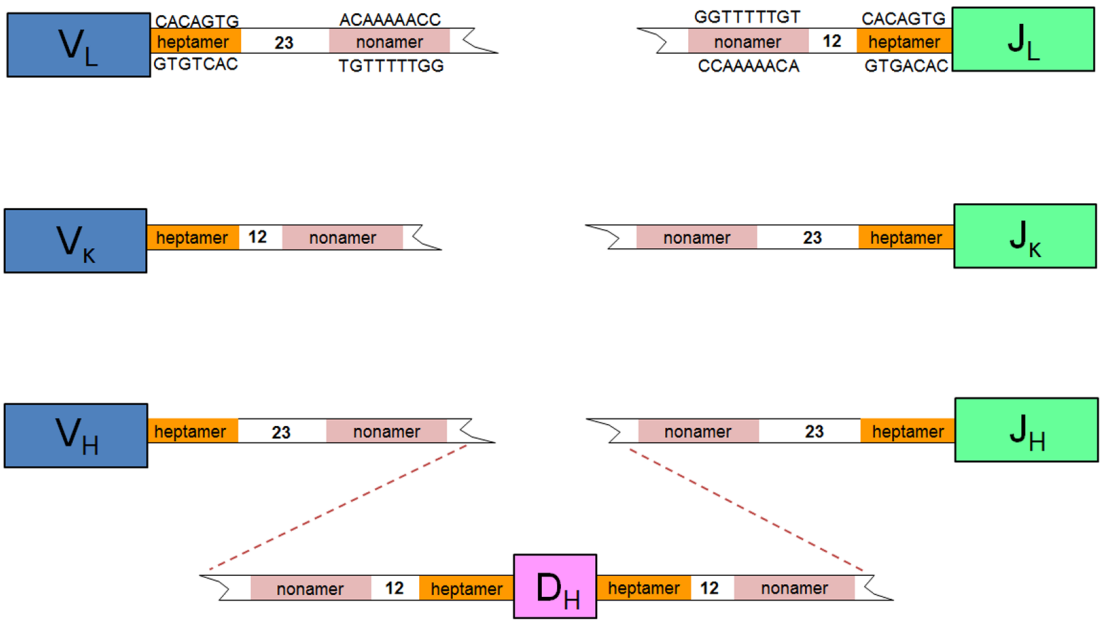

Figure 1.4 - V(D)J region. RSS sequences are composed by an heptamer (in orange) and nonamer (in pink) interspersed with a non-conserved spacer sequence (12/23 bp) (in white). RAG1/RAG2 genes induce the double strands binding the conserved sequences (heptamer and nonamer) and introducing single nicks in 12/23 fragments. The microhomology due to nonamer allow to repair the lesion through NHEJ, thus rearranging the immunoglobulin locus.

the V(D)J (naïve B-cells) recombination are able to survive and to migrate to the secondary lymphoid organs where they will undergo the secondary diversification processes to increase the affinity and the specificity to the antigen.

Antigen - dependent: secondary diversification

The antibody repertoire in naïve B-cells can potentially recognize a large amount of antigens with low affinity. Upon encountering the antigen a specific program initiates in the B-cell involving proliferation and diversification of the Ig genes. Cells that successfully conclude this multi-step process, differentiate in two sub-type: plasma cells and memory B-cells, whereas those cells which end up with a non-functional BCR, are selected against {LeBien and Tedder, 2008}. Plasma cells are able to produce large amounts of high affinity antibodies, whereas memory B-cells are fundamental for long term immunity.

Immunoglobulin gene conversion

This molecular mechanism is the archetypal diversification process in tetrapods, which is found in species such as birds, rabbits, cattle and horses. Diversification of the variable region through IGC requires the presence of Ig pseudogenes upstream to the functional Ig loci. These pseudo genes are then used as templates for homologous recombination (HR) to repair the variable region. This diversity within the array of pseudogenes allows the introduction of templated mutations that are responsible for the increase in the affinity for the antigen (**Figure 1.5**) {McCormack, 1991; Reynaud, 1985; Reynaud, 1987; Reynaud, 1991; Reynaud, 1989; Thompson and Neiman, 1987}.

Somatic Hypermutation (SHM)

SHM acts on the productive allele on V(D)J region of immunoglobulin genes in a short window during B-cells development (**Figure 1.5**). Similarly to IGC, the aim of this process is to introduce mutations on the CDR of the variable regions of heavy and light chains. In contrast to IGC, SHM introduces non-templated mutations, insertions (ins), or deletions (del)) {Di Noia, 2007; Peled, 2008}. Sequence analysis reveal that the frequency of mutations at the immunoglobulin loci is orders of magnitude higher (10^{-5}-10^{-3} mutations/bp/generation) than in the rest of the genome (estimated at 10^{-9}). The region targeted for mutations extends from about 150 nucleotides after the promoter to 1-2 kilobases (kb) downstream {Saribasak, 2012}. The mutation frequency at C:G and A:T pairs appears to be similar, while - if we consider the single targeted base - the mutation frequency between bases is different {Jansen, 2006}. With regards to A:T basepairs there is a bias between the transcribed and the un-transcribed strand: adenines on the transcribed strand present a mutation rate two fold higher than that of thymines {Rada, 1998}; on the other hand, there is no strand preference with regards to C and G mutations. Nonetheless, while no mutational hotspots have been identified for A:T pairs, mutations at C:G occur preferentially within a WRCY motif (W=A or T, R=G or A, C=targeted C, Y=C or T) or the inverse RGYW {Wagner, 1995; Petersen-Mahrt, 2003}.

Class Switch Recombination (CSR)

Class Switch Recombination is a deletetional/recombinational process that changes the isotype of the expressed antibody through an alteration of the heavy chain primary sequence. In contrast to SHM, CSR promotes the rearrangements of both immunoglobulin heavy chains alleles. The outcome of the process is the excision of the region between two S regions and the joining of the resulting ends {Chaudhuri, 2007; Ono, 2007}. This is possible due to the organization of the IgH locus, which is formed by an array of

regions encoding for the constant domains (Cμ-Cδ, Cγ, Cε and Cα) spaced by repetitive regions called Switch regions (S).
The array is precede by an intronic enhancer, and each region is preceded by a promoter. Similar to the V(D)J recombination, also this process requires the generation of two DBS. The first DBS occurs on the donor region - always Sμ -, the second one occurs in the acceptor regions (Sγ, Sε or Sα).
Naïve B-cells express IgM/IgD and the CSR allow them to change the isotype to IgG, IgE or IgA depending on the chromosomal rearrangement (**Figure 1.5**).

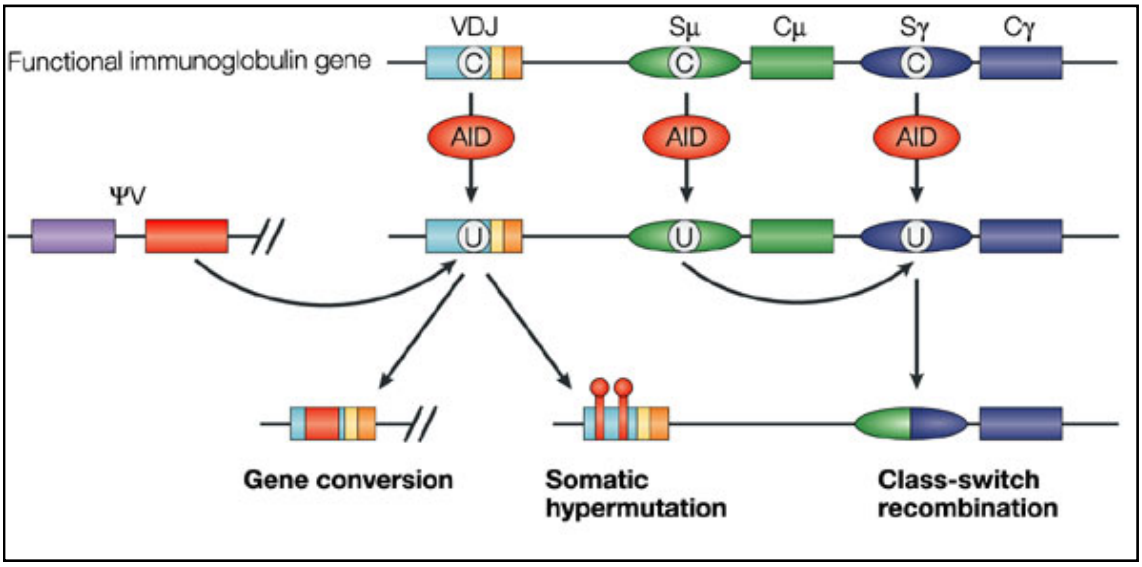

Figure 1.5 - **Secondary immune diversification of B-cells in germinal center.** Activation-induced deaminase (AID) deaminates cytosines to uracils in the immunoglobulin gene and triggers gene conversion, somatically hypermutation and class-switch recombination. The uracil (U) is removed by the cellular uracil DNA glicosylase UNG2. V (variable), D (diversity), J (joining) forming the Variable region of the immunoglobulin gene; S and C, Switch and Constant regions, respectively {Reuben, 2004}.

In 1999, Tasuku Honjo and Anne Durandy, published their studies about a new member of RNA editing deaminase family identified in B-cells from germinal centers in humans and mice. The protein - activation-induced cytidine deaminase (AID) is the key player in all three molecular processes responsible of secondary immune response: IGC {Arakawa, 2002} SHM and CSR {Muramatsu, 2000; Revy, 2000}.

The discovery of AID opened new avenues in the study and in the understanding of these important and peculiar processes, and in their mis-regulation.

Activation Induced Deaminase (AID)

Discovery

Activation-induced deaminase (AID) was first identified in a murine lymphoma cell line (CH12-F3) by a subtractive cDNA screen in which unstimulated and CSR-stimulated cells were compared {Muramatsu, 1999}.
The demonstration that patients affected by type 2 Hyper-IgM Syndrome and - experimentally - knock-out mice for AID lacked the ability to undergo antigen-driven antibody diversification revealed AID as the key regulator in both CSR and SHM {Muramatsu, 2000; Revy., 2000}; subsequently studies on chicken B-cell line (DT40) demonstrated that also IGC directly depended on the action of AID {Arakawa, 2002}.
Initially, AID was thought to be an RNA-editing enzyme, because of its similarity with APOBEC1, an RNA editing enzyme .
Apolipoprotein B Editing Complex (APOBEC1) is expressed in the small intestine and liver (in mice) and deaminates a specific cytosine (6666 in humans) on the mRNA of apolipoprotein B; the deamination introduces a stop codon that leads to the translation of a shorter protein involved in the biosynthesis of chylomicrons {Navaratnam 1993; Teng, 1993}. The proposed model involved RNA editing mediated by AID that would activate a putative recombinase involved in the SHM and CSR processes {Muramatsu, 1999; Muramatsu, 2000}.

An alternative model was proposed by Michael Neuberger, who suggested that AID was a DNA editor acting directly on the Ig loci. To understand the physiological target of AID a mutation assay was performed in *Escherichia Coli* and it showed how the presence of AID could induce a mutator phenotype by deaminating deoxycytosines to deoxyuracils {Petersen-Mahrt, 2002}. More evidence linking AID to direct targeting of cytosines was produced immediately afterwards by showing that the repair pathway

downstream to the action of AID involved excision of the deaminated cytosine through the Uracil-DNA-Glycosilase (UNG) {Di Noia, 2002; Rada, 2002}. Mice B-cells lacking for UNG gene reduce the efficiency of SHM and CSR {Rada, 2002}; these experiment demonstrated that the physiological substrate of AID is the DNA strand.
To clarify whether the mutational pattern introduced by AID was random or not, *in vitro* experiments were performed to demonstrate that AID preferentially targets cytosines within a WRC hotspot. This hotspot explained the WRCY/RGYW hotspot found on the Ig loci {Wagner, 1995} as these latter ones can be regarded as WRC motifs overlapping on the opposite strand {Beale, 2004; Rogozin and Diaz, 2004}. This observation also suggested the origin of the DSBs observed in the S regions, as AID-mediated DNA damage on the opposite strands could lead to the formation of DSBs during the CSR. From an evolutionary perspective, it is indicative the coevolution between AID as the trigger of the secondary immune diversification processes and the intrinsic characteristics of the gene locus of the immunoglobulin.

Structure and function

The structure and the function of activation-induced deaminase is still not totally characterized, because we are not able to obtain a sufficient amount of active recombinant protein to create a three-dimensional structure of AID; anyway sequence and mutagenesis analysis lead scientists to be able to obtain important informations.

The gene encoding for AID is located on chromosome 12 in *homo sapiens* and codes for a 198 amino acid protein. AID belongs to a family of cytosine deaminases {Conticello, 2005} which comprises several members: APOBEC1, previously mentioned, APOBEC2, and the APOBEC3 cluster.
APOBEC2 is expressed in muscle {Liao, 1999}, and seems to be related to muscle fibre development {Sato, 2010}. Yet, its function is still not clear. The APOBEC3s {Bishop 2004; Yu, 2004}, in particular APOBEC3G {Conticello, 2003; Mangeat, 2003} and APOBEC3F {Zheng, 2004; Liddament, 2004} are involved in the innate immune response against retrovirus {Harris 2003; Lecossier 2003; Mangeat 2003}. Phylogenetic sequence analysis performed

to understand the origin and evolution of these deaminases revealed that the AID/APOBECs are a subgroup of deaminases restricted to vertebrates. Homologues of AID and APOBEC2 were found in the genome of cartilaginous and bony fish and they were identified as the ancestral members of the AID/APOBEC family. On the other hand APOBEC1 is present in all tetrapods. While its physiological role in mammals is established as an RNA editing enzyme, it is not clear yet whether an additional role targeting DNA could be present {Severi, 2011}. The APOBEC3 genes are restricted to mammals, where the cluster has undergone a rapid evolution and takes origin directly from AID. In humans it has expanded to include seven different APOBEC3s (A to H) {Conticello, 2005}. Catalytic core of AID resembles that of others Zinc-dependent deaminases.

Indeed, starting from the three-dimensional structure of *E. coli* cytidine deaminase it was possible to identify the catalytic domain (Zn^{2+} binding domain) common to all cytidine deaminase. This domain is characterized by an His-X-Glu motif flanked by another motif Pro-Cys-X-X-Cys {Betts, 1994). The tertiary structure of the deaminases bring these two motives nearby, and the two cysteine and the histidine serve to bind the Zinc atom necessary for the reaction. The deamination takes place in this site, where the nucleotide is held in the catalytic pocket {Carter, 1995; Xiang, 1996; Xiang, 1997; Ireton, 2003; Ko, 2003}.

The deamination reaction starts from a nucleophilic reaction in which the zinc atom in the catalytic pocket of the deaminase binds an activated water molecule through its fourth bond to generate an hydroxyl nucleophile; this nucleophilic attack on the C4 position of the targeted cytosine is followed by tetravalent intermediates that end up with the release of ammonia and the generation of uracil (deoxycytosine+H_2O = deoxyuracil+NH_3) {Carter; 1995}.

AID also contains an N-terminal nuclear localization signal (NLS) (amino acids 9 to 20) and a C-terminal nuclear export signal (NES) (amino acids 183 to 190) {Ta, 2003; Ito, 2004; Brar, 2004; Geisberger, 2009}. Deletion of the terminal regions of the protein reveals other functions: the deletion of the N-terminal domain damages the SHM process {Shinkura, 2004}; on the other hand the

loss of C-terminal domain directly affects the CSR {Barreto, 2003; Imai, 2005; McBride, 2004; Ta, 2003}.

Alternative splicing of AID

Alternative splicing is a process in eukaryotic cells fundamental for the regulation of gene expression and for the generation of diversity starting from a single gene. The eukaryotic gene sequences are composed by exons and introns that are processed during the maturation of messenger RNA: exons may be included or excluded from the final mRNA thus generating different coding sequences, which - upon translation - can display differences in their biological function (**Figure 1.6**).

Moreover alternative splicing is physiologically used to regulate the expression of a given gene, the processing of its mRNA, for example by inducing nonsense-mediated decay or the translation of an inactive protein {Black, 2003}.

Alternative splicing of AID was first observed in lymphoproliferative disease {Albesiano, 2003; Greeve, 2003; McCarthy, 2003; Oppezzo, 2003; Babbage, 2004; Forconi, 2004; Heintel, 2004; Wu, 2008; Marantidou, 2010; Palacios, 2010}. AID splice variants were detected as well in normal B-cells in both humans and mice {Greeve, 2003; Oppezzo, 2003}. Although the presence of AID splice variants in B-cell tumor do not affect the mutational pattern, It has been hypothesized that they may play a role in the regulation of full length AID. Subsequent experiments demonstrated that B-cells expressed the isoforms of AID cannot reconstitute the CSR process {van Maldegem, 2009; van Maldegem, 2010}. These observation support the hypothesis that AID isoforms lack for deamination activity.

The isoforms in *homo sapiens* are four and they were hypothesized to play a role in the regulation of full-length AID reducing the amount of AID-FL mRNA or controlling the localization of AID-FL inside and outside from the nucleus of the cell {van Maldegem, 2009}.

The full length isoform of AID is encoded by five exons. the other ones resulted from the deletion of exon 4 or of all exons coding for the active domain, or insertion of intron 4. Similar to the full length, AID-ΔE4a lacks on the first part of exon 4 of few amino acid; AID-ΔE4 presents a stop codon

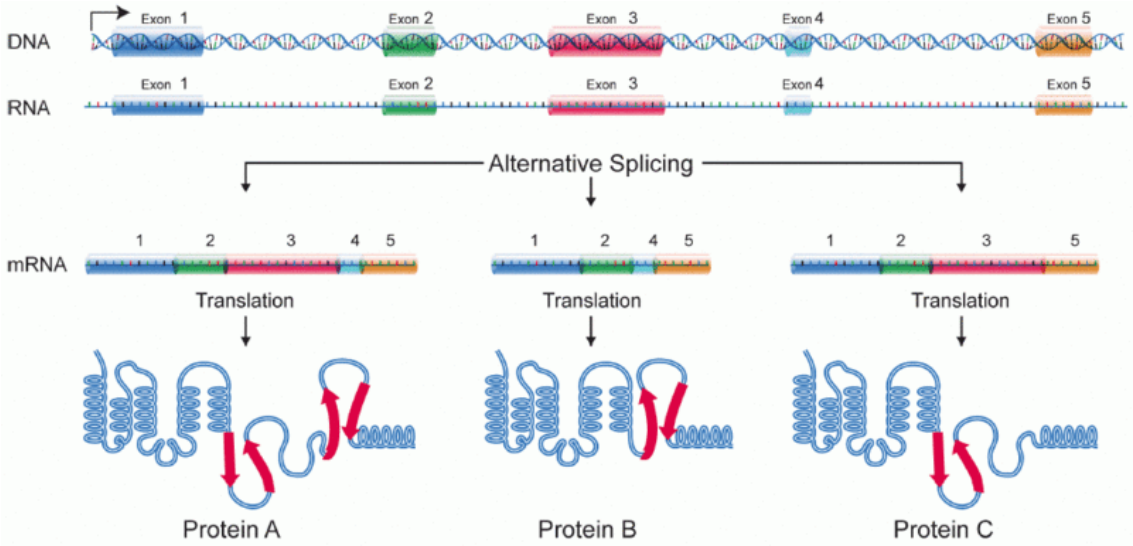

Figure 1.6 - Schematic representation of alternative splicing. Different transcripts take origin from the same coding gene leading the cell to produce different proteins from the same gene.

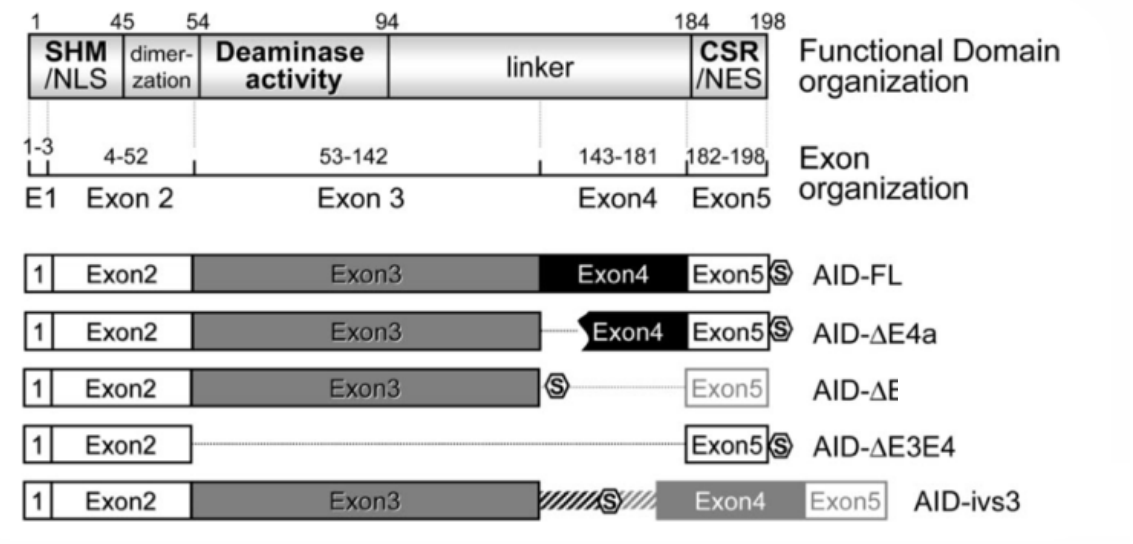

Figure 1.7 - Schematic organization of human AID isoforms. The alignment of the splice variants with the full-length protein reveals different domain of AID. Activate-induced deaminase is encoded by five exons, some of which are selectively excluded from AID splice variants. The numbers indicate the amino acid positions. Hexagon-shaped stop signs mark the stop codons of the various products. Exon and intron sequences that are not translated are in light gray color. Splice variants indicated with the arrow have been used in the experiments reported in this chapter. Picture modified from {Wu, 2008}.

after exon 3; AID-ΔE3ΔE4 formed by exons 1, 2 and 5 and AID-ivs3 isoforms gains the intron 4 changing the frameshift introducing a stop codon approximately 50 amino acids before the normal one (**Figure 1.7**).

AID in secondary-immune response

AID in Somatic Hypermutation (SHM)

As mentioned above, SHM is a process that occurs in germinal center B-cells and introduces mutations, insertions and deletions on the complementarity determining region (CDR) of the variable regions in the Ig genes. This process results in the evolution of an antibody with a higher affinity for a specific antigen. This process is strictly dependent on the DNA lesion introduced by AID and the pathways involved in the repair of the damage.
The dU:dG mismatch generated by AID is recognized by Uracil-N-Glycosylase (UNG), an enzyme devoted to remove the uracil leaving an apurinic/ apyrimidinic site (AP site). This abasic site is cleaved by the apurinic/ apyrimidinic endonuclease (APE) inducing a single strand break which can be processed and successfully repaired by DNA polymerase β. On the other hand the lesion can be processed through mutagenic repair pathways that can insert the mutations in the V(D)J region.
Even before the identification of AID as the trigger of SHM, studies suggested that the SHM process takes place in two distinct phases that are directly dependent from the different repair pathways involved. Whereas the proportion of C:G and A:T mutations in the V(D)J region from normal B-cells was roughly equal, the mutational pattern observed in B-cells deficient in Mismatch repair (MMR) was heavily biased towards mutations at C:G pairs {Rada, 1998}. This suggested that mutations at A:T pairs were introduced in a distinct, possibly later phase during the SHM process.
Eventually the mutational processes occurring during SHM could be classified in phases: those resulting in C:G mutations, phase 1A and 1B, respectively, and phase 2, responsible for the AT mutations.

The phase 1A occurs during the replication (phase G1-S of the cell cycle) and is characterized by a preference for C to T or G to A transitions (**Figure 1.8**).

the transitions occur because the uracil is read by the DNA polymerase as a thymine, bypassing UNG, and thus an A is inserted in place of a G. Phase 1B is triggered by UNG: UNG removes the uracil, introducing an abasic site which is repaired by base excision repair (BER). BER can introduce either transitions or transversions (replacement of a pyrimidine with a purine and *vice versa*)

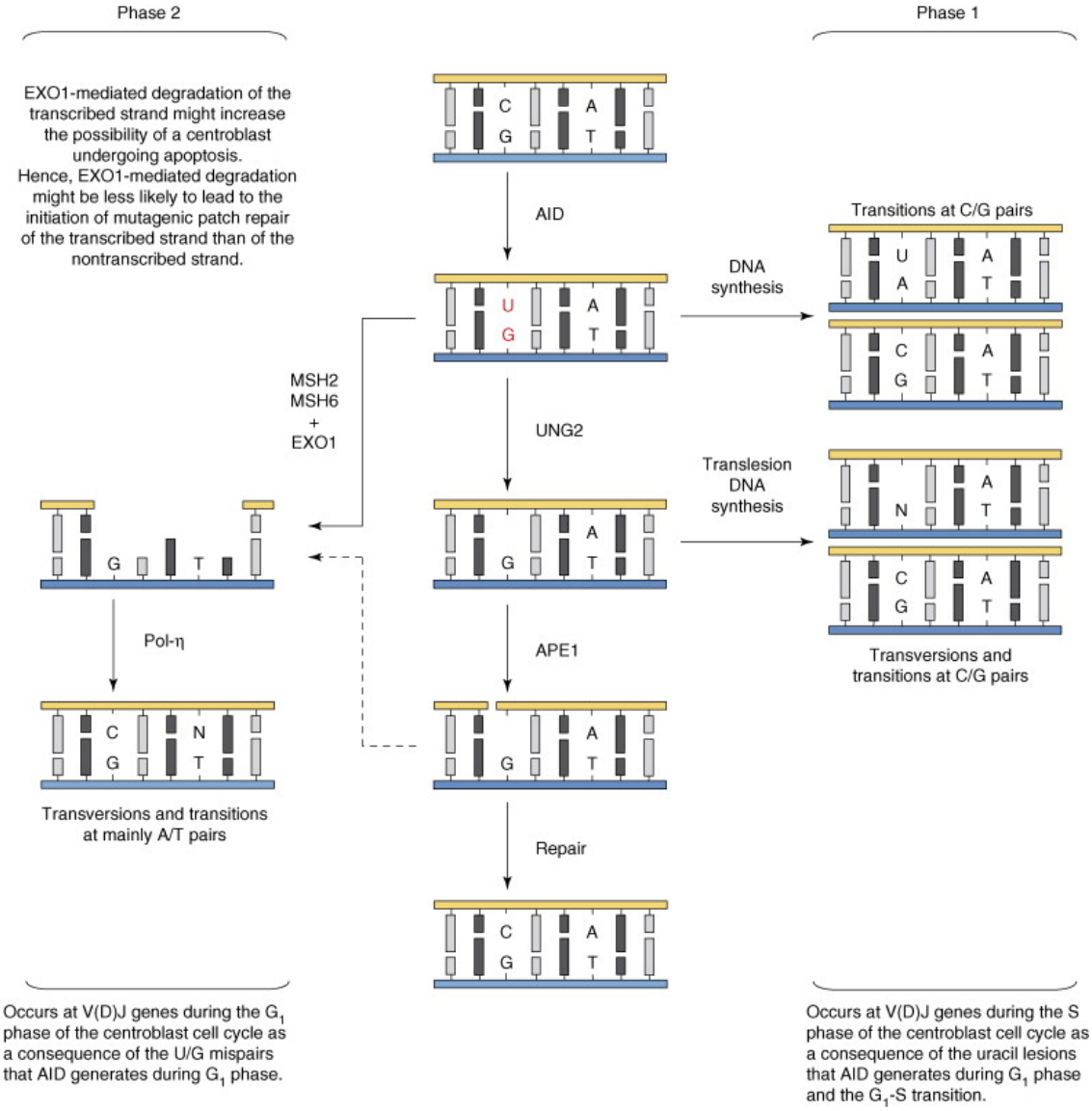

Figure 1.8 - Somatic Hypermutation models. DNA deamination model of immunoglobulin gene diversification, emphasizing Somatic Hypermutation and indicating some of the key enzymes implicated in each pathway. AID deaminates the cytosine generating a mismatch in the primary sequence. The alteration can be processed by base excision Repair (BER): UNG removes the uracil and the resulting abasic site is recognized by DNA polymerase β which repairs the lesion. The mutagenic profile can be acquired when replication occurs after deamination over the dU:dG mismatch or after UNG base removal. Moreover, the dU:dG mismatch can be processed by the Mismatch Repair Pathway (MMR) which leads the introduction of mutations biased towards A:T basepairs. Modified from {Di Noia, 2007}.

depending from DNA polymerase involved. Transversions seem to be mediated by, deoxycytidyl transferase Rev1 polymerase {Nelson, 1996; Zan, 2012} that prefers the insertion of cytidines during the synthesis of the new DNA leading to C to G or G to C mutations {Petersen-Mahrt, 2002; Seki, 2005; Jansen, 2006}.

In Phase 2 the dU:dG mismatch is detected by Msh2/Msh6, which belong to the post-replicative DNA repair system, mismatch repair (MMR) {Frey, 1998; Rada, 1998; Phung, 1998; Martomo, 2005} (**Figure 1.8**).
Phase two occurs during G1 phase and the mutational pattern involves A:T pairs; *pol-η* knock-out mice show an accumulation of mutation at A:T pairs. This suggest that DNA polymerase-η is involved in the repair of the AID-induced lesion. {Delbos, 2005; Di Noia, 2007).

AID in Class Switch Recombination (CSR)

CSR is a multi-step process that occurs in germinal center B-cells, promoting the switch of the antibody isotype, from IgM to a different immunoglobulin, such as IgA, IgG or IgE, maintaining the same antigen specificity.
Evolutionarily, the CSR process is the last acquisition in the development of the adaptive immune system. It is present in all tetrapods and absent in cartilaginous and bony fishes {Barreto, 2005}.
Considering that AID appears at the beginning of the vertebrate radiation, in cartilaginous and bony fish in which CSR is not present, it is intriguing that fish AID can be used to induce CSR in B-cells from *AID*$^{-/-}$ mice {Barreto, 2005}. This suggests that even though the CSR process evolved at a later stage, all the machinery to exploit the system was already present.

CSR is initiated by the introduction of uracil into the DNA as a result of the AID-mediated deamination in the Switch regions. These repetitive regions are rich in mutational hotspots for AID and are actively transcribed during the cellular programme that begins with the activation of the naïve B-cells {Chaudhuri, 2003; Chaudhuri, 2004}. It is hypothesized that, due to the high density of mutational hotspots, AID is recruited and deaminates both DNA strands. The processing of the AID-induced deamination by UNG generates

an abasic site, which is further process to generate a nick in the DNA by the activity of the AP endonuclease {Doetsch, 1990; Mol, 2000; Xu, 2014}. Overlapping deaminations close enough on both strands could therefore generate DSBs {Stavnezer, 2008}.

The DSB lesions on separate S regions is then repaired through Non-homologous end-joining pathways (NHEJ) {Kotnis, 2009}. While PARP1 or LIG3 - genes involved in a-NHEJ pathways - promote the translocation {Wray, 2003; Soni, 2014}, B cells lack for XRCC2, a gene involved in HR pathways, show an increase in the DNA breaks; thus suggest that HR could mediate the correct restoration of sequence during the immunoglobulin rearrangement, and induces a decrease in the efficiency of CSR {Hasham, 2010} (**FIgure 1.9**).

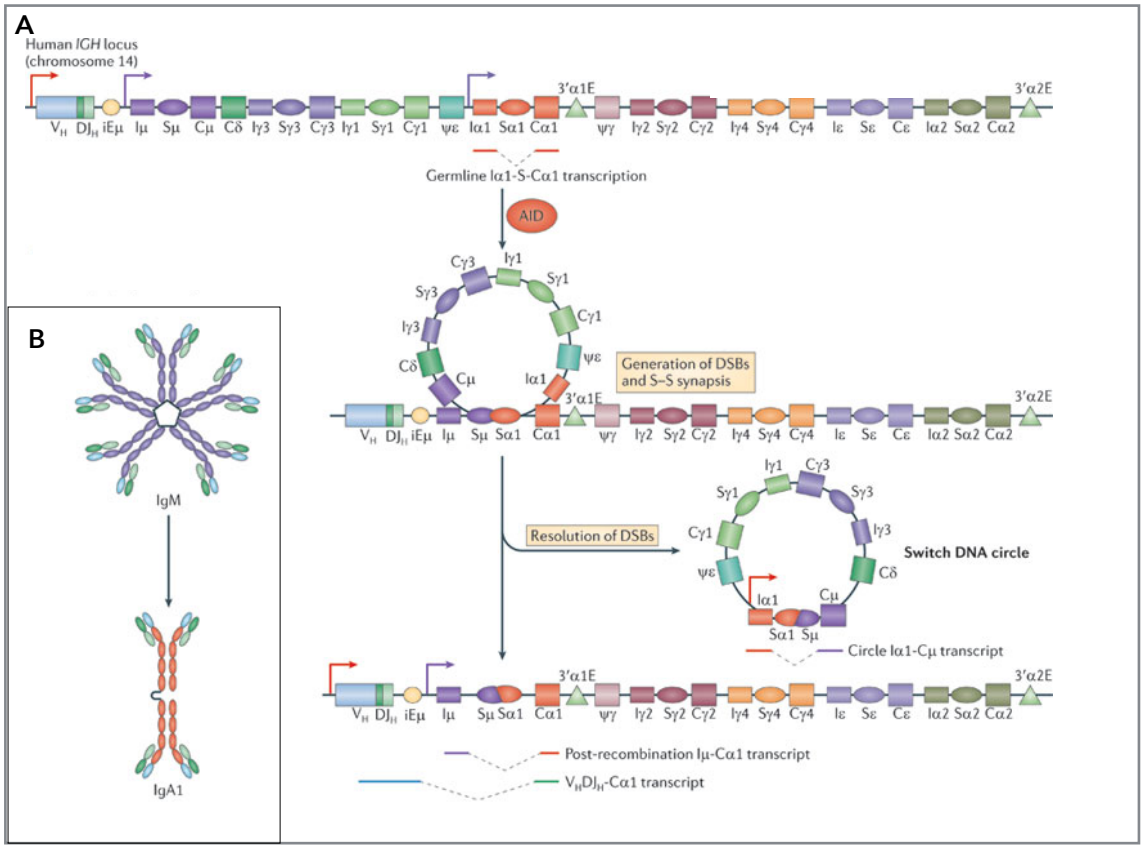

Figure 1.9 - Schematic representation of Class Switch Recombination mechanism. A) AID triggers the deletional-recombination inducing a DSBs at S regions of immunoglobulin locus followed by the action of the repair pathways. This allows the rearrangement of C region of heavy chain and the deletion of the intervening sequence. **B)** Representation shows the structural modification from IgM to IgA protein ad a result of the CSR.

In fact the possibility to be a perfect substrate for AID, the induction of a controlled DBS followed by the capability to obtain a chromosome rearrangement able to produce a functional new antibody, is solely attributable to the sequence of the Switch regions.

Recruitment of AID

Recruitment of AID to the immunoglobulin loci during SHM and CSR is highly regulated to prevent uncontrolled mutagenic activity of AID on non-Ig genes that could culminate with the onset of the genome alterations.
Despite an overall balance between B- and T- cells in the blood, 95% of lymphomas originate from B-cells. This paradox is explained by the molecular processes specific to the B-cell lineage {Küppers, 2005; Alt, 2013}. AID physiologically targets the Ig heavy and light chain loci, yet it has been shown to mutate non-Ig genes {Robbiani, 2008; Robbiani, 2009; Chiarle, 2011; Klein, 2011Hakim, 2012;}. Indeed the particular mutational status of some of the oncogenes involved in B-cell tumors had been identified even before the identification of AID as a DNA mutator. AID mutational pattern could be thus identified in oncogenes such as BCL6, MYC, PAX5, MIR142, CD95 and BCL7 {Pasqualucci, 1998; Shen, 1998; Müschen, 2000; Hakim, 2012}. Yet, all these off-target genes were targeted at a much lower frequency than the Ig loci {Liu 2008}. Moreover, the damage induced by AID on other genomic loci could lead to over-expression of proto-oncogenes through onset of chromosomal translocation that juxtaposes them to the *Ig* enhancer {Nussenzweig and Nussenzweig, 2010}.
Studies revealed that AID action on *Ig* gene seems to be facilitated by at least three related factor. The first one is the presence of enhancers, required for the hypermutation and the recombination of variable and Switch regions {Buerstedde, 2014}. The second one is the pausing of RNA PolII during transcription process {Wang, 2009}, which allows the hypermutation and recombination through the interaction of Spt5, a PolII stalling factor, with AID {Pavri, 2010}. The third element is the role of the RNA exosome complex, which enhances the activity of AID on the template strand by freeing the DNA strand of the RNA transcript and thus making it accessible to AID {Basu, 2011}.

Studies in murine B-cells and human lymphomas show that the predominant characteristic shared by the targets of AID is their clustering within RNA factories in which are transcribed genes with highly active super-enhancers (SEs) and regulatory clusters {Meng, 2014; Qian, 2014}. Super-enhancers are identified as a special subgroup of regulatory elements {Whyte, 2013} normally associated with highly transcribed genes; alteration on these enhancers can lead to the onset of disease {Hnisz, 2013; Parker, 2013}.

Homologous Recombination and Non-homologous end-joining: two pathways to repair DBSs

The number of DNA lesions estimated to occur every day in a single cell is in the order of tens of thousands. This is why organisms have developed an array of reliable DNA repair systems that are fundamental to preserve the genetic information. {Krejci, 2012}.
Genetic alterations that involve breaks in the DNA strand leads to the activation of two different repair pathways: Homologous Recombination (HR) and the Non-homologous end joining (NHEJ) (**Figure 1.10**).

Homologous Recombination (HR) is necessary to accurately repair breaks on DNA strands and prevent loss of genetic information. It plays a crucial role in the crossing over during meiosis, thus permitting the exchange of genetic material between two homologous chromosomes each derived from parental gametes, and during mitosis, by permitting the correct replication of the genome. This high fidelity repair pathway is a post-replicative event which uses the homologous chromosome as a template to restore the break {Krejci, 2012}. Gene targeting, a fundamental technique for genetic engineering based on the HR system was developed thanks to the studies on Homologous Recombination. The importance of these studies and of the gene targeting technique is highlighted by the Nobel prize shared in 2007 by Capecchi, Smithies and Evans.

The other pathway involved in repair of DSBs is the Non-homologous end joining (NHEJ), which joins the broken DNA ends without the need of a donor homologous sequence. This repair pathways is active throughout the cell cycle, especially in G1 due to the lack of sister chromatids. This form of repair is not as accurate as HR, in fact the onset of chromosomal translocations in

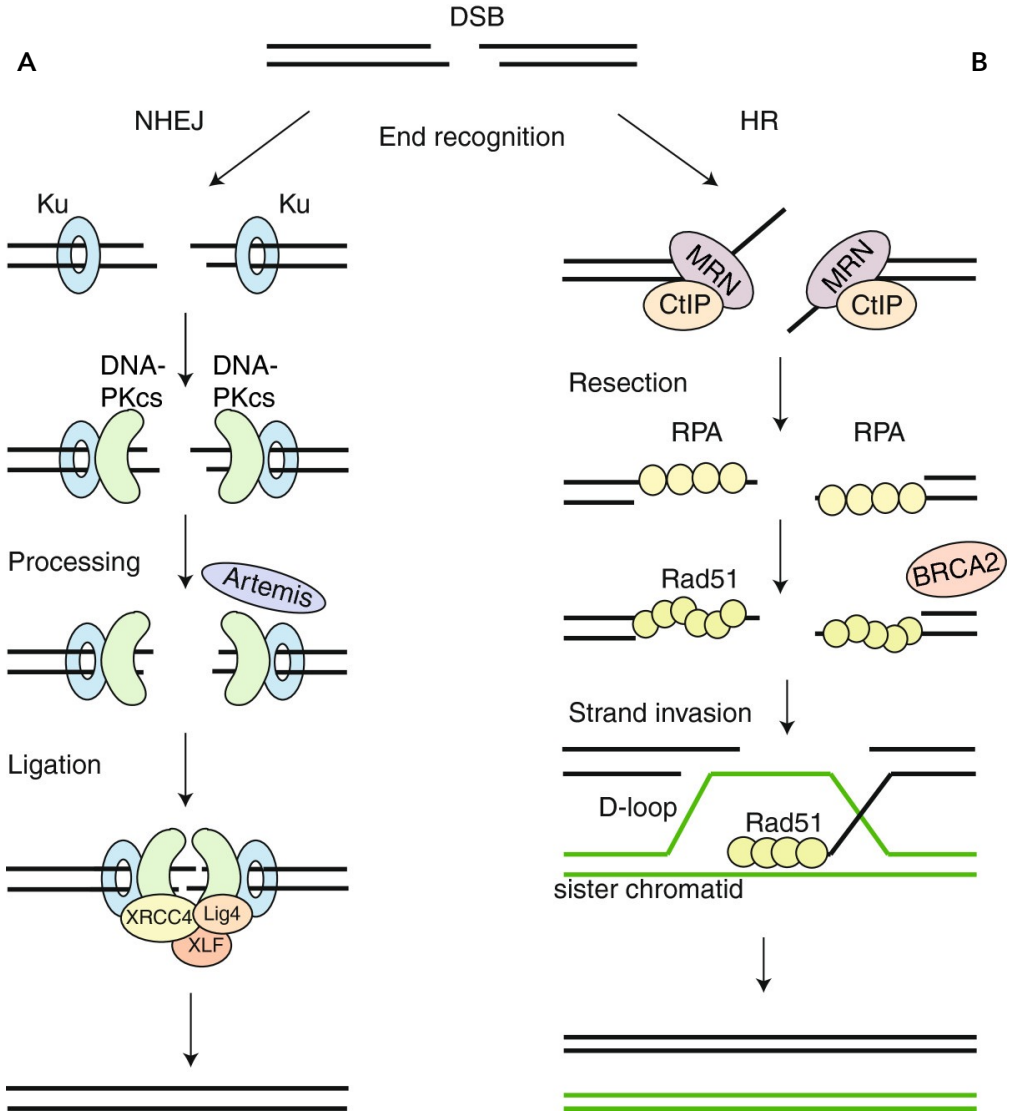

Figure 1.10 - Schematic diagram for NHEJ and HR. The double strand break can be repaired by two different systems depending on various factors, phase of the cell cycle and entity of lesion among them. **A)** NHEJ pathways starts with recognition of the DNA ends by the Ku70/80 heterodimer, which recruits the DNA-PKcs. Nucleases, such as Artemis, trim the ends. and the XRCC4-DNA Ligase IV-XLF ligation complex joins the break. **B)** HR pathway begins with the MRN-CtIP-complex, which resects the breaks to generate single stranded DNA (ssDNA) overhangs. The ssDNA is covered by RPA and subsequently replaced by Rad51 with the help of BRCA2. Rad51 nucleoprotein filaments mediate strand invasion on the homologous chromosome - used as template - and the broken strand is resynthesized. {Brandsma, 2012}.

cancer has been correlated to events repaired through NHEJ {Karanam, 2012}.

The repair of DSBs in CSR

Classical- and alternative-NHEJ (c- and a-NHEJ) repair pathways are used by the cell to solve the DSBs when the sister chromatid is not available. NHEJ pathways join DNA ends without any homology and are active during specific phases of the cell cycle {Rothkamm, 2003}. C-NHEJ is the preferred repair process during the G1 phase when the chromatids are not available. Studies revealed its fundamental role for the repair of DSBs induced by ionizing radiation {Lieber, MR, 2010} or generated during the development of the B- and T-cells. It is the physiological pathway used during the SHM on V(D)J region of B receptor (BCR) and during CSR in B-cells {Stavnezer, 2008}. Mutations in the key elements of c-NHEJ, such as the Ku70/Ku80 heterodimer or the DNA ligase complex XRCC4/ligase-4 (LIG4), reduce the level of CSR {Casellas, 1998; Pan-Hammarström, 2005}.

B-cells deficient for c-NHEJ components are still capable level of level of CSR and sequence analysis, on the rearranged Switch regions revealed a preference for microhomology at the DNA ends junctions {Yan 2007; Boboila, 2010}. This repair pathway is not correlated with c-NHEJ, it is called microhomology-mediated-EJ, alternative end- joining (A-EJ) or alternative-NHEJ {Guirouilh-Barbat, 2007}. Thus, even if the c-NHEJ pathway joins DNA breaks with short (1-3 nt) or no homology, the a-NHEJ prefers to join DNA ends which present a microhomology (4-20 nt). Actually, it is still unclear if a-NHEJ is one or more pathways, but it is accepted its role in repair DBSs during CSR. Studies reported the importance of the role of a-NHEJ during the development of mature B-cells; B-cells deficient in C-terminal binding protein interacting protein (CtIP), a protein responsible for the resection step of the DNA ends from DSBs, which is normally associated with HR {Yun, 2009}, and proficient for c-NHEJ pathway show a considerable reduction in CSR level. {Lee-Theilen, 2011}. During CSR both c- and a-NHEJ pathways (**Figure 1.11**) are active at the same time and they are used by the cell to correct different types of DBSs at the Switch regions. The nature of the DSBs depends by the

density of the deaminated cytosines, due to the action of AID on opposite strands. This defines which repair pathway is more efficient for any given DSB event {Cortizas, 2013}.

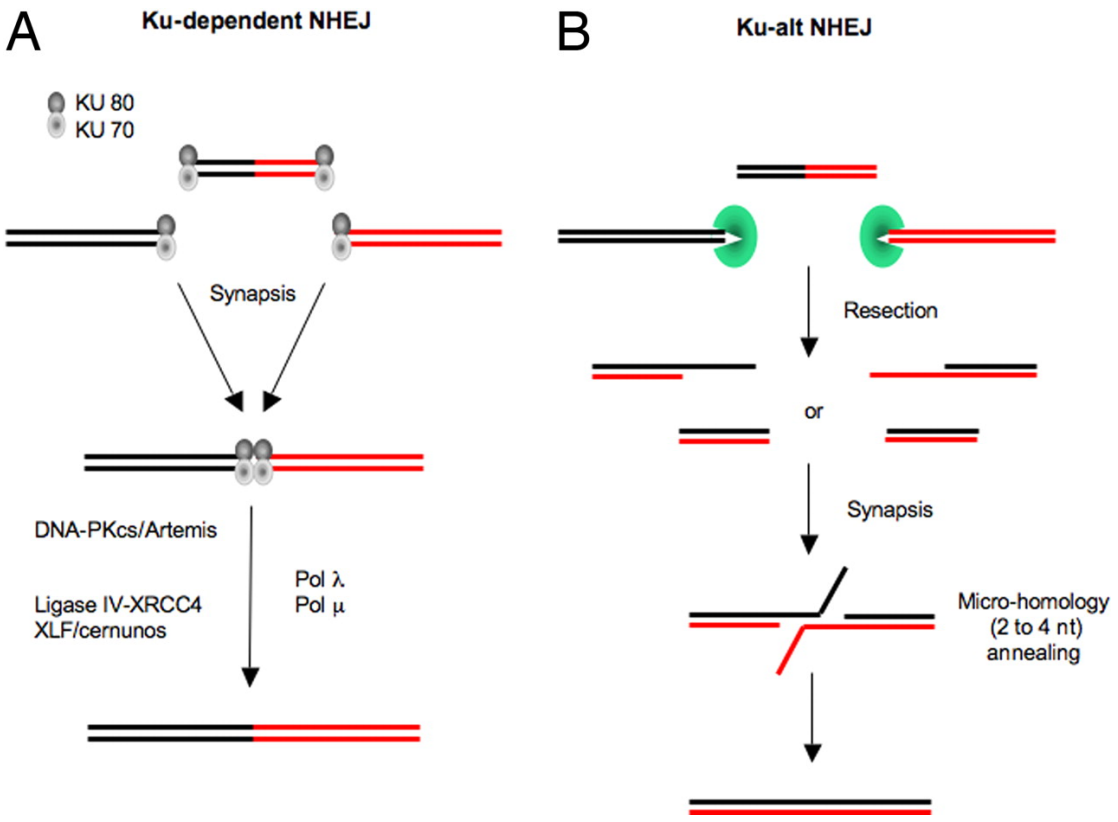

Figure 1.11 - Schematic mechanism of classical- and alternative-NHEJ. A) c-NHEJ restores the break through the Ku70/Ku80 and XRCC4. This repair pathway joins DNA ends without homology. **B)** a-NHEJ Ku-indipendent pathway exploits the internal microhomologies distant from the ends. to repair the damage {Guirouilh-Barbat, 2007}.

Chromosomal Translocations

Translocations arise from the combination of two factors: DNA double-strand breaks (DBSs) and a trans-chromosomal ligation of the broken strands. Endogenous and exogenous mutagens agents, such as metabolism products like oxygen, chemical agents, radiations or ultraviolet light (UV) may trigger the onset of breaks on DNA strand.

This genomic alteration can result in the overexpression of oncogenes or in generation of novel fusions genes {Jasin, 2000; Mills, 2003}.
The overexpression arises when erroneous rearrangement juxtaposes a strong promoter in proximity of the coding sequence of a gene, such as happens in LDPs, in which the expression of proto-oncogenes (BCL2, BCL6, MYC) is controlled by IgH intronic enhancer {Robbiani, 2008; Lu, 2013; Sungalee, 2014}. This conditions can lead to the onset of tumors as happen in blood cancers. On the other hand, the translocation can promote the creation of a fusion gene, which encodes for a novel protein, directly implicated in the origin of disorder. For example, the Philadelphia (Ph) chromosome (or translocation) arises from the reciprocal translocation between chromosome 9 and 22, and leads to the formation of a fusion gene BCR-ABL, which is strictly associates with Chronic Myeloid Leukaemia (CML) {Soverini, 2014}. There a lot of examples of deleterious translocations, most of them recorded in the Mittelmann database {Mitelman, 2015}.

AID and B-cell tumors

AID plays a crucial role during B-cell development triggering the SHM and CSR processes; AID is an efficient DNA mutator strictly controlled and regulated by many factors. Its regulation is very important because its mutagenic activity links immunity, chronic inflammation and genomic instability resulting, in some cases, in tumorigenesis {Pasqualucci, 2001; Kim, 2007; Marusawa, 2008; Endo, 2008}.

In the past years it has been demonstrated that AID can act as a genome-wide somatic mutator on non-Ig genes such as proto-oncogenes, tumor suppressors, genes involved in genome stability maintenance or pluripotency genes. These lesions are hallmarks of blood cancers: for example in more than 50% of diffuse-large B-cell lymphomas, proto-oncogenes such as PIM1, MYC, PAX5 are mutated, and the mutation pattern observed in these tumors correspond to that found on Ig genes during SHM and CSR {Wang, 2004; Dorsett, 2007; Robbiani, 2008; Robbiani 2009; White, 2011; Greisman, 2012}.

Genomic instability in often associated with a characteristic translocation that arise during the SHM or CSR; studies shown how these rearrangements are directly connect with AID action on Ig and non-Ig genes {Klein, 2011}. The nature of the translocation normally is cell specific and permits to define the molecular features of the disorder, used as a marker for prognosis and as indicator for therapy {Kersey, 1997; Biondi, 2000}.

Diffuse large B-cell lymphoma (DLBL), a non-Hodgkin's lymphoma (NHL), is characterized by the presence of different chromosomal translocation, which often involve a translocation between the Immunoglobulin (Ig) loci and the BCL6 gene (**Figure 1.12**). *BCL6* plays an important role during germinal-center formation through its capability to suppress p53 gene and regulate the apoptosis during the CSR {Phan, 2004}. The rearrangement brings the Ig promoter in the upstream region of *BCL6* {Lu, 2013} inducing an over expression that is necessary for the cancer development. The over expression of *BCL6* acts as a negative regulator of apoptosis blocking the cell death and leading to the accumulation of DNA lesions.

Follicular lymphoma (FL) is a disease that affects follicle centre B-cells; with respects to DLBL, it presents a characteristic reciprocal translocation t(14;18), in approximately 80-90% of cases. This translocation involves the IgH locus and the proto-oncogene BCL2 (**Figure 1.12**). Sequence analysis of *BCL2* reveals similarities with the mutational pattern induced by AID during SHM;

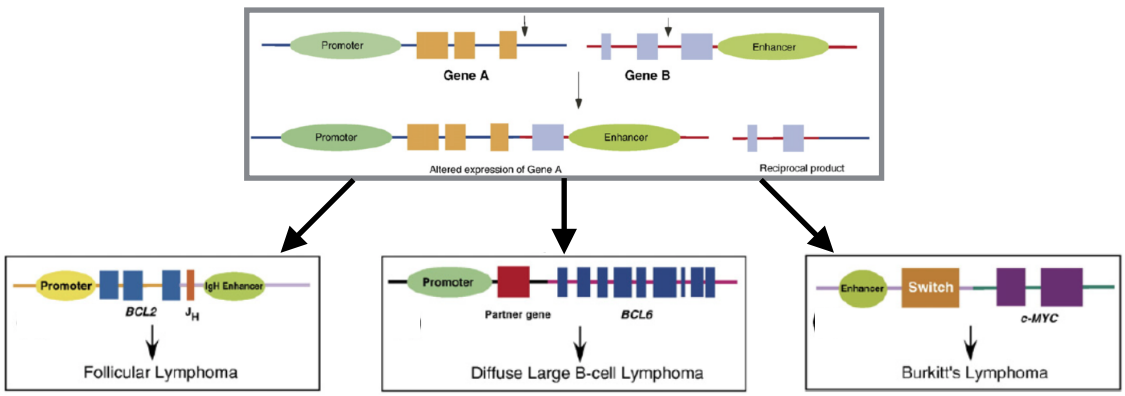

Figure 1.12 - Schematic representation of common translocations in LPDs. A) The follicular lymphomas (FL) are characterized by the presence of *BCL2/IgH* translocation; this normally occurs during the SHM process and involve the V(D)J region of the immunoglobulins. **B)** Diffuse large B-cell lymphomas (DLBL) are characterized by translocations that involve BCL6 gene. **C)** The most common translocation in Burkitt's lymphomas involve the immunoglobulin loci and the MYC gene; this kind of translocation is directly connected with the CSR process.

normally the amount of lesions has be correlated with the survival of patients affected by FL {Correia, 2014}.
Burkitt's lymphoma (BL) is the most common NHL with an incidence of 70% in children and adolescents, and 30% in patients older than 60 years; most of BL cases are characterized by a reciprocal translocation that arises during CSR between the proto-oncogene MYC and the immunoglobulin Switch regions classified as t(8;14) {Robbiani, 2012; Said, 2014} (**Figure 1.12)**; this alteration can also be found in others LPDs, such as follicular lymphoma, diffuse large B-cell lymphoma and "B-cell lymphoma unclassifiable".

The translocation t(8;14) in Burkitt's lymphoma

Discovered in 1958 by Denis P. Burkitt during his studies on equatorial African's children, the Burkitt's Lymphomas is an uncommon type of Non-Hodgkin Lymphoma (NHL) and its characteristic is an uncontrolled growth of white blood cells in particular B-cell. Similarly to other cancers, this phenotype takes origin from genetic alterations {Nambiar, 2008}. Two different variants of Burkitt's lymphomas have been discovered: the endemic and sporadic variant.
The endemic variant is the most common cancer among children in Africa and normally it is related to the presence Epstein Barr Virus (EBV); in fact the 95% of cases of Burkitt's lymphomas in Africa shows the EBV infection that often is associated with another disease really common in that area, the chronic malaria, that seems to lower the resistance versus the virus.
The sporadic one (also called "non-Africa") is another variant that can be found in Western World; it hasn't a strong correlation with EBV, in fact only 20% of patient affected by sporadic variety shows the presence of the virus.

A characteristic chromosomal alteration associated with Burkitt's lymphoma normally involves the oncogene MYC and one of the immunoglobulin genes encoding for the IgH gene {Erikson, 1982; Dalla Favera, 1982}, Ig kappa {Malcolm 1982} and Ig lambda {de la Chapelle, 1983}. The translocation *MYC/IgH* appears in the 40% BL studied and although it has been defined as a marker of the disease, it has been found in other human LPDs, as DLBCL and

FL. The same translocation was described in pristane or IL-6-induced murine plasmacytomas {Taub, 1982}.

AID deaminates cytosines outside Ig loci and triggers the onset of the translocation {Dorsett, 2007; Robbiani, 2008; Robbiani, 2009; White, 2011; Greisman, 2012}. On the other hand, mice lacking active AID are less susceptible of acquiring the translocation *MYC/IgH* {Gazumyan, 2011}.

Sequence analysis performed on patients with t(8; 14) show the presence of two specific regions (clusters) on the first intron of myc, in which can be found most of the breakpoints involved in the translocation. These two regions, respectively 560 bp and 779 bp in length, appear to have analogies with the immunoglobulin Switch regions. In particular, while the donor region (Sμ) recombines equally with both clusters, Sγ regions associate mostly with cluster 1 breakpoints, and Sα regions with cluster 2 ones (**Figure 1.13**) {Busch, 2007}.
These clusters share a common characteristic with other AID-targeted proto-oncogenes, such as BCL6, RhoH, PIM1 and PAX5 {Pasqualucci, 1998; Liu, 2008}. The instability on these proto-oncogenes does not depend on the density of WRC motifs, but it has been correlated with G-richness sequences. During transcription the G-regions form single-strand structures (G-loops) which contain the targeted sequence by AID {Duquette, 2007}.

The MYC gene was discovered as gene homologous to the retroviral v-myc oncogene {Vennstrom 1982} and, subsequently, it was found to be expressed in different animal and human cancers {Cole MD, 1986}. *MYC* is a transcription factor that promotes or repress the expression of other genes (it is estimated 15% of all human genes) through binding to the Enhancer Box sequence (E-boxes). The expression of *myc* results in various physiological processes and controls genes involved in the cell cycle, in the control of cellular metabolism, signaling, and apoptosis {Dang, 1999; Menssen and Hermeking, 2002}. *MYC* upregulation was found in different tumors, such as lymphoproliferative diseases, lung adenocarcinoma {Little, 1983} , breast {Escot, 1986; Munzel, 1991}, cervical {Sagawa Y, 2001} and ovarian carcinomas {Wang, 1999}. Its misregulation can lead to, or facilitate the onset of different diseases.

MYC gene is located on chromosome 8 in humans and on chromosome 15 in mice and its conservation between humans and mice is 81%. The gene consists of three exons, with the start codon being located at the beginning of exon 2 in humans, and at the end of exon 1 in mice {Dang, 1999}.

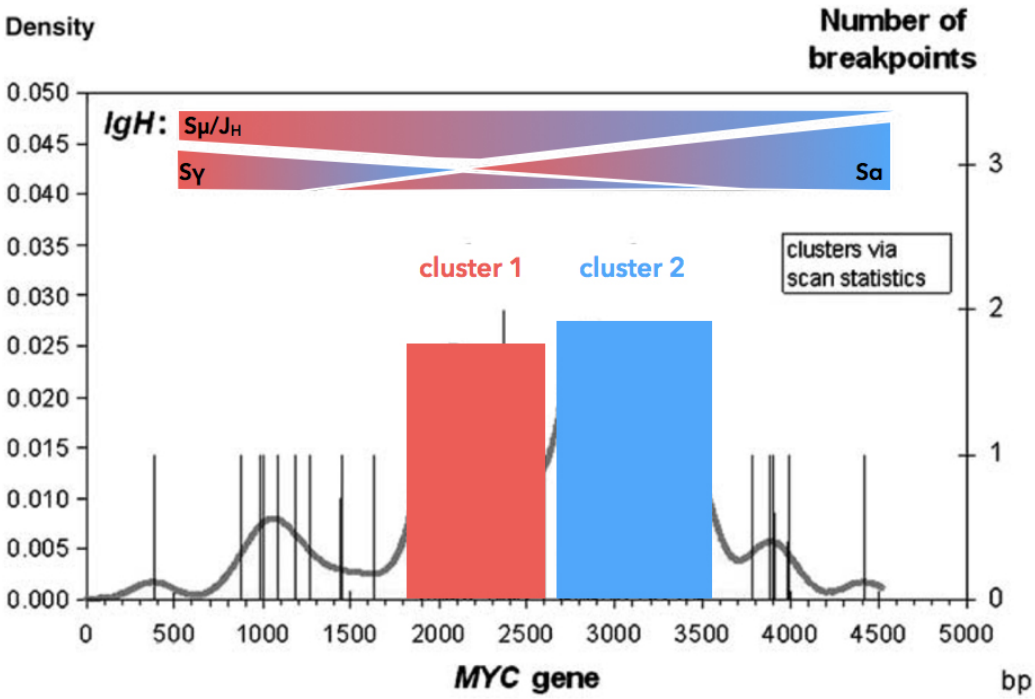

Figure 1.13 - Schematic presentation of the MYC clusters. The clusters region starts from position 381 located 5′ of exon 1 until 4415 of intron 1. The distribution of the breakpoints within the IGH regions are indicated schematically in red and light blue to point out the relevant association with the MYC clusters. Image modified from {Busch, 2007}.

In the past years the commons techniques used to detect the presence of the translocation MYC/IgH was Southern blot and molecular cloning, but these were not suitable to detect chromosomal alterations in a small number of cells and to detect the breakpoints. A new approach, based on PCR, was developed to allow scientists to detect rearrangements in a small fraction of the cells and to investigate which part of the MYC gene was involved in the *MYC/IgH* translocation. {Shiramizu, 1990}. While the regions of the immunoglobulins involved in the translocation were known (S regions), on the other hand the breakpoint regions of c-myc gene were not described. To detect and distinguish different breakpoint regions close or within to c-myc, different primers were designed to bind the 5′ region of c-myc gene at

different position. Thus, this system permitted the detection of translocation t(8;14) in a cell population and the localization of the breakpoint regions of the c-myc locus {Ramiro, 2004}.

Aid outside B-cell

Physiologically AID is expressed in activated B-cells during their development and triggers the molecular mechanism that drive the SHM and CSR.
The expression of AID is controlled and regulated through different layers: a tightly controlled promoter, post-transcriptional regulation by microRNA {Dorsett, 2008; Teng, 2008}, post-translational modifications {McBride, 2008; Cheng 2009} and subcellular localization {Geisberger, 2009; Patenaude, 2009}.
However, mutations ascribed to AID have been identified in tumors which take origin from different cellular type. Moreover, studies revealed the presence of AID in healthy spermatocytes in human testis and oocytes {Morgan, 2004; Schreck, 2006}

On the other hand, Helicobacter pylori infection in gastric epithelium triggers the expression of AID {Matsumoto, 2007}. The comparison of deep sequences analysis of whole-exomes of gastric epithelium mucosal tissue of patients and mice knock-in for TP53 and constitutively expressing AID show that the mutations were predominantly transition C:G to T:A in GpCpX motif; this motif correspond to hotspots of AID found on Ig genes after SHM and CSR {Shimizu, 2014}. Moreover the aberrant expression of AID has been found in human cholangiocarcinoma-derived cells. AID expression is induced by tumor necrosis factor-(TNF-) via IkappaB kinase-dependent nuclear factor (NF)-kappaB pathways. This erroneous expression of AID resulted in the generation of somatic mutations implicated in tumorigenesis {Endo, 2007; Komori, 2008}. Moreover, AID expression is also regulated by estrogens {Pauklin S, 2009}.

Aim of the work

Chromosomal translocations in Lymphoproliferative diseases (LPD) often involve the immunoglobulin locus, and are related to the secondary diversification process specific of B-cells. These erroneous rearrangements can place proto-oncogenes under the control of the intronic super-enhancer of Ig, thus leading to their upregulation.
The *MYC/IgH* translocation is the most common alteration in Burkitt's Lymphomas, and can also be found in murine pristane or IL6–induced plasmacytomas. This genetic alteration is characterized by the presence of well defined breakpoint regions. Moreover it is directly connected to the CSR process, which takes place during the development of B-cells. These unique characteristics make *MYC/IgH* translocation a perfect model for the onset of chromosomal translocations in human cancers.

Many factors - mostly correlated to DNA repair pathways - have been identified as involved in the onset of *MYC/IgH* translocation. Yet, many others have been suggested but not proved yet, and other unknown ones could play an even more prominent role in the onset of this alteration. The experimental systems usually used to study chromosomal rearrangements, PCR amplification and Southern blot, have been used since a long time. Yet, they are not feasible to be used as a screening tool for unknown factors. Moreover their poor sensitivity when studying cellular populations in which only a small part bears the chromosomal translocation makes them difficult to use as tools to investigate the onset of chromosomal rearrangements.

The aim of my PhD was to setup a tool to discover and study factors involved in the onset of the chromosomic translocation *MYC/IgH* during the secondary diversification process.

To this aim we used a murine lymphoma cell line (CH12-F3) in which CSR can be induced through stimulation with TGF-β, IL-4 and CD40 stimulation. *MYC/IgH* translocations have been observed in this cell line at a low frequency {Nakamura, 1996}. We thus replaced the second exon in one of the MYC alleles with the reporter gene Enhanced Green Fluorescent Protein (EGFP),

placed out-of-frame with respects to the wild-type MYC cds. The rationale of our approach is that the onset of the *MYC/IgH* translocation will place the full EGFP cds under the control of the Ig locus, and thus induce the appearance of GFP(+) cells easily detectable by FACS analysis. We have set up the system, and we are now testing its feasibility to assay the role of specific factors in the onset of *MYC/IgH* translocation.

In parallel, during my PhD I collaborated with Dr. Cesare Sala e Dr. Silvia Pietrobono to study the effects of AID splice variants in Class Switch Recombination.

A number of splice variants of AID have been identified in B-tumor cells first and in normal B-cell later {Muramatsu, 1999}. The physiological role of these splice isoforms is controversial, as some studies deemed them catalytically active while others did not. We hypothesized that - even in absence of catalytic activity - these splice isoforms might have an effect on the efficiency of the CSR.

We thus tested two of these splice variants, which we find catalytically inactive, for their ability to modulate the activity of endogenous AID in CH12-F3 cells. In contrast to full-length AID, neither these splice variants or a catalytically impaired AID mutant affect the efficiency of CSR. Thus, while a role for these splice variants at the RNA level remains possible, it is unlikely that they, as proteins, exert any regulatory effect on the function of AID.

Chapter II

Materials and Methods

Solution and buffers

Transformation Buffer (TB)

1 liter
Pipes 10 mM
$CaCl_2$ 15 mM
KCl 250 mM
brought to pH 6.7 with KOH
MnCl2 55 mM

2X Cracking Buffer

50 ml
NaOH 1 M (5 ml)
EDTA 0.5 M (1 ml)
10% SDS (5 ml)
Glycerol (5 ml)
ddH_2O to adjust the final volume

5X TBE

1 liter
Tris Base 0.5 M (54 gr)
Boric Acid 0.5 M (27.50 gr)
EDTA 20 mM (4.56 gr)

Loading Buffer (DYE)

1 ml
15 % Xylene Cyanole
15% bromophenol blue
30% glycerol

Sodium phosphate buffer

Na_2HPO_4 1 M (342 ml)
NaH_2PO_4 1 M (158 ml)
ddH_2O to adjust the final volume

Church Buffer

Sodium phosphate buffer 0.25 M
EDTA 1 mM
1% BSA
7% SDS

Depurination Buffer

1 liter
HCl 11 ml
ddH_2O to adjust the final volume

Denaturation Buffer

1 liter
NaCl 1.5 M (87.7 gr)
NaOH 0.5 M (20 gr)

20x SSC

NaCl 3M
Trisodium Citrate 0.3 M
HCL 1 M to adjust pH to 7.2

Bacteria

Bacterial strains		
Bacterial Strain	**Genotype**	**Info**
DH5α	F-φ80d/*lac*ZΔM15Δ (*lacZYA -argF)* U169 endA1 rec A1 *sd*R17 (r_K^{-m} K^+) *deo*R *thi* 1 *sup*E44 λ-*gyr*A96 *rel*A1	host strain for general cloning Invitrogen (1986)
Top 10F	F'{*lac*Iq Tn10 (TetR)} mcrA Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal/U, gal/K* *rpsL end*A1 *nupG*	host strain for high efficiency cloning Invitrogen, US Yao Sylvia, 2002
Stable3	F- glnV44 recA13 mcrB mrr hsdS20(rB-, mB-) ara-14 galK2 lacY1 proA2 rpsL20 xyl-5 leu mtl-1.	host strain for retroviral/lentiviral cloning Invitrogen, US

Bacteria Medium	
Medium	**Composition**
Luria - Bertani (LB)	bacto tryptone 2% yeast extract 0.5% NaCl 10 mM KCl 2.5 mM $MgCl_2$ 10 mM
SOB	LB $MgSO_4$ 10%
SOC	SOB Glucose 1 M
Freezing medium	LB 40% glycerol

Bacteria Antibiotics		
Antibiotics	**Concentration**	**Cell Type**
Ampicillin	50-100 µg/ml	bacteria
Kanamycin	50 µg/ml	bacteria

Bacteria Techniques

Preparation of chemically competent bacteria

Day 1:
inoculate overnight bacteria culture in 5 ml of LB media.
Day 2:
Inoculate 1 ml from the overnight culture in 250 ml SOB and let the cells grow into the shaker (200rpm) at 37°C until OD_{600}=0.6 (lecture made with a Biophotometer Eppendorf spectrometer). The growing time depends on the bacterial strain (around 2 hours and half for DH5α).
When the culture reaches the correct OD, cool the bacteria on ice 10 minutes and spin them down at 4°C, 2500 g for 10 minutes. From now the cells must be kept in ice for all the next steps.
Remove the supernatant and resuspend the pellet in 80 ml of cold TB; store the bacteria on ice for 30 minutes and then pelleted at 4°C 2500g for 10 minutes. Remove the supernatant and resuspend the pellet in in 20 ml of cold TB and keep the cells on ice 15 minutes. Add DMSO up to 7% final concentration and after 5 minutes; aliquot the competent bacteria in 1.5 ml tubes, snap-freeze them in liquid nitrogen, and store at -80°C.

Transformation of chemically competent bacteria

Day 1:
Thaw chemically competent bacteria on ice and add up to 100ng of plasmid DNA and wait 30 minute. Heat shock the bacteria at 42°C for 45 seconds and immediately cool the samples on ice for 2 minutes. Add 1 ml of SOC media and culture the cells for at least 1h at 37°C, to recover and to express the antibiotic resistance present on the plasmid. Gently spin down the bacteria, resuspend the pellet in ~100 µl of SOC and plate them on plates containing the appropriate antibiotic. Place the plates for overnight growth at 37°C.
Day 2:
Pick the colonies and inoculate them for the plasmid extraction.

Storage of bacterial strains

An equal amount of freezing medium 2X (LB + 40% glycerol) was directly added to the bacterial culture. The culteres were then stored at -80°C.

Cracking System: a massively screening for bacteria transformation

For screening purposes, picked colonies are resuspended in 10 µl of sterile 0.9% NaCl; transfer 5 µl to 200 µl of LB media for the maintenance of the bacteria; mix the others 5 µl with cracking buffer 2X to lyse the cells. Load directly the lysate into agarose gel for electrophoresis run.

Eukaryotic Cells

Cell Lines

Name	Derivation	Reference
CH12-F3	a mouse lymphoma B cell line	Honjo et al., 1996
Hek 293T	human embryionic kidney cell line	

Cell Medium

Cell line medium	Composition	Selection
CH12-F3	RPMI 500 ml 10% of fetal bovine serum (FBS) 1% of Pen/Strep 100x (50 U/ml penicillin and streptomycin sulphate) 1% of L-Glutamine (200 mM) β- mercaptoethanol. 0.1 mM	puromycin 0.6 µg/ml blasticidin 30-50 ng/ml
HEK293T	DMEM 500 ml 10% of fetal bovine serum (FBS) 1% of Pen/Strep 100x (50 U/ml penicillin and streptomycin sulphate) 1% of L-Glutamine (200 mM)	puromycin 1.8 µg/ml
Freezing medium	90% FBS 10% DMSO	

Cell Antibodies

Name	Source	Feature(s) and usage	Concentration
Anti-mouse IgA-RPE	Southern blot Biotech	CH12-F3 switch analysis	1:100
Anti-mouse IgM-RPE	Bethyl	CH12-F3 switch analysis	1:100

Cellular techniques

Maintenance of cells

When adherent or suspended cell cultures reach confluence, the cells are split into new flasks containing fresh medium and incubated in 37°C incubators with 5% CO2. Cell counting was performed using a Neubauer haemocytometer counting chamber.

Freezing of cells

Spin down a healthy logarithmic culture (3 to 5x10^6 cells) at 1,150 rpm for 3 minutes; aspirate the supernatant and resuspend the pellet in 1 ml of sterile, ice cold freezing medium. Transfer them into cryogenic tubes and freeze at -80°C.
For long-term storage, the cryogenic tubes are stored into liquid nitrogen.

Thawing of the cells

Warm the cryogenic tubes at 37°C in a water-bath until the freezing medium is thawed. Transfer the cells into a 15-ml falcon tube with 10 ml of warm culture medium and spin down the cells at 1,150 rpm for 3 minutes; resuspend the pellet in warm tissue culture medium and transfer the culture into an appropriate flask or a Petri-dish.

Transfection of eukaryotic cells

Transient transfections of HEK293T cells:
Lipofectamine 2000 (Invitrogen) or Gene juice (Novagen) were used following the procedures recommended by the manufacturer.

Stable transfection of CH12-F3
Day 1:
Wash 20x10^6 cells and spin them down at 1,150 rpm for 3 minutes; aspirate the supernatant and resuspend in 0,5 ml sterile PBS and transfer the cells in 0.4 cm electroporator cuvette (Biorad) with 20 µg of linearized plasmid; put

the cuvette on ice for 10 minutes. Electroporate the cells with Gene Pulser II (Biorad) using the following parameters: Voltage = 250V; Capacity = 500 μF; Resistance = ∞; incubate the cuvette on ice for 10 minutes. Transfer the cells into a new flask content warm medium and place them in the incubator for 24 - 48 hours.

Day 2:

Plate the cells into 96 wells F-type (four at least) using selection medium containing the appropriate antibiotic.

When colonies are visible (white specks), pick them and expand them into 24-well plates.

Analysis of Class Switch Recombination

Class Switch Recombination in CH12-F3 cells was induced with TGF-ß (2 ng/ml), IL4 (2 μg/ml) and anti-CD40 antibody (0.5 mg/ml) as described in Nakamura et al. {Nakamura 1996}, in medium containing 30% FBS. After 72 hours in culture CSR was assayed in stimulated cells by FACS using an anti-IgA antibody conjugated with RPE (Southern blot Biotech, Birmingham, AL; 1:100). Flow cytometry analysis was performed on a Accuri C6 flow cytometer with a standard configuration (BD Biosciences, San Jose, CA).

Analysis of the expression levels

The protein levels of AID variants from induced bacteria or transiently transfected HEK293T cells were assayed by western blot analysis after lysis with RIPA buffer and SDS-PAGE. A monoclonal antibody raised against AID N-terminus was used (clone 52-1; 1:1000){Conticello 2008b}. ß-Actin was used as loading control (monoclonal Ab, clone AC-15, lot 117K4873, Sigma, St Louis, MI, USA; 1:10000).

Quantitative real-time PCR (qPCR) was carried out to analyze the transcription level of endogenous and exogenous AID in CH12-F3 cells. Total RNA was isolated with TriPure Isolation Reagent (Roche Diagnostics) and subjected to DNase I treatment (Roche Diagnostics). Reverse transcription was performed with High Capacity cDNA Reverse Transcription Kit (Life Technologies, Paisley, UK). The qPCR reactions were performed at 60°C using SsoFastTM EvaGreen

Supermix (Bio-Rad, Hercules, CA, USA) on a Rotorgene-Q (Qiagen) and analyzed using β-Actin as housekeeping gene.

CRE-Lox system

Cre recombinase is an enzyme that triggers the recombination between two loxP sites. These sites have a specific sequence with 8 bp core sequence, that represents the recombination site, flanked by two 13 bp inverted repeats.
CH12-F3 cells were plated at a density of 3-5x10^5/ml and starved overnight. Then we added TAT-CRE (100 µg/ml) and after 1 hour we added chloroquine (100µM). The cells were incubated for 30 minute; after washing we plated the cells in fresh media in a limited dilution in order to obtain single clones to be tested for the success of the CRE treatment. In our case, we checked the loss of the blasticidin resistance conferred by the antibiotic cassette. The TAT-CRE utilized in our experiments was a gift from Dr. Svend Petersen-Mahrt.

Molecular Biology techniques

PCR

Polymerase Chain Reaction (PCR) protocols

AB-Taq			
Reagents		**Quantity/Volume**	
DNA template		100 ng	
10x AB-taq buffer		5 µl	
dNTPs mix (10 mM each)		1 µl	
$MgCl_2$		1.5 µl	
Primer Forward (100 µM)		0.3 µl	
Primer Reverse (100 µM)		0.3 µl	
AB-taq (AB Analitica)		1 µl	
ddH_2O		to 50 µl	
Cycling Step	**Temperature (°C)**	**Time**	**Cycles (n°)**
Initial denaturation	94	2 min	1
Denaturation	94	35 sec	35
Annealing	depends on primers	40 sec	
Extension	72	1 min/kb	
Final Extension	72	5 min	1
Hold	4	∞	1

Cloned *Pfu* DNA polymerase AD			
Reagents		**Quantity/Volume**	
DNA template		100 ng	
10x Cloned *Pfu* Reaction Buffer		5 µl	
dNTPs mix (25 mM each)		0.4 µl	
Primer Forward (100 µM)		1.25 µl	
Primer Reverse (100 µM)		1.25 µl	
Cloned *Pfu* DNA polymerase		1 µl	
ddH_2O		to 50 µl	
Cycling Step	**Temperature (°C)**	**Time**	**Cycles (n°)**
Initial denaturation	94	1 min	1
Denaturation	94	1 min	25-30
Annealing	T_m primer - 5°C	1 min	
Extension	72	1 - 2 min/kb	
Final Extension	72	10 min	1
Hold	4	∞	1

Kod hot start DNA polymerase reagents (Novagen)			
Reagents		**Quantity/Volume**	
DNA template		50-100 ng	
10x Kod hot start Reaction Buffer		5 µl	
dNTPs mix (2 mM each)		1 µl	
$MgCl_2$ (25 mM)		3 µl	
Primer Forward (100 µM)		0.15 µl	
Primer Reverse (100 µM)		0.15 µl	
Kod hot start DNA polymerase		1 µl	
ddH_2O		to 50 µl	
Cycling Step	**Temperature (°C)**	**Time**	**Cycles (n°)**
Initial denaturation	95	2 min	1
Denaturation	95	20 sec	25
Annealing	depends on primers	10 sec	
Extension	70	15 sec/kb	
Final Extension	70	5 min	1
Hold	4	∞	1

PCR programs

Nested PCR used to detect the knock-in	
Reagents Reaction #1	**Quantity/Volume**
DNA template	depending on DNA polymerase
10x Reaction Buffer	
dNTPs mix (2 mM each)	
$MgCl_2$ (25 mM)	
External Primer Forward (100 µM)	
External Primer Reverse (100 µM)	
DNA polymerase	
ddH_2O	
Reagents Reaction #2	
DNA template from PCR #1	
10x Reaction Buffer	
dNTPs mix (2 mM each)	
$MgCl_2$ (25 mM)	
Internal Primer Forward (100 µM)	
Internal Primer Reverse (100 µM)	
DNA polymerase	
ddH_2O	

Touchdown PCR program to decrease aspecific bands			
Cycling Step	**Temperature (°C)**	**Time**	**Cycles (n°)**
Initial denaturation	94	depending on DNA polymerase	
Denaturation	94		
Annealing	from 65 to 55 (-1° each cycle)		
Extension	72		
Final Extension	72		
Hold	4		

Reverse Transcription (RT-PCR)

For Reverse Transcription PCR (rtPCR), the RNA was extracted from cells using TriPure isolation reagent (Roche), and then purified from eventual DNA contamination using the DNase I recombinant kit (Roche) according to manufacturer's instructions. The retrotranscription was performed using the following PCR set-up and program, using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystem).

Plasmid DNA extraction

Nucleospin plasmid kit (Macherey- Nagel), for miniprep and High pure plasmid maxiprep kit (Invitrogen) for maxiprep according to manufacturer's instructions.

DNA extraction from eukaryotic cells

Wizard Genomic Purification Kit (Promega) was used according to manufacturer's instructions. The DNA obtained is quantified by spectrophotometer and by gel electrophoresis.

Agarose gel electrophoresis

For analysis and purification of DNA fragments.
Dissolve 1 gr of agarose into 100 ml of 1x TBE to obtain a 1% gel and add 1 µl/ml ethidium bromide (1.0 µg/ml) or SYBR green and pour the gel into an appropriate slide and wait until it solidifies. Mix 1-5 µl of DNA sample with loading dye (section 2.1). The samples and the DNA marker (100-plus or 1 kb ladder - New England Biolabs) are then on the gel, and apply current to run the electrophoresis. After the run, place the gel on UV light box to be able to visualized the bands. The ethidium bromide binds the DNA and when absorbs UV radiation (245 nm) is able to re-emit light at 590 nm (380-750 nm visible spectrum).

DNA fragment purification from gel

DNA purification from gel was performed using Nucleospin Extract II kit (Macherey-Nagel), following manufacturer's instructions. The obtained DNA was quantified by spectrophotometer or gel run.

Dephosphorylation and Ligation

Dephosphorylation and Ligation of plasmids and inserts was performed with Rapid DNA Dephos & Ligation Kit, (Roche) following manufacturer's instructions.

TOPO cloning kit

TOPO TA® cloning kit (Invitrogen, US) was used to clone directly PCR fragments. Both TOPO plasmids (see table above) carry the lacZ gene for Blue-white screening.

DNA sequencing

DNA sequencing service was performed by the following companies:
PRIMM Biotech Services (www.primmbiotech.com) and Macrogen (www.macrogen.com) for Sanger method sequencing.
The identity and integrity of sequenced fragments were confirmed by aligning them to the ENSEMBL sequence database by nucleotide-nucleotide BLAST search (http://www.ncbi.nlm.nih.gov/BLAST). Putative mutations were checked against the original chromatogram.

Statistical analysis

All statistical analyses were performed using Prism (Graphpad, La Jolla, CA USA). The Kruskal Wallis Test with *post hoc* Dunn's multiple comparison test was used for the rifampicin assay. One-way ANOVA with Tukey's multiple comparison test was performed to analyze the Class Switch Recombination.

Southern Blot

In this work we used southern blot analysis (Southern, 1975) to screen clones for gene targeting. The protocol used is as follows:

Day 1

Digest 10 µg of genomic DNA, extracted by singles clones with an appropriate restriction enzyme. Run the samples on agarose gel at 80 V for 5-6 hours. After the run a photo of the gel was taken to keep track of the running distances of all marker bands. The gel was then incubated with gentle agitation in Depurination Buffer for 15-30 min. The gel was then incubated with Denaturation Buffer for other 30 min. The gel was then prepared for the blotting to the "tower" to transfer the DNA from the gel to nylon membrane overnight.

Day 2

After disassembling the tower, the position of the wells on the membrane were marked and the membrane was cooked in a 3MM paper envelope at 80°C for 2 hours. After a wash in 2X SSC buffer the membrane was prehybridized with 15-20 ml of Church buffer at 65°C in the hybridizer (Techne) for 1 hour.

The probe was prepared by PCR amplification using primers 43 and 44 on CH12-F3 genome and radiolabeled with ^{32}PCTP Amersham Ready-To-Go DNA Labelling Beads (-dCTP) (GE Healtcare). After purification with ProbeQuant G-50 MicroColumns (GE Healtcare) to eliminate not incorporated radioactive the probe was denatured.

The prehybridization buffer was substituted with new Church buffer and the radioactive probe was added. The hybridisation was performed overnight at 65°C.

Day 3

The radioactive buffer was removed and the membrane was washed:

2 washes with 2X SSC + 0,1% SDS for 10 min each

2 washes with 0.1X SSC + SDS 0.1% SDS for 20-30 min each.

Excess buffer was removed from the membrane using 3MM paper, and the membrane was exposed in the phosphoimager cassette after being enveloped in plastic film. The radioactive bands were detected with a phosphoimager Typhoon FLA 7000 (GE).

Genome editing

TALENs

TALEs and TALENs plasmids were prepared using the Golden Gate TALEN and TAL Effector Kit 2.0. The kit was a gift from Daniel Voytas and Adam Bogdanove (Addgene kit #1000000024). To obtain different TALEs chimaera, we cloned the DNA binding domains into different final vectors. For TALENs we used as a final plasmid pCAGGS-TAL-NC2 (a present from Takashi Yamamoto - Addgene plasmid #43856).

TALEs targets	Sequence
MYC_LA1	CAACGTGAACTTCACC
MYC_LA2	GTACGGAGTCGTAGTCG
MYC_RA1	GAAAAGGACTTATTGA
MYC_RA2	TCGAGTTTGTGTTTCA
PARP1	TGCACAGGCGCAGTCGCG
LIG3	TAGGATCACTGCAATGAT

CRISPR/Cas9

The CRISPRs vectors used in this work were obtained by insertion of oligos (Primers Table) into the backbone plasmid pX330. This plasmid is a gift form Feng Zhang (Addgene plasmid #42230).

Plasmids construction

pAIDS3-ΔC: phAID-ΔC was digested with NheI/BsrgI to remove the hAID-ΔC +IRES. To amplify (1100 bp) we used GM84_hAIDS3_3F and GM83_IRES_R primers and the PCR product was digested NheI/BamHI. We obtained the mcherry cDNA by digestion of pMcherry vector with BamHI/BsrGI (700 bp). We assembled a ligation containing pBuerstedde-NheI/BsrGI, hAIDS3-ΔC +IRES-NheI/BamHI and mCherry gene/BamHI/BsrGI. We transformed chemically competent DH5α with the ligation mix.

LITE system plasmids: to construct the plasmid we removed the EGFP from pCIB1(deltaNLS)-pmEGFP (pCIB1_EGFP) and pLITE2.0_pAAV-EF1a-NLS-C2P-NLS-VP64_2A_EGFP. We digested both plasmids respectively with AgeI/EcoRI and XbaI and EcoRI. The plasmid pLITE2.0 woGFP lacks for eukaryotic resistance cassette so we linearized the vectors through digestion with AgeI and we cloned the puromycin cassette which is removed from pBML4 plasmid.
To modify the LITE plasmid we amplified the cDNA of *FokI* with primers GM81_FokI_F/GM82_FokI_R and the band was purified and cloned, obtaining pLITE2.0-FokIwoVP64-EGFP. We inserted a puromycin cassette and we transfected the cells.

Ecdysone-Inducible: we constructed a vector in which the expression of Cas9 is under the control of heat shock (HS) promoter, followed by IRES and reporter gene; HS is bind by ponasterone A that triggers the expression.
To do this we use several PCR reaction to amply the DNA fragment of interest. HS promoter and Neo/Kan resistance were obtained through a PCR respectively with GM_102 / GM_103 and GM_108/GM_109 primers that uses as template pIND vector. To monitor the induced expression we planned to clone report gene blue-fluorescent protein (BFP) obtained through PCR on pHR-SFFV-dCas9-BFP-KRAB plasmid with GM_106 and GM_107 primers; the IRES sequence was amplified from pBuerstedde vector with GM_104 and GM_105 primers.
To be able to clone them in our backbone vector (pX330) the amplified products were digested in according with our strategy. pX330-Cas9 and HS promoter were digested with XbaI/AgeI and the vector pX330-HS-Cas9 was

linearized with NotI to clone the Neo/Kan resistance treated with the same restriction enzymes. We opened the new plasmid pX330-HS-Cas9—Neo/Kan through a digestion with EcoRI to clone the IRES sequence and the BFP gene both digested with EcoRI/NcoI.

Knock-out of P53 in CH12-F3: This reporter vector was obtained through the insertion of annealed and phosphorylated oligonucleotides (41-42) into pBS-CMV-mcherry-EGFP; subsequently we digested the plasmid with BsmBI enzyme to insert the CRISPR target. The inserted sgRNA, homologous to p53 sequence, followed by an out-of-frame blasticidin cassette. If the CRISPR/Cas9 system work correctly, the Cas9 can mutates the target sequence inserted in the region upstream the blasticidin gene; these lesions can be repaired and the frame of the resistant gene can be restored allows us the selection of clones in which the CRISPR/Cas9 system worked.
The CRISPRs to target p53 loci on intron 4 were prepared as per protocol in the same way of CRISPRs used to introduce single nick on myc loci; the sequence are described in "Materials and Methods, section Primers".
CH12-F3 $myc^{+/EGFP-}$ were expanded and transfected with pX330-p53ko and pBS-CMV-mcherry-p53ctrl-BSR through electroporation; the cells were inoculated for 24 hours in CH12-F3 medium plus blasticidin antibiotics (30-50 ng/ml). This passage permits to Cas9 to act on p53 locus and on reporter plasmid on which can restore the frame of blasticidin gene to confer the resistance phenotype.

Chapter III
Results and Discussion

The recurrent chromosomal translocation *MYC/IgH*, which can take place during CSR is a perfect model to study the factors involved in chromosomal alterations. To detect and analyze this translocation in samples of clonal origin (e.g. cancer tissue) the common techniques are Southern blot analysis or fluorescent in situ hybridization (FISH). These are techniques feasible only under the assumption that cells bearing the chromosomal translocation represent a sizeable portion of the sample. Of course, these techniques cannot be used to study the onset of chromosomal translocations. On the contrary, when studying the frequency of chromosomal translocations the most suitable approach is the Long Range PCR, which uses nested primers on the MYc gene and on the Switch regions of the IgH locus to amplify the DNA fragments deriving from the translocation. Using these technique with samples at different dilutions, it is possible to infer the translocation frequency. Nonetheless, the method is rather expensive, due the high number of PCR reaction with long range polymerase to be performed. Moreover, being based on extracted DNA, it is not possible to go back or to dissect the original cell/genome where the translocation has been found.
During my PhD we developed a system to allow us the study other factors involved in the chromosomal translocation *MYC/IgH* in a more effective way. Namely, we wanted to make the onset of the *MYC/IgH* translocation in single cells visible by FACS analysis, eventually coupled with Cell Sorting.

We chose the mice lymphoma cell line (CH12-F3) as the basis for our system. CSR can be readily induced in this cell line using a mix of cytokines and

stimulation of CD40 signaling. After induction of CSR the switch from surface IgM to surface IgM is readily analyzed by FACS analysis.
To this aim we planned to replace one of the two alleles of the MYC gene with an out-of-frame Enhanced Green Fluorescent Protein (EGFP). The frame of the reporter gene could be restored only in cells bearing the chromosomal translocation, with the result that the fluorescence can be detected by FACS analysis.
The development of this system could allow us to study new factors involved in the onset of chromosome translocations. We could use this system to prove the role of specific genes, or we could use it to screen libraries to identify novel factors involved in their onset.

Designing a vector for Knock-in CH12-F3

The designing of the targeting construct is crucial for the engineering of our tool **(Figure 3.1)**. The vector for the knock-in (KI) to produce a CH12-F3 cell line CH12-F3-myc$^{+/EGFP}$ is composed by three main components:

- the backbone of the plasmid
- the 5' and 3' homology arms
- the cassette to be inserted at the location of the intended knock-in.

We used the plasmid pBS-SK(+) as backbone for the targeting construct, an often used plasmid.
The position of the homology arms depends on the region we aim to target in the MYC gene. Most common breakpoints in *MYC/IgH* translocations occur within intron 1. We therefore planned to insert the cassette to replace exon 2 and exon 3 of *MYC*, which encode for most of the MYC cds, downstream to the breakpoint cluster.
The homology arms were designed on the sequence of murine MYC gene, flanking the region encompassing the end of exon 1 until the end of exon 3. We aimed to replace one allele of MYC gene with an out-of-frame EGFP downstream the breakpoints cluster regions on intron 1 of the gene.

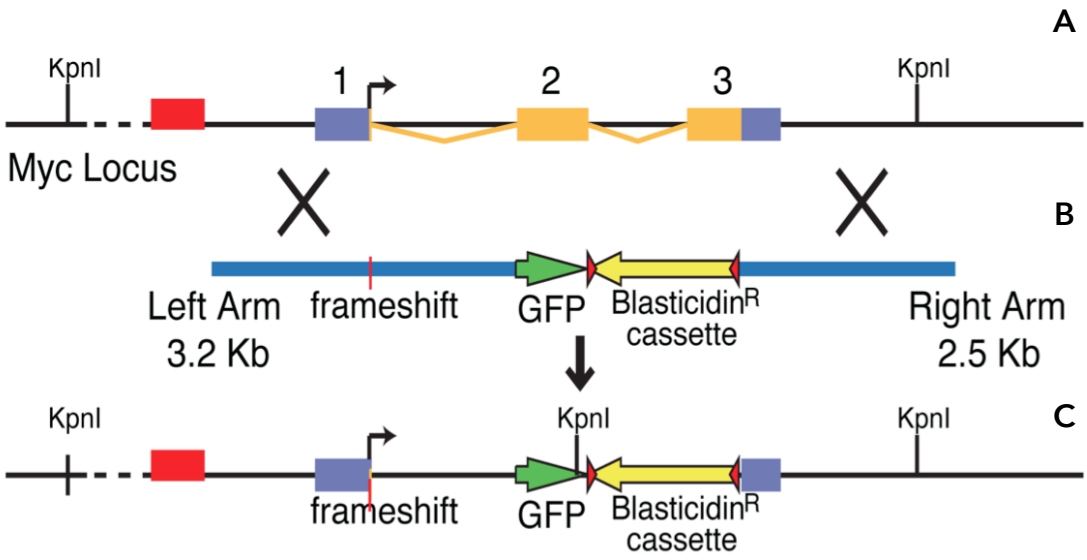

Figure 3.1 - Targeting strategy to knock-in an EGFP cds in place of one of the two alleles of MYC. A) Representation of the murine MYC gene; the boxes 1, 2 and 3 indicate the exomes (orange boxes indicate the coding sequence). **B)** Schematic representation of the vector for the Knock-in. The homologies arms are indicated in blue and are interspersed with out of frame GFP and blasticidin cassette. **C)** Representation of successful knock-in, in which our construct replaces the entire cds of MYC gene.

Moreover, in designing the targeting of the MYC locus, we had to take in consideration also the transcription and translational features of the MYC gene. Contrarily to the human MYC, the start codon of the murine MYC cds is a CTG codon **{**Dang, 1999**}** located at the end of the first exon. That meant that - in order to avoid expression of the EGFP transgene in non-translocated cells, the end of the first exon in the MYC gene had to be modified as well **(Figure 3.2)**. We therefore planned to insert a frameshift downstream to the start CTG to make unlikely the synthesis of the EGFP while located in the original knock-in location. To make sure that the ATG start site on the GFP cds could not be used we further created another ATG upstream to the frameshift, so that - if the CTG codon had failed, our extra ATG could serve our purpose.
In our plan only translocation of the second exon of the MYC gene near the intronic enhancer of the IgH locus could trigger transcription of a translatable EGFP cds.

With regards to the cassette to be knocked-in, we used an EGFP cds, placed right after the exon 2 splice junction. A blasticidin resistance cassette (BSR-cassette) was placed after the stop codon of the GFP cassette. This cassette,

to be used for the selection of the stably transfected clones, is flanked by two lox sites in order to make it possible its ablation after selection of the targeted clones {Peitz, 2002}.
After excision of the BSR cassette the GFP cds would perfectly replace the cds of the MYC included between the exon 2 and the stop codon in the exon 3.

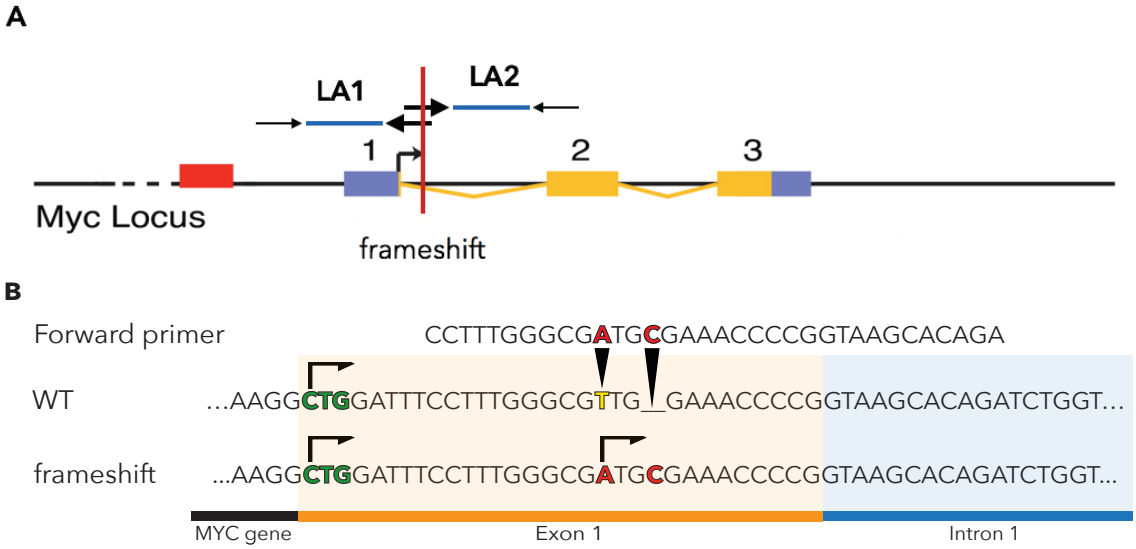

Figure 3.2 - Schematic representation of the frameshift. A) The primers used to introduce the frameshift (bold arrows) bind the MYC sequence at start codon **B)** frameshift sequence details; the red letters indicate the nucleotides that we introduce. The green CTG represent the WT starting codon.

Building the Knock-in vector

The order of the cloning was set to the insertion in the pBS of the 3' homology arm (right arm, RA, 2239 bases) preceded by EGFP cds and the blasticidin resistance cassette (BSR-cassette). Last part is the insertion of 5' homology arm (left arm, LA, 3286 bases) mutated at the end of the exon 1.

The designed arms were obtained through a nested PCR with a different set of oligonucleotides.

The Right Arm was obtained by amplification through, mycRAnf/mycRAnr and mycRAf/mycRAr primers. The obtained band was purified and digested NheI/NotI.
We linearized the pBS−SK(+) with NheI/NotI, two restriction enzymes that recognize sites downstream the BSR-cassette, and we cloned the RA fragment into the plasmid.
The 5' homology arm, which includes the end of exon 1 to be mutated was amplified by stitching together two DNA fragments obtained by PCR. After evaluation of the site in which introduce the two points mutation to induce the frameshift, we divided the left arm sequence in two part: left arm 1 (LA1) and left arm 2 (LA2).
We amplified both DNA fragments using a nested PCR, with a first PCR amplification using external primers (mycLA1nf/mycLA1nr and mycLA2nf/mycLA2nr), followed by another one with a pair of internal oligonucleotides. We joined them through another PCR reaction in which the first 10 cycles were run without primers to allow LA1 and LA2 to extend on each other through the complementary sequence represented by primers mycLA1r and mycLA2f. The last 20 cycles of the PCR program were run after addition of primers mycLA1f/mycLA2r. The band was purified and digested with NcoI for the cloning in the backbone vector.
We digested the plasmids pTarget-MYC missing for left arm with NcoI to open the plasmid in the region immediately upstream of the coding sequence of EGFP gene to clone the LA.

Transfection of the Knock-in construct and first screens

The final plasmid pTarget-MYC was linearized with ApaI and transfected in CH12-F3 through electroporation. The cells were plated in RPMI medium with blasticidin (30 ng/ml) and after ~10 days we picked up the colonies grown.
We extracted genomics DNA for each samples; we digested the DNA with KpnI an enzyme previously selected to obtain two possibles bands through Southern blot. The probe was obtained by PCR amplification from genomic DNA with mycprobeF/mycprobeR primers.

Overall we performed several transfections on CH12-F3 cells, and we screened by Southern blot ~500 clones, but we could not detect any targeted clones (**Figure 3.3**).

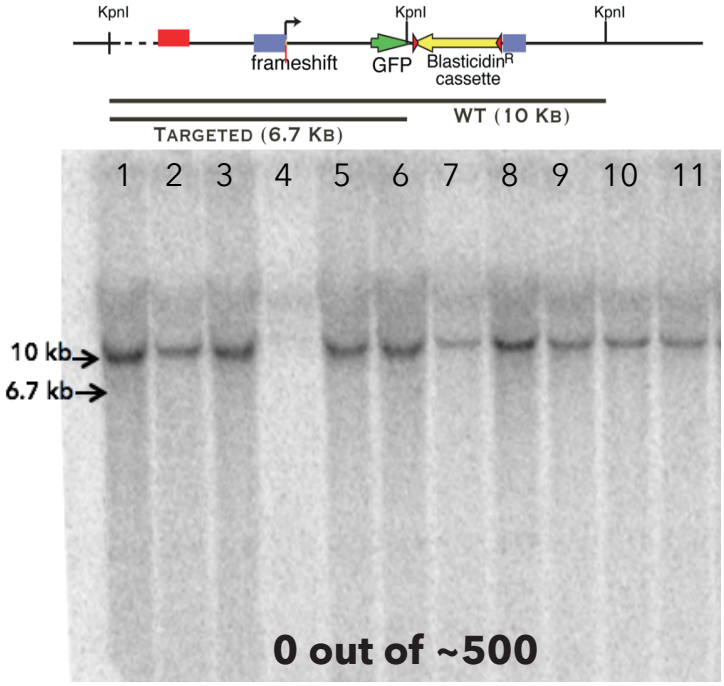

Figure 3.3 - Souther blot. The resistant clones were expanded extract and digest genomic DNA. Samples were analyzed by Southern blot to reveal the positive band. The probe was designed to bind a sequence located in 5' region respect to left arm homology sequence (red). No positive clones have been detected.

TALEs and TALENs

Due to the difficulty in obtaining the clones, and given the availability of novel tools for genome editing, we decided to use TALENs to improve our chances to obtain targeted clones.

Transcription activator-like effectors (TALEs) are DNA binding proteins discovered in *Xanthomonas,* a plant pathogenic bacteria {Boch et al. 2009; Moscou and Bogdanove 2009}. TALEs recognize and bind specific promoters and trigger the transcription of downstream genes. Bacteria use TALEs to colonize plants through the induced-expression of specific genes to increase susceptibility of the host. The ability of TALEs to recognize and bind a specific DNA sequence is mediated by a modular motif.

These modulars motif are composed by more the 30 tandem repeats and each repeat contains a sequence of 33 to 35 amino acid; the sequence is largely invariant with the exception of two amino acid in Repeat Variable

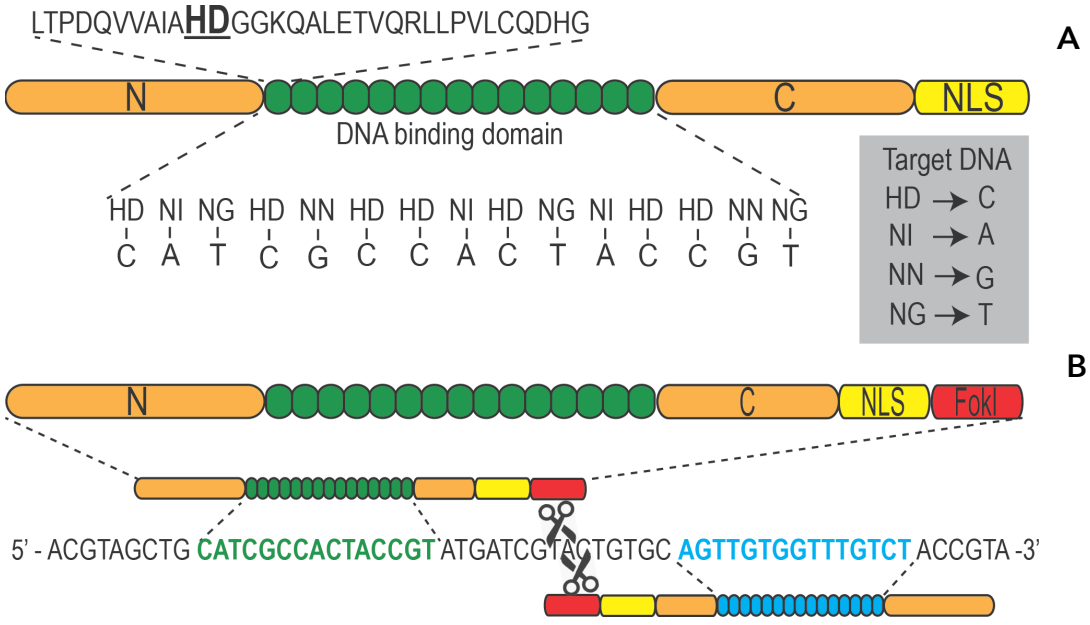

Figure 3.4 - The schematic represents the TALEN system. A) the tandem repeat binds specific DNA sequence (in green).B) TALEs fuses N-terminally with the FokI (red box). Two subunits of FokI nuclease can dimerize and cut double strand DNA . Modified from {Zu, 2013}.

Diresidue (RVD). RVDs play a key role in recognition of the target sequence as each of them recognize a specific single DNA base (NI=A, HD=C, NN=G, and NG=T). The characteristic allows researchers to build and customize modular domain specific for almost any DNA sequence of interest.
Transcription activator-like effector nucleases (TALENs) are customizable specific molecular DNA scissors composed by a modular domain of TALEs fused directly with the active domain of FokI, a nuclease able to cleave double strand DNA upon dimerization {Zu, 2013; Chen 2013} (**Figure 3.4**).

To increase the homologous recombination for the knock-in, we assembled two pairs of TALENs respectively TAL1-LA / TAL2-LA and TAL3-RA / TAL4-RA able to bind and induce breaks at the LA and RA homology sequence of the myc gene.
We transfected both pairs of TALENs (pTAL1-LA / pTAL2-LA and pTAL3-RA / pTAL4-RA) together with the targeting construct pTarget-MYC in CH12-F3 cells. We obtained 36 selected colonies, and Southern blot analysis revealed that 6 out of 36 clones could be contain the cassette targeted into the myc locus (**Figure 3.5**).

We expanded the clones and, after genomic DNA extraction, we performed PCR analysis using a pair of primers (mycLAf/EGFP2R) designed to recognize

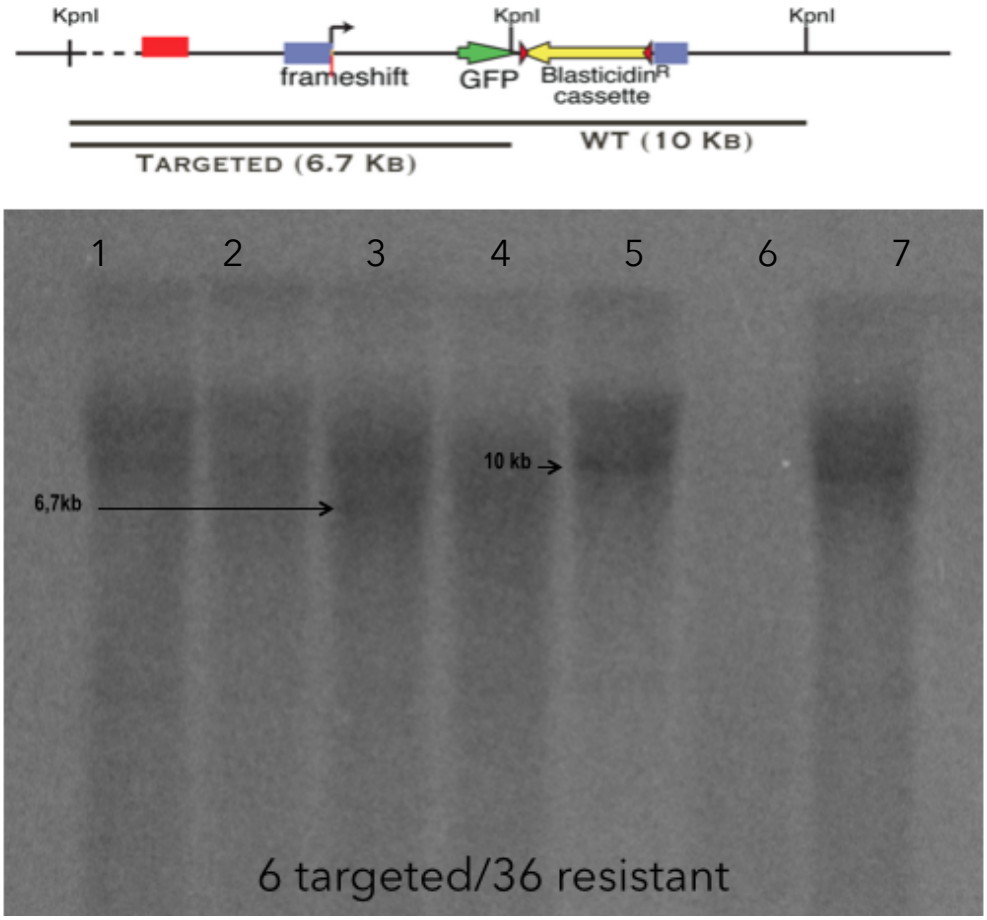

Figure 3.5 - **Southern blot membrane.** The resistant clones were expanded extract and digest genomic DNA. Samples were analyzed by Southern blot to reveal the positive band. The probe was designed to bind a sequence located in 5' region respect to left arm homology sequence (red). In position 3 we detected the positive band (6.7 kb) which reveals the presence of successful knock-in.

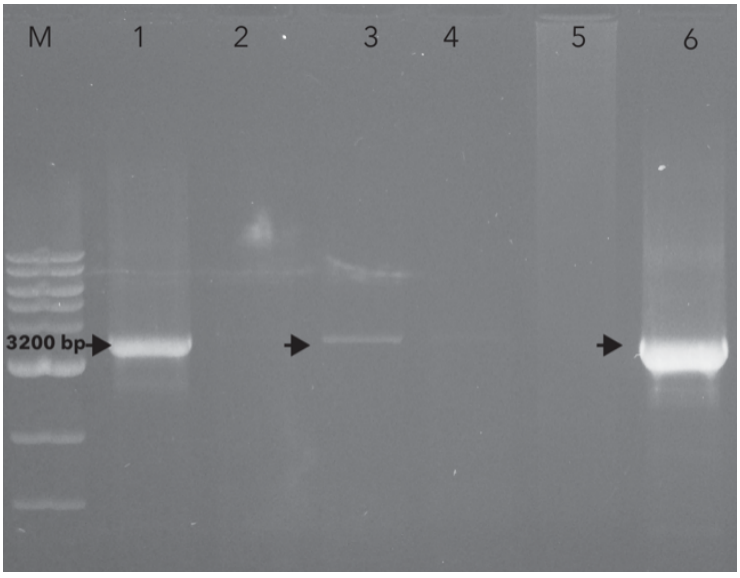

Figure 3.6 - PCR screening. PCR was performed on the 6 positive clones for Southern blot. Primers were designed to amplify a fragment of 3.2 kb that indicates the positive samples for the knock-in (lines 1, 3 and 6). The DNA fragment was sequence verified, and two clones obtained bore the expected sequence.

a region upstream to the LA homology sequence and the EGFP (**Figure 3.6**). We cloned the amplified bands into pCR-BluntII TOPO for sequence analysis. The sequence results confirmed the presence of two positive clones CH12-F3 $myc^{+/EGFP}$ and one negative, in which the sequence was mutated. CH12-F3$^{myc\ +/EGFP}$ clones were stored at -80°C.

We treated the cells with TAT-Cre recombinase to remove the BSR-cassette from the genome, and we selected clones in which the resistance was lost.

Visualizing MYC/IgH translocation

The next crucial step in my work was to prove that the CH12-F3 *myc*$^{+/EGFP}$ could be used to visualize *MYC/IgH* chromosomal translocations, and thus validate the system.
The *MYC/IgH* translocation frequency is 2 x 10^{-4} in normal murine B-cells {Sernandez, 2008}, and 1 x 10^{6} in CH12-F3 {Boboila, 2011}, which could prove difficult to assess by FACS. On the other hands a number of factors had been shown to increase translocation frequency {Wray J, 2003; Ramiro, 2006}. We therefore set to assay the cells under different conditions that might increase the frequency of chromosomal translocations.

Expression of AID mutants to increase the translocation frequency

The C-terminal domain of AID bears a NES which prevents its accumulation in the nucleus {Ito, 2004}. This region contains also features that are necessary for CSR. Thus, an AID mutant with a mutated NES cannot complement AID deficient B cells with respects to CSR {Sabouri, 2014}. Beyond hindering the CSR, the expression of a mutant of AID lacking its C-terminal region (hAID-ΔC) increases the frequency of *MYC/IgH* translocation in CH12-F3 up to 12 per 1.5 x 10^{6} respect to 1 per 1.5 x 10^{6} {Doi, 2009}.

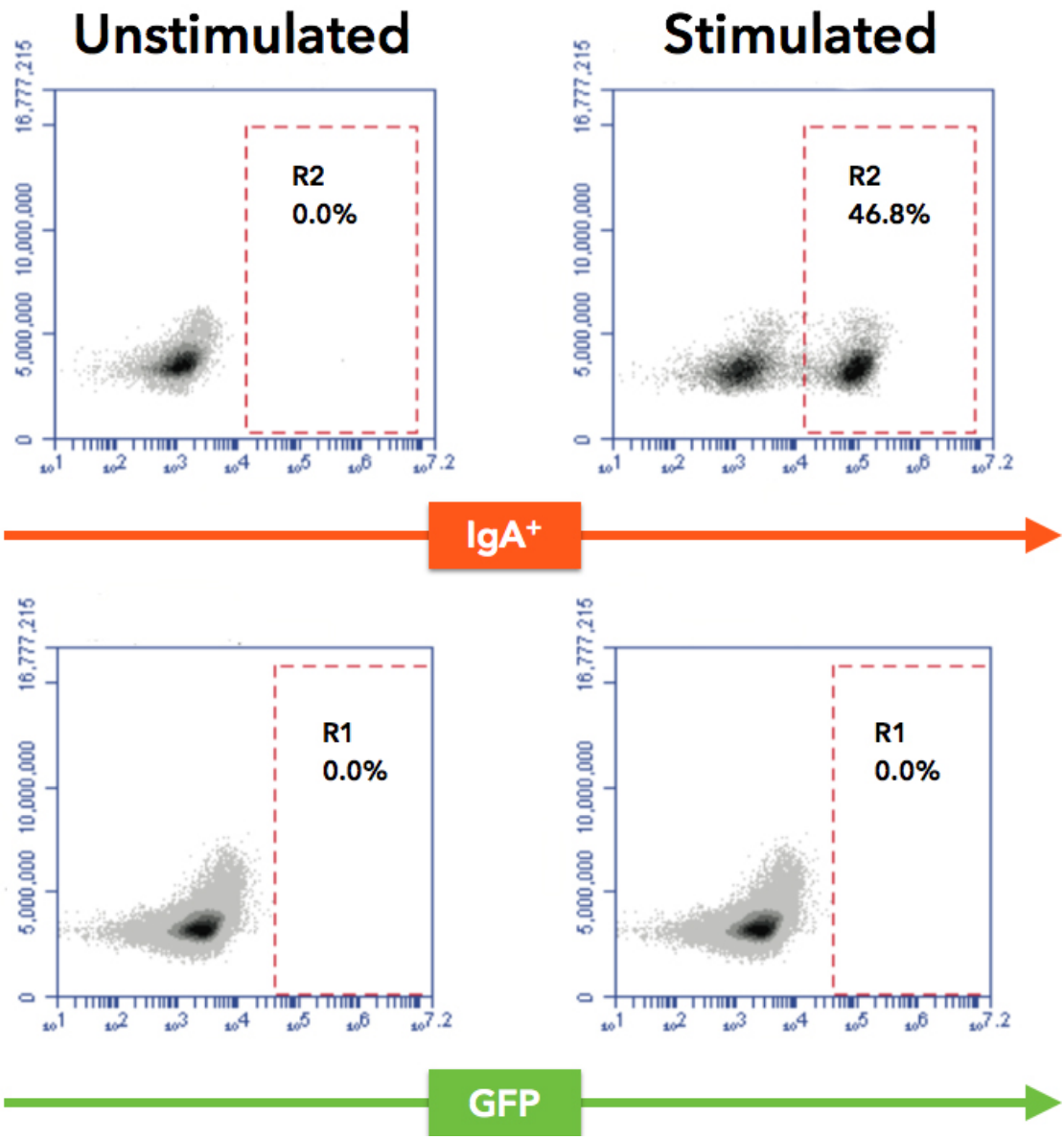

Figure 3.7 - FACS analysis of stimulated CH12-F3 myc$^{+/EGFP}$. The vector pBuerstedde-hAIDdeltaLastExon was transfected in CH12-F3 myc$^{+/-EGFP}$. Resistant cells were stimulated for 48 hours with TGF-β, IL-4 and anti-CD40 antibody. Subsequently the cells were stained with Anti-mouse IgA-RPE and analyzed by FACS.

We thus transfected CH12-F3 myc$^{+/EGFP}$ cells with the plasmid phAID-ΔC and we plated them in selection medium. Despite several transfection attempts only few colonies could be picked after ~2 weeks and expanded. The low number of the resistant colonies could be justified by the chronic genotoxic damage induced by this AID upmutant. The expanded clones were stimulated to induce CSR, as it had been shown that this would increase the translocation frequency. We analyzed the cells by FACS but we did not observe any GFP(+) cell even if the percentage of the switched cells was

correct (25 - 50%) (**Figure 3.7**). Yet, further analysis of the clones showed that none of them expressed the mutant AID, suggesting that its expression was shut during selection of the cells.

Hyperactive AID

Previous mutagenic studies showed that phosphorylation of serine 3 in the amino terminal part of AID reduces both CSR and onset of *MYC/IgH* translocation {Gazumyan, 2010}; the substitution of Serine3 with an Alanine prevents phosphorylation increasing the mutagenic activity of AID.
We thus modified the mutant hAID-ΔC introducing two single point mutations on the third codon to replace the serine with alanine; in addition we inserted a reporter gene in our plasmid to monitor the expression of the hAID-ΔC. To this aim we used an IRES sequence to drive the translation of a second coding sequence downstream to that of the AID mutant {Jang, 1988; Pelletier, 1988}.
We designed the plasmid to carry the modified mutant of AID flanked by an IRES sequence and by an mCherry reporter gene (pAIDS3-ΔC).

We linearized the plasmid pAIDS3-ΔC with EcoRI and we transfected CH12-F3 myc$^{+/EGFP-}$cells. The selected clones expressing hAIDS3-ΔC were stimulated to triggers the CSR and after 48 hours we analyzed the cells by FACS to detect translocated cells. Despite the successful induction of Class Switching, the FACS did not reveal any RED or GFP(+) cells.

Caffeine treatment to promote translocation

Previous work focused in understanding which factors are involved in DNA repair systems during the cell cycle demonstrated that cells treated with caffeine presented an increase in DNA lesions, including chromosomal translocation.
Caffeine directly inhibits Ataxia telangiectasia mutated (ATM), a serine/threonine protein kinase activated by DBSs. Caffeine triggers the DNA damage checkpoint through phosphorylation of several proteins involved in the repair process {Blasina, 1999}.

Caffeine-ATM inhibition represses the homologous recombination (HR) pathway {Zelensky, 2013} through inhibition with *p53*. Studies demonstrated that activated mice B-cells lacking for p53 gene showed an increase in the *MYC/IgH* translocation frequency with respect to the wild type {Ramiro; 2006}. We thus treated CH12-F3 *myc*$^{+/EGFP}$ cells with both the stimulation cocktail and caffeine 2.5 mM for 48 hours. After expansion of the cells for two weeks, FACS analysis showed the presence of GFP(+) cells only in the cells in which CSR had been induced together with caffeine treatment. (**Figure 3.8**).

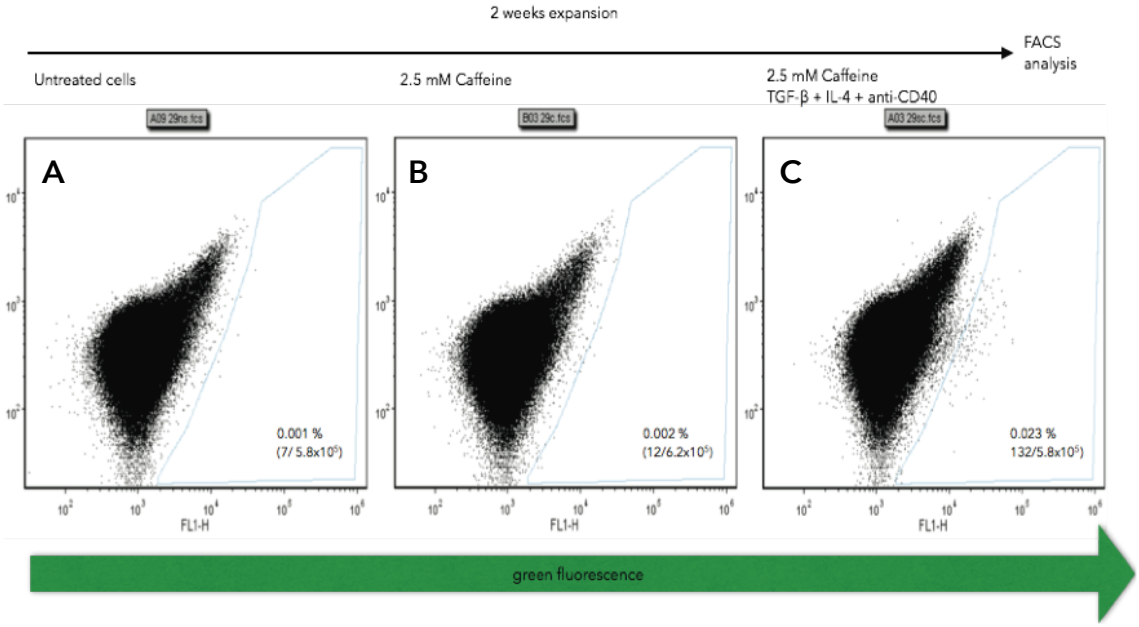

Figure 3.8- Flow cytometric analysis of stimulated and treated cells. A) The first plot represents the CH12-F3 *myc*$^{+/EGFP}$ untreated; **B)** Cells treated with caffeine; **C)** Cells treated with caffeine and stimulation mix.

This experiment suggests that our approach works. Yet, the frequency of chromosomal translocations is still lower than that reported {Boboila, 2011}, probably due to the crudeness of the caffeine treatment which does not specifically impinge on the specific pathway that leads to *MYC/IgH* chromosomal translocations. We thus planned to set a proof of principle using an artificial system through which DSBs could be induced at the *MYC* intron. To this aim we used the recently discovered CRISPR/Cas system.

CRISPR-Cas system

The CRISPRs-Cas system can be defined as an immune pathway in prokaryotic organisms which provides the means to recognize and destroy foreign DNA elements such as phages, virus and plasmids. Functionally similar to memory B-cells, CRISPRs provide the cells a heritable information of past infections allowing them to increase their response in case of a second infection.

Clustered regularly interspaced short palindromic repeats (CRISPRs) are DNA partially palindromic repeats interspersed with unique short DNA segments. These peculiar repeat elements of 24-50 bp have been described, using a bioinformatic approach, in different kind of bacteria and archaea, and are easily distinguishable in very distant phylogenetic groups {Mojica, 2000}. On the other hand, the origin "DNA spacer" between repeated motives arise from foreign genetic elements. Studies performed in bacteria revealed how sensibility to phage infection is strictly correlated to the number of extrachromosomal spacers on the CRISPRs locus and that the spacer elements showed a strong homology with phage genome suggesting that they arise from previous infections and provide cell immunity against further infections {Bolotin, 2005; Mojica, 2005; Pourcel, 2005}.

Genome sequencing of different bacteria strains identified the presence of four CRISPR-associated genes (*cas*), designed Cas1 to Cas4, adjacent to CRISPR loci suggesting a functional relation with the repeated sequence as operon {Jansen, 2002}. Different Cas protein families were discovered in multiple prokaryotic species {Haft, 2005} and their functions ranging from helicase to nuclease.

The CRISPR/Cas "immune" system is divided in three stages: adaptation, CRISPR expression, and "immunity". The Adaptation involves the acquisition and integration of DNA spacers that correspond to the "protospacer", the original sequence from virus, phage or plasmids genome. These sequences are flanked by the PAM (protospacer adjacent motif), a conserved short sequence at 3' end (5'-NGG) that represents the recognition motif for the endonuclease factor in the system. In the second stage the bacteria transcribe the CRISPR loci; the single-guide-RNA (sgRNA) is processed from the CRISPR and specifically recognizes other DNA complementary to its sequence. During the "immunity" phase, the formed hybrid sgRNA-DNA flanked by the PAM is recognized by Cas9, an RNA-guided DNA endonuclease enzyme able

to cut DNA by introducing single nicks on both strands of the sequence adjacent to the PAM.
The ability of Cas9 to recognize a specific RNA-DNA sequence composed by N_{20}-NGG allowed scientists to develop a new specific customizable gene editing system based on CRISPR/Cas9 system.

We thus designed sgRNAs for CRISPR/Cas9 system able to recognize and bind the intron 1 on both strand of Myc gene to introduce a double-nick that may promote the break of double strand; we cloned the sgRNAs for MYC gene into the pX330 as per protocol) we performed a transient transfection in CH12-F3 *myc*$^{+/EGFP}$ cells. After 48 hours we controlled the cells by FACS but we could not ever obtain a transfection efficiency higher than 2%. This means that even a translocation frequency of 1 in 100 could not have been detected.

Inducible promoters

We next tried to use inducible systems to drive the expression of either mutants of AID or the CRISPR/Cas9 system. To this aim we have constructed expression constructs based on either the Tet-on system or the Ecdysone-Inducible one.
To this aim we have built expression vectors to test the expression in CH12-F3 cells. Unfortunately, we found that induction of the Tet-on construct with Doxycycline led to cell death in CH12-F3 cells. On the other hand induction with Ponasterone-A to trigger transcription of the Ecdysone-Inducible constructs did not succeed in our cell line.

Work in Progress

LITE system

Another attempt we made to induce expression in CH12-F3 cells was based on the LITE inducible system, which - being based on blue light exposure - is not invasive or harmful to the cell {Konermann and Brigham, 2013}.
Characterized few years ago, the interaction of plant and microbial-derived light-sensitive proteins (from *Arabidopsis thaliana)* CIB1 and cryptochrome 2 (CRY2) is induced by blue-light (466 nm).
The modified system we tried to employ is based on chimeras fused with either CIB1 or CRY2. These were engineered to contain customizable TALE DNA-binding domain to become an optogenetic inducer. TALE-CRY2 protein binds the upstream region of targeted gene; CIB1 is fused with VP64 domain able to recruit the transcription complex. In presence of blue-light CRY2 becomes active and can interact with CIB1 carrying the VP64 domain to trigger transcription {Konermann and Brigham, 2013} (**Figure 3.9**).
We tried to use this system to induce the transcription of genes involved in a-NHEJ pathway, linked to chromosomal translocation formation. Exploiting the induction of the expression by blue light in combination with the stimulation to promote CSR an increase in the translocation frequency might have been observed in CH12-F3 *myc*$^{+/EGFPr}$ expressing constitutively CIB1 and CRY2.
We decided to induce the transcription of PARP1 and LIG3 genes. Previous studies demonstrated the both PARP1 and LIG3 play a role in the a-NHEJ pathway, promoting the translocation {Paul, 2013; Mansour, 2010; Simsek, 2011; Soni, 2014; Wray, 2013}. We thus built TALEs to recognize the promoter region of PARP1 and LIG3 and we cloned them into pCIB1.

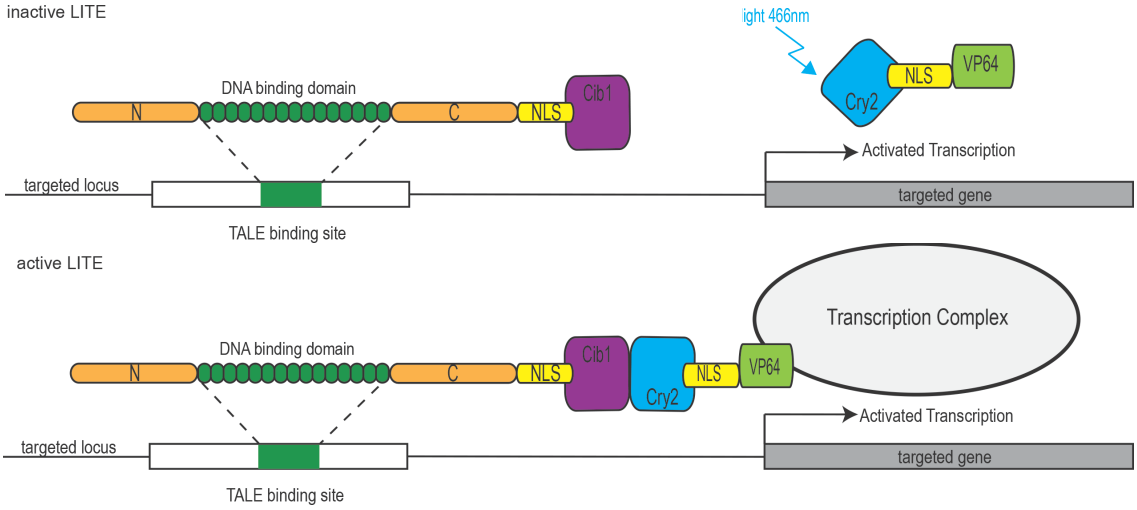

Figure 3.9 - Schematic of the LITE system. TALE module binds target sequence (green) and locates the CIB1 (purple) upstream the transcription start site. B) Light stimulation (466 nm) induces dimerization of CRY2 and CIB1 that leads to activator VP64 to recruiting the transcription factor complex, triggering the transcription of the targeted gene.

We first transfected pLITE2.0woEGFP_PURO into CH12-F3 *myc*$^{+/EGFP}$and we obtained stable clones expressing the chimaera CRY2-VP64.

We next transfected the plasmids encoding the TALEs targeted at PARP1 and LiG3 (pCIB1_woEGFP_TAL_LIG3 and pCIB1_woEGFP_TAL_PARP1) in CH12-F3 myc$^{+/EGFP}$-CRY2-VP64 cells.
We obtained clones expressing the TALEs, but we are still setting up the light system to induce the interaction.

In parallel we prepared plasmids combining the LITE system with the FokI nuclease. Such combination should allow us to target the induction of DSBs at will to the selected locus. To do this we replaced the transcription-activator VP64-EGFP with FokI gene; as TALEs we used the same couple designed for the knock-in; so as to bind the intron 1 of MYC gene which contains the cluster region.

CRISPR KNOCK-OUT P53 in CH12-F3 myc$^{+/EGFP}$

Chromosome translocations occurring in B-cells involving immunoglobulin Switch regions are triggered by AID during the Class Switch Recombination process {Muramatsu, 2000; Petersen-Mahrt 2002}. The abasic sites generated by UNG that removes uracil from DNA, are processed to double strand break {Rada, 2002; Schrader, 2005}. The repair systems that mediate the aberrant joining of IgH and non-Ig genes, is different from intrachromosomal repair pathways during Class Switch Recombination that requires histone H2AX {Petersen, 2001}, p53 binding protein 1 (53BP1) {Manis, 2004; Ward, 2004} or non-homologous end-joining protein Ku80 {Casellas, 1998}. Moreover tumor suppressors, as ATM, Nbs1, p19 and p53, activated by DNA damage and oncogenic stress during the physiological switching, inhibit the translocation, protecting the cell healthy.
Previous works showed how mice knock-out for one allele of the p53 gene display an increased translocations frequency of *MYC/IgH* in B-cells {Ramiro, 2006}.

We thus prepared the CRISPR/Cas9 to target the exon 4 of the murine p53 gene; the second one is a reporter plasmid able to control the specificity and activity of CRISPR/Cas9 system. We are currently performing the transfection of the constructs in CH12-F3 cells.

Conclusions

We think that the model that we are developing to study the onset of translocation *MYC/IgH* could be an interesting approach to clarify how these genetic alteration occur. Once properly set up, it will allow us to get insights on the relation between antigen-driven antibody diversification processes and the onset of cancer-inducing chromosomal translocations. While we were able to observe the rise of events in our cells engineered to visualize the *MYC/IgH* translocation, we could not observe the same frequency observed in previous studies. We are currently working at setting up the system to obtain a higher number of events.

Considering the problem we faced to set up the assay, we have prepared a new targeting construct in which there is a puromycin resistance cds fused to the EGFP one. This will allow us to select the cells in which the translocation has occurred by antibiotic.

We could use the assay to screen for novel factors involved in the generation of chromosomal translocations. To this aim we would transduce the myc/GFP cell line with (i) expression libraries to identify genes that promote chromosomal translocations, or (ii) CRISPR libraries to identify genes that protect against their onset. Once the cells are transduced with these libraries and CSR is induced, we will sort/select the ensuing fluorescent cells in which the translocation has taken place. After expansion of the cells, we will amplify the construct that had been transduced in order to identify the gene/shRNA in it. Transduction of the identified gene/Ablation of the gene will be used to confirm its ability to trigger chromosomal translocations.

Splice variants of Activation Induced Deaminase (AID) do not affect the efficiency of Class Switch Recombination in murine CH12-F3 cells

Introduction

Activation Induced Deaminase (AID) is the DNA editing enzyme that triggers the secondary immune diversification process in activated B cells. {Muramatsu ,2000; Petersen-Mahrt 2002; Neuberger 2008}. Whereas AID action is physiologically exerted on the immunoglobulin locus, AID dependent damage can induce mutations and chromosomal translocations in a cohort of

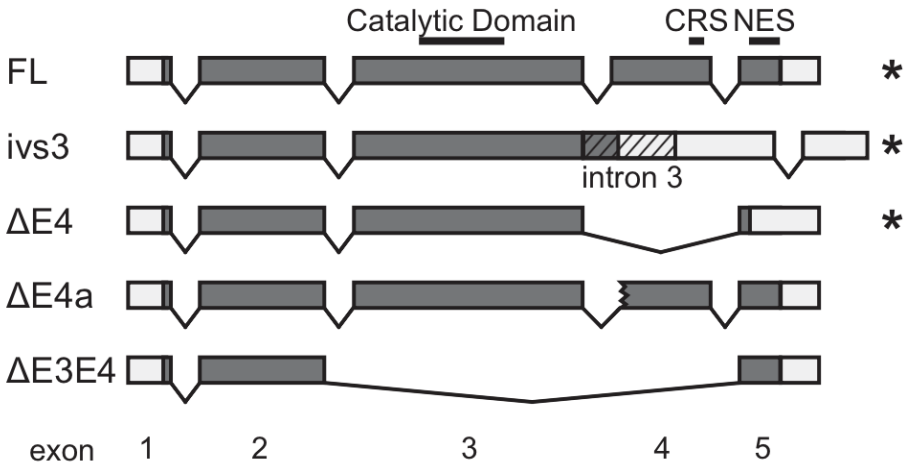

Figure 3.10 - Schematic representation of the splice variants of AID The exonic structure of the splice isoforms is shown. The position of functional features of the full-length AID (AID-FL) is indicated: the catalytic domain, the cytoplasmic retention signal (CRS) and the nuclear export signal (NES). The coding sequence appears in grey and the retained intron 3 is indicated (AID-ivs3) by the angle-striped pattern. The asterisks indicate the AID isoforms tested in the study.

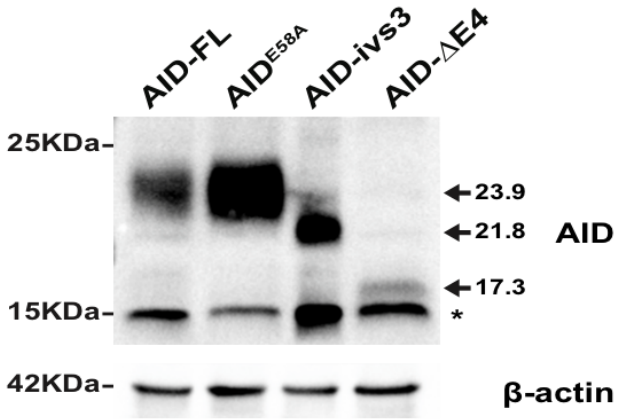

Figure 3.11 - Expression of the AID variants in HEK293T cells Western blot analysis of HEK293T cells transiently transfected with the various constructs (AID-FL, AID-ΔE4, AID-ivs3, AID^{E58A}). Equal amounts of protein lysates (30 µg) were loaded on SDS-PAGE. ß-actin was used as a loading control. The arrows indicate the expected molecular weight of the AID isoforms. The asterisk indicates an unspecific band.

other loci {Pasqualucci, 2001; Kotani 2007; Okazaki 2007} as well as from experimental systems {Liu, 2008; Ramiro, 2004; Wang, 2004; Parsa, 2007; Dorsett, 2007; Robbiani, 2008; Robbiani, 2013}. Expression of AID has been found in a number of B cell tumors {Albesiano, 2003; Greeve, 2003; Oppezzo, 2003; McCarthy, 2003; Babbage, 2004; Heintel, 2004; Forconi, 2004; Wu, 2008; Marantidou, 2010; Palacios, 2010}. Together with the full-length form of AID, other splice transcripts have been identified in B cell tumors, initially, and then in normal B cells {Noguch,i 2001; Albesiano, 2003; Greeve, 2003; Oppezzo, 2003; McCarthy, 2003; Babbage, 2004; Heintel, 2004; Forconi, 2004; Wu, 2008; Iacobucci, 2010; Marantidou, 2010}. Indeed, the presence of the various splice variants has been inversely correlated to the mutational status of immunoglobulins in B-cell chronic lymphocytic leukemia {Albesiano, 2003; McCarthy, 2003; Oppezzo, 2003; Babbage, 2004; Marantidou, 2010}. This has opened the possibility that at least some of these splice variants might be part of the regulatory network of AID.
While initially deemed as catalytically active {Wu, 2008}, subsequent reports clarified that AID splice variants are neither able to support deamination of DNA nor to trigger antibody diversification on their own {van Maldegem, 2009; van Maldegem, 2010}. Yet, even an inactive isoform could play a role in the regulation of AID: interaction with the catalytically active AID itself or with other molecules in its pathway could modify the physiologic activity of AID. In order to clarify this point we have thus tested the ability of some of these splice isoforms to affect the efficiency of Class Switch Recombination (CSR) in CH12-F3 cells, a murine lymphoma cellular model in which the efficiency of CSR can be assessed {Nakamura, 1996}.

Results and Discussion

We cloned two of the splice isoforms of AID (AID-ivs3 and AID-ΔE4) that encode for truncated forms of AID in which either the fourth exon is skipped or the fourth intron is retained (**Figure 3.10**). In either case such alternative splicing results in the lack of the C-terminal 57 amino acids of the catalytically active AID. These isoforms lack both the Nuclear Export Signal, necessary for an efficient CSR, and a cytoplasmic retention signal {Ta, 2003; Ito, 2004;

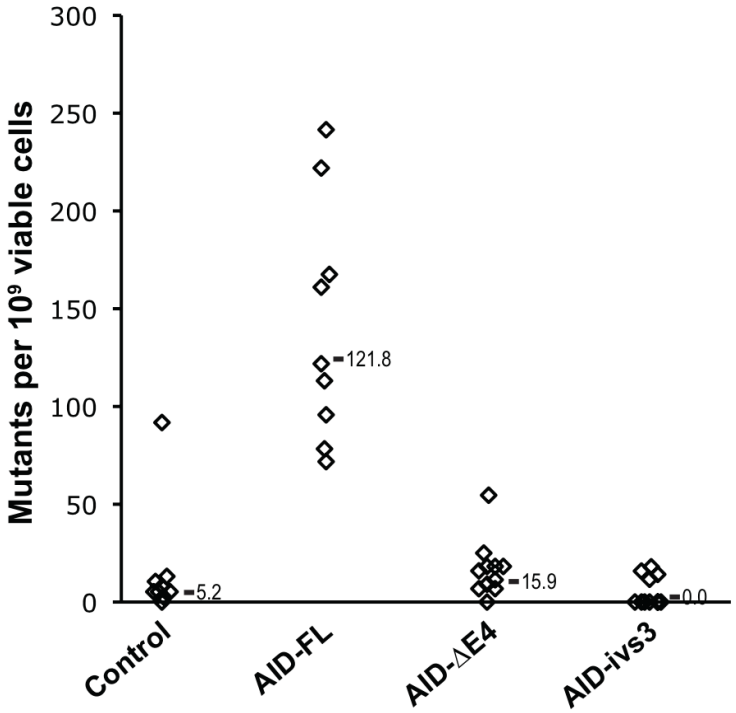

Figure 3.12 - Rifampicin assay using the various AID isoforms. Only the revertants resistant to rifampicin can grow. While AID induces a mutator phenotype ($P<10^{-3}$ by Dunn's multiple comparison test), the tested splice variants display levels of revertants similar to the negative control (empty plasmid). The mutation rate is calculated after normalization with Ampicillin resistant viable colonies. The median value is indicated.

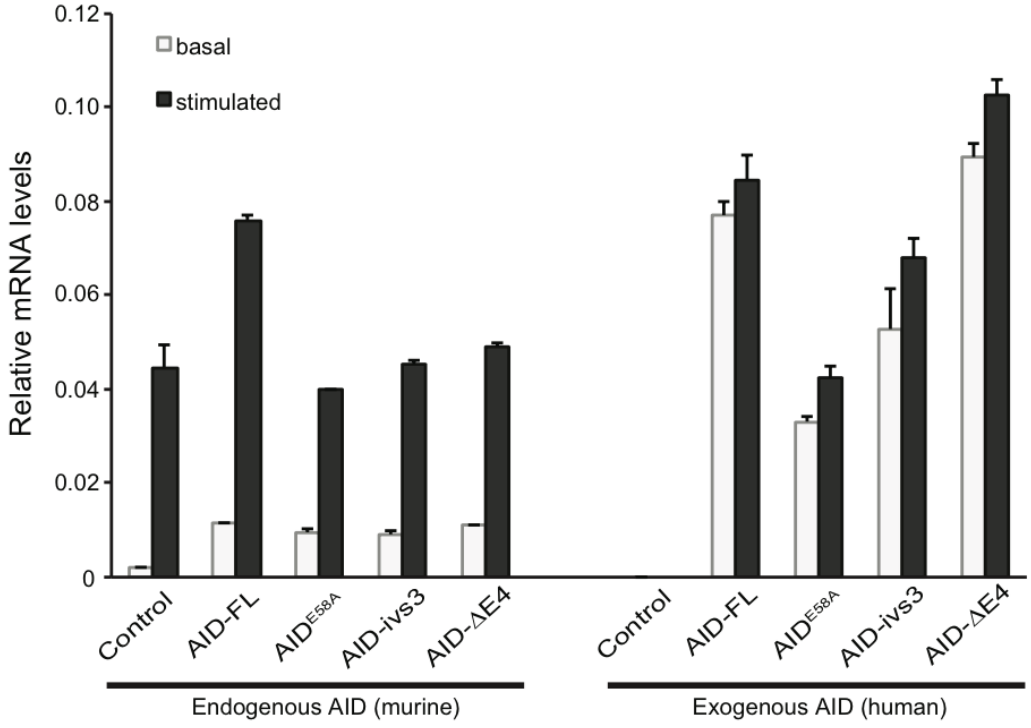

Figure 3.13. Expression of the AID variants in HEK293T cells
Quantitative real-time PCR showing the expression levels of endogenous (murine) and exogenous (human) AID from representative clones of CH12-F3 cells stably transfected with the constructs for AID-FL, AID-ΔE4, AID-ivs3, AID^{E58A}, or from control cells. Basal expression levels (gray) are compared to those in stimulated cells (dark columns).

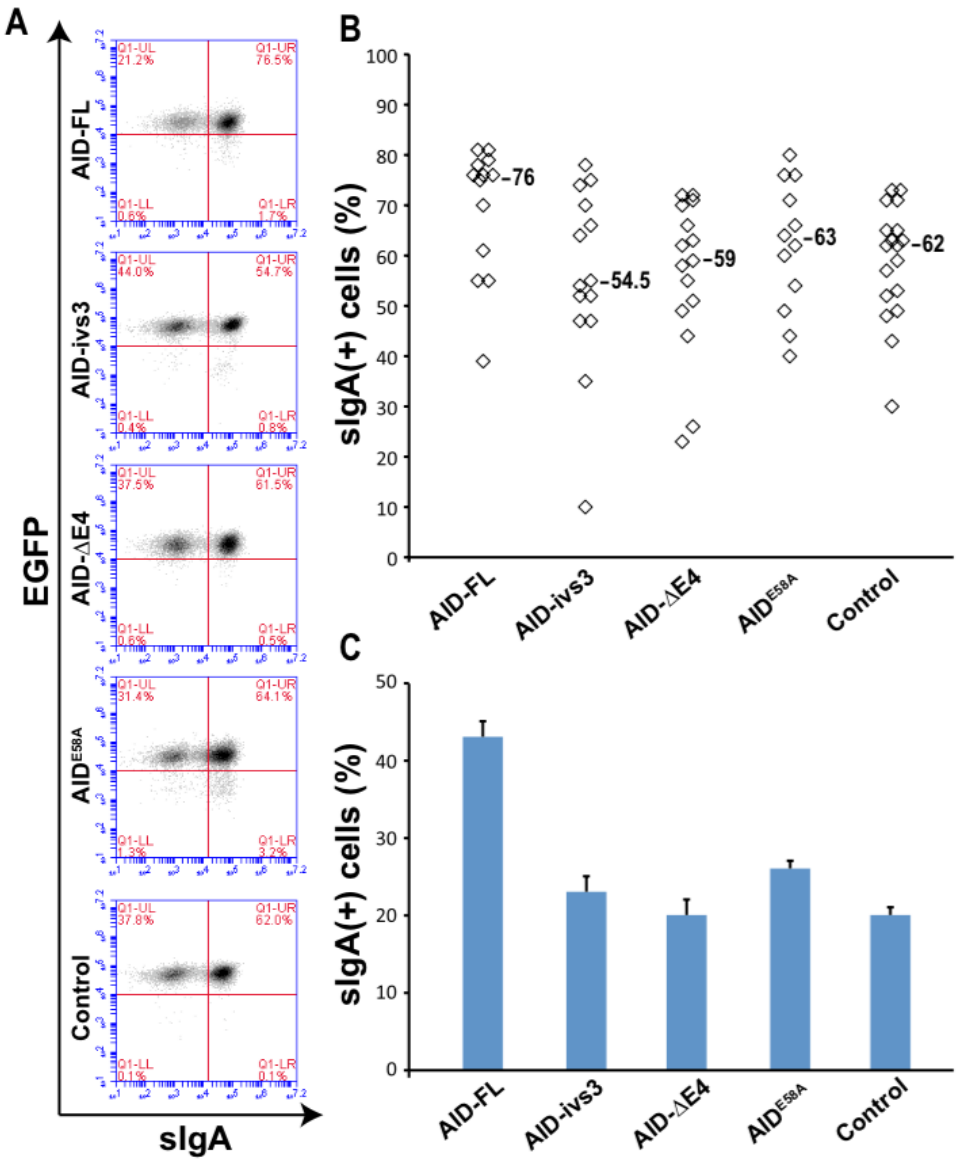

Figure 3.14 - AID variants do not affect Class Switch Recombination in CH12-F3 cells

(A) FACS profiles of representative CH12-F3 clones stably expressing AID-FL, AID-ivs3, AID-ΔE4 AID^{E58A}, or an EGFP negative control. Cells were induced for CSR and assayed by FACS after 72 hr. The numbers in the upper-right quadrants indicate the percentage of cells within the scatter window that are sIgA(+)/EGFP(+). (B) Each datapoint in the plot represents the percentage of sIgA(+)/EGFP(+) cells in an independent clone expressing the various constructs. While there is an increase in the efficiency of the CSR in the clones transfected with full-length AID, none of the samples reaches statistically significance (one-way ANOVA coupled with Tukey's test). The median level of CSR is indicated. (C) CSR analysis of cells transiently transfected with the constructs encoding for AID FL, AID-ΔE4 and AID-ivs3, AID-E58A, and a plasmid containing only the EGFP negative control. The analysis was performed on the bulk of transfected cells after 48 hours of selection with puromycin and 72 hours after induction with the stimulation cocktail. Transfected EGFP(+) cells represented ~2% of the total population. The only sample affecting the CSR is the one expressing the full-length AID (one-way ANOVA coupled with Tukey's test, *P=0.0004*). The histograms represent the mean of four different experiments and the standard error is shown.

McBride, 2004; Brar, 2004; Patenaude, 2009}. After cloning these splice variants in bacterial expression vectors (**Figure 3.11**), we tested their ability to induce a mutator phenotype in bacteria in the rifampicin-resistance reversion assay {Petersen-Mahrt, 2002}. The inability of the splice variants of AID to induce an increased level of revertants suggests that they are catalytically inactive, as suggested by the previous *in vitro* analyses {van Maldegem, 2009; van Maldegem, 2010} (**Figure 3.12**).

In order to assess the effects of these splice variants on Class Switch Recombination, we cloned them in a mammalian expression vector and we tested them in HEK293T cells. Compared to the full-length AID, the AID-ΔE4 variant consistently appeared expressed at lower levels, whereas the AID-ivs3 variant, which differs in its C-terminus from AID-ΔE4, was expressed at slightly higher levels (**Figure 3.11**). This result is compatible with the increased instability of the splice variants observed by Rebhandl et al. {Rebhandl,, 2014}. We next transfected all the mammalian expression constructs in CH12-F3 cells, and we have selected clones expressing them by puromycin-selection and expression of the EGFP reporter. Quantitative real-time PCR of transfected CH12-F3 cells before and after induction with TGF-ß, IL-4 and CD40 stimulation suggests that expression of the exogenous AID variants, at least at the transcriptional level, matches that of endogenous AID in stimulated cells (**Figure 3.13**). Due to technical limitations, it is not possible to assess AID protein levels in CH12-F3 cells, but it is reasonable to assume that protein levels of AID splice variants should be proportional to those observed in HEK293T cells and they might be preferentially degraded through the proteasome {Rebhandl, 2014}. On the other hand, it must be taken into account that – contrarily to full-length AID – the splice variants are predominantly nuclear {van Maldegem, 2009}. The probable stoichiometric ratio between wild-type and splice variant would not support a direct interaction between the isoforms. Nonetheless, even though steady-state levels of the splice variants could be lower than those of endogenous AID in stimulated CH12-F3 cells, it could still be possible that the splice variants could interfere with the physiological activity of AID through the catabolic pathway.

To test whether any of these constructs could modulate the activity of the induced endogenous AID, we have induced CSR in the selected clones with

TGF-ß, IL-4 and CD40 stimulation. Cells were assayed 72 hours after stimulation, a time-point that usually provides high levels of CSR (~50%). Indeed, only overexpression of the full-length AID elicited a slight increase (not statistically significant) in the efficiency of CSR. All other constructs showed levels of CSR very similar to those of the controls (**Figure 3.12 - 3.13**). Similar results were obtained analyzing the EGFP(+) population from transiently transfected CH12-F3 cells (**Figure 3.14**). Thus, our data suggest that overexpression of the catalytically inactive natural isoforms of AID is not able to modulate the activity of the endogenous AID, at least with regards to the Class Switch Recombination process.

Conclusions

A number of splice variants of AID have been identified in normal B cells and in lymphoproliferative diseases beyond the full-length one {Albesiano 2003; Greeve 2003; Oppezzo 2003; McCarthy 2003; Babbage 2004; Heintel 2004; Forconi 2004; Wu 2008; Marantidou 2010; Palacios 2010}, and their presence correlated to the prognosis of the disease {Albesiano 2003; McCarthy 2003; Oppezzo 2003; Babbage 2004; Marantidou 2010}. However, their functional significance is controversial {Wu 2008; van Maldegem 2009}. The absence of catalytic activity ({van Maldegem 2009}, our data) suggests that – if translated – their effect might be indirect, through their interplay with the pathways involved in the antigen-driven antibody diversification processes. Indeed, our results exclude that co-expression with endogenous AID substantially alters the efficiency of CSR (as opposed to the overexpression of full-length AID). Our results do not rule out the possibility that these alternatively spliced isoforms of AID could still play a role, either as transcripts or simply because their transcription competes with that of the full-length AID.

Appendices

A - Oligonucleotides

Oligonucleotides, were synthesized by the company PRIMM Biotech Services.

Name	Sequence	Enzyme(s)	Purpose
mycLA1f	aaaccatggaattcTGGGTTGGCAATTCACTCCT	NcoI-EcoRI	knock k-n
mycLA1nf	AGGGGAAAGCTTGGGTTTGT		knock k-n
mycLA1nr	taaccggccgctacattcaa		knock k-n
mycLA1r	accggggtttcGcatcgcccaaaggaaatccagcc		knock k-n
mycLA2f	cctttgggcgatgCgaaaccccggtaagcacaga		knock k-n
mycLA2nf	attgcagcgggcagacactt		knock k-n
mycLA2nr	gtcgaggtcatagttcctgt		knock k-n
mycLA2r	aaagctagccatgggcatcgtcgtggctgtctg	NheI-NcoI	knock k-n
mycRAf	aaagctagctcgaacagcttcgaaactc	NheI	knock k-n
mycRAnf	ttttgccctgcgtgaccagat		knock k-n
mycRAnr	TTCACCCCTCATCGTGATTG		knock k-n
mycRAnr	aaattgcggccgcTGGCCGATGTCCTGTTTTCA	NotI	knock k-n
mycprobeF	TAAACCGGGTAGCAGTGAGA		Southern blot
mycprobeR	AGGAGTGAATTGCCAACCCA		Southern blot
GM80_CIB1_F	GctagcCCTCTGCTAACCATGTTCATGCCT	NheI	
GM81_CIB1_R	GctAGCTGGGATCTGATCAATTCCG	NheI	
GM81_FokI_F	GctagcaCGGAATTGATCAGATCCCAGCT	NheI	
GM82_FokI_R	ctatagTGCACCTGAGGAGTGAATTCCA	EcoRV	
GM83_IRES_R	ggatccGGTATTATCGTGT		

GM_84_hAIDS3_3F	AAAGCTAGCACCATGGACGccCTCTTGATG		
GM99_cry2_F	tgcagggagctcaaaatgga		to check CRY2 expression
GM100_cry2_R	CGCGAGCCCAAATGATTTGCTT		to check CRY2 expression
GM102_HS_pr_F	AAAtctagaCGCCGGAGTATA	XbaI	to amplify HS promote
GM103_HS_pr_R	AAAaccggtTGAGTTCTTCTTT	AgeI	to amplify HS promote
GM104_IRES_F	AAAgaattcggatccGGCGCATCCTT	EcoRI-BamHI	to amplify IRES
GM105_IRES_R	aaaCTTGCTCACCATGGTATTAT	NcoI	to amplify IRES
GM106_BFP_F	AAAccatggacATGCACATGAAGCTGT	NcoI	to amplify BFP
GM107_BFP_R	AAAgaattcTTACTAGGGCTCTT	EcoRI	to amplify BFP
GM108_Kan_Neo_F	AAAAgcggccgcGGTGTGGAAAGTCCC	NotI	to amplify Neo/Kan resistance
GM109_Kan_Neo_R	AAAAgcggccgcACGGTATACAGA	NotI	to amplify Neo/Kan resistance
GM121_p53CrisprF	caccgAGTGAAGCCCTCCGAGTGTC		Knock-out of mouse TP53
GM122_p53CrisprR	aaacGACACTCGGAGGGCTTCACTc		Knock-out of mouse TP53
GM123_p53_BSR_F	GTCGAGTGAAGCCCTCCGAGTGTCagg		
GM124_p53_BSR_R	CAGAcctGACACTCGGAGGGCTTCACT		
GM125_dCas9F	aaaGAGCTCGTGAAAGTGATGGG	SacI	to amplify half Cas9
GM126_dCas9R	ctcgagaagcttggatccCCCGGGGAGCATGTC	BamHI-HindIII	to amplify half Cas9
GM127_KrabF	ggatccaagcttctcgagATGGATGCTAAG	BamHI-HindIII	to amplify KRAB from 48237
GM128_KrabR	AAAgagctcTTACTAGGGCTCTTCTCCCT	SacI	to amplify KRAB from 48237
GM129_p53_crispr_kdF	caccgCCTGCTGAGGGCAACATCTC		to downregulate p53 in mouse
GM130_p53_crispr_kdR	aaacGAGATGTTGCCCTCAGCAGGc		to downregulate p53 in mouse
GM131_mycCRISPRn0_F	caccgGTCCCCGCCTCCTTCGAGCA		CRISPR nick
GM132_mycCRISPRn0_R	caaaTGCTCGAAGGAGGCGGGGACc		CRISPR nick

GM133_mycCRISPRn5_F	caccgAAGTCCCTGCTCGAAGGAGG		CRISPR nick
GM134_mycCRISPRn5_R	caaaCCTCCTTCGAGCAGGGACTTc		CRISPR nick
GM135_mycCRISPRn14_F	caccgGCTCGAAGGAGGCGGGGACT		CRISPR nick
GM136_mycCRISPRn14_R	caaaAGTCCCCGCCTCCTTCGAGCc		CRISPR nick
GM137_mycCRISPRn20_F	caccgGCGATTCCTGGGCGTCTGCA		CRISPR nick
GM138_mycCRISPRn20_R	caaaTGCAGACGCCCAGGAATCGCc		CRISPR nick
Manuscript Primers			
AID_F_{pTrc99}	AAAAAGCTTACCATGGACAGCCTCTTGATG	SnaBI	to amplify and clone AID into pTrc99
AID_R_{pTrc99}	TTCTCGAGTCAAAGTCCCAAAGTACGAAATG	XhoI	to amplify and clone AID into pTrc99
AID_F_{pAID}	AAAGCTAGCACCATGGACAGCCTCTTGATG	BglII	to amplify clone AID into pAID-expreass-puro2
AID_R_{pAID}	AAAGCTAGCACCATGGACAGCCTCTTGATG	NheI	to amplify clone AID into pAID-expreass-puro2
AID_E58A_F	CTGCCACGTGGCATTGCTCTTCCTCCGC		to amplify AID_E58A
AID_E58A_R	GGAAGAGCAATGCCACGTGGCAGCCGTT		to amplify AID_E58A
murine AID_F	CTCCTGCTCACTGGACTTCG		Quantitative real-time PCR (qPCR)
murine AID_R	AGGCTGAGGTTAGGGTTCCA		Quantitative real-time PCR (qPCR)
human AID,_F	CAGCCTCTTGATGAACCGGA		Quantitative real-time PCR (qPCR)
human AID,_R	CGTGGCAGCCGTTCTTATTG		Quantitative real-time PCR (qPCR)
β-Actin_F	GGCTCCTAGCACCATGAAG		Quantitative real-time PCR (qPCR)
β-Actin_R	GAAAGGGTGTAAAACGCAGC		Quantitative real-time PCR (qPCR)

B - Plasmids and vectors

Plasmids and vectors used in this work

Plasmid	Feature(s) and usage	Marker (s)	Source
pCR-BluntII TOPO	routine cloning of blunt PCR products	Kan	Invitrogen
Zero Blunt TOPO	routine cloning of sticky PCR products	Amp / Kan	Invitrogen
pBuerstedde	expression vector for eukaryotic cells		
pBS-SK(+)	expression vector for knock-in		
pIND-EGFP	ponasterone inducible system	Amp	
LITE 2.0_pAAV-EF1a-NLS-C2P-NLS-VP64_2A_EGFP	LITE inducible system	Amp	{Konermann and Brigham, 2013}
pCIB1(deltaNLS)-pmEGFP	LITE inducible system	Kan	{Konermann and Brigham, 2013}
pLITE2.0woVP64-EGFP/ FokI			
pX330			CRISPR vector
pTrc99a			

References

Agrawal A, Eastman QM, Schatz DG (**1998**). Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. ***Nature*** 394: 744-751.

Albesiano E, Messmer BT, Damle RN, Allen SL, Rai KR, Chiorazzi N (**2003**). Activation-induced cytidine deaminase in chronic lymphocytic leukemia B cells: expression as multiple forms in a dynamic, variably sized fraction of the clone. ***Blood*** 102: 3333-3339.

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

Arakawa H, Hauschild J, and Buerstedde JM (**2002**). Requirement of the activation-induced deaminase (AID) gene for immunoglobulin gene conversion. ***Science*** 295: 1301-06.

Babbage G, Garand R, Robillard N, Zojer N, Stevenson FK, Sahota SS (**2004**). Mantle cell lymphoma with t(11;14) and unmutated or mutated VH genes expresses AID and undergoes isotype switch events. ***Blood*** 103: 2795-2798.

Basu U, Chaudhuri J, Alpert C, Dutt S, Ranganath S (**2005**). The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. ***Nature*** 438: 508-511.

Basu U, Meng FL, Keim C, Grinstein V, Pefanis E, Eccleston J, Zhang T, Myers D, Wasserman CR, Wesemann DR (**2011**). The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. ***Cell*** 144: 353-363.

Barreto V, Reina-San-Martin B, Ramiro AR, McBride KM and Nussenzweig MC (**2003**). C-terminal deletion of AID uncouples Class Switch Recombination from Somatic Hypermutation and gene conversion. ***Mol. Cell*** 12: 501-8.

Barreto VM, Pan-Hammarstrom Q, Zhao Y, Hammarstrom L, Misulovin Z, Nussenzweig MC (**2005**). AID from bony fish catalyzes Class Switch Recombination. ***J Exp Med*** 202(6):733-8.

Beale RC, Petersen-Mahrt SK, Watt IN, Harris RS, Rada C and Neuberger MS (**2004**). Comparison of the differential context-dependence of DNA deamination by APOBEC enzymes: correlation with mutation spectra in vivo. ***J Mol Biol*** 337: 585-596.

Betts L, Xiang S, Short SA, Wolfenden R and Carter CW (**1994**). Cytidine deaminase. The 2.3 A crystal structure of an enzyme: transition-state analog complex. ***J Mol Biol*** 235: 635-656.

Biondi A, Valsecchi MG, Seriu T, D'Aniello E, Willemse MJ, Fasching K, Pannunzio A, Gadner, H, Schrappe M (**2000**). Molecular detection of minimal residual disease is a strong predictive factor of relapse in childhood B-lineage acute lymphoblastic leukemia with medium risk features. A case control study of the International BFM study group. ***Leukemia*** 14: 1939-943.

Bishop KN, Holmes RK, Sheehy AM, Davidson NO, Cho SJ, Malim MH (**2004**). Cytidine deamination of retroviral DNA by diverse APOBEC proteins. ***Curr Biol*** 14(15):1392-6.

Black DL (**2003**). Mechanisms of alternative pre-messenger RNA splicing. ***Annu Rev Biochem*** 72:291-336.

Boboila C,Jankovic M, Yan CT, Wang JH, Wesemann DR, Zhang T, Fazeli A, Feldman L, Nussenzweig A, Nussenzweig M and Alt FW (**2010**). Alternative end-joining catalyzes robust IgH locus deletions and translocations in the combined absence of ligase 4 and Ku70. ***Proc. Natl. Acad. Sci. USA*** 107: 3034–3039.

Boboila C, Oksenycha V, Gostissaa M, Wanga JH,Zhaa S, Zhanga Y, Chaia H, Leea CS, Jankovicb M, Albertorio Saezc LM, Nussenzweig MC, McKinnond PJ, Alt FW and Schwer B (**2011**). Robust chromosomal DNA repair via alternative end-joining in the absence of X-ray repair cross-complementing protein 1 (XRCC1). Proc ***Natl Acad Sci USA*** 109(7): 2473–2478.

Boerma G, Siebert R, Kluin PM and Baudis M (**2009**). Translocations involving 8q24 in Burkitt lymphoma and other malignant lymphomas: a historical review of cytogenetics in the light of todays knowledge. ***Leukemia*** 23: 225–234.

Bödör C, Bognár A, Reiniger L, Szepesi A, Tóth E, Kopper L and Matolcsy A (**2005**). Aberrant Somatic Hypermutation and expression of activation-induced cytidine deaminase mRNA in mediastinal large B-cell lymphoma. ***Br J Haematol*** 129: 373-76.

Brandsma I, Gent DC (**2012**). Pathway choice in DNA double strand break repair: observations of a balancing act. **Genome Integr.** 3(1):9.

Brar SS, Watson M and Diaz M (**2004**). Activation-induced cytosine deaminase (AID) is actively exported out of the nucleus but retained by the induction of DNA breaks. T***he Journal of biological chemistry*** 279: 26395-26401.

Buerstedde JM, Lowndes N, Schatz DG (**2014**). Induction of homologous recombination between sequence repeats by the activation induced cytidine deaminase (AID) protein. ***Elife.***

Busch K, Keller T, Fuchs U, Yeh RF, Harbott J, Klose I, Wiemels J, Novosel A, Reiter A Borkhardt A (**2007**). Identification of two distinct MYC breakpoint clusters and their association with various IGH breakpoint regions in the t(8;14) translocations in sporadic Burkitt-lymphoma. ***Leukemia*** (8): 1739-51.

Capello D, Vitolo, U., Pasqualucci, L., Quattrone, S., Migliaretti, G., Fassone, L., Ariatti C, Vivenza, D., Gloghini, A (**2000**). Distribution and pattern of BCL-6 mutations throughout the spectrum of B-cell neoplasia. ***Blood*** 95: 651-59.

Carter CW (**1995**). The nucleoside deaminases for cytidine and adenosine: structure, transition state stabilization, mechanism, and evolution. ***Biochimie*** 77: 92-98.

Casellas R, Nussenzweig A, Wuerffel R, Pelanda R, Reichlin A, Suh H, Qin XF, Besmer E, Kenter A, Rajewsky K and Nussenzweig MC (**1998**). Ku80 is required for immunoglobulin isotype switching. ***EMBO J.*** 17: 2404– 2411.

Chapelle de la CA, Lenoir G, Boue J, Boue A, Gallano P, Huerre C (**1983**). Lambda Ig constant region genes are translocated to chromosome 8 in Burkitt's lymphoma with t(8;22). ***Nucleic Acids Res*** 11: 1133–1142.

Chaudhuri J, Tian M, Khuong C, Chua K, Pinaud E and Alt FW (**2003**). Transcription- targeted DNA deamination by the AID antibody diversification enzyme. ***Nature*** 422: 726- 730.

Chaudhuri J, Khuong C and Alt FW (**2004**). Replication protein A interacts with AID to promote deamination of Somatic Hypermutation targets. ***Nature*** 430: 992-98.

Chaudhuri J, Basu U, Zarrin A, Yan C, Franco S, Perlot T, Vuong B, Wang J, Phan RT (**2007**). Evolution of the immunoglobulin heavy chain Class Switch Recombination mechanism. ***Adv Immunol*** 94: 157-214.

Chen Z, Gu J (**2007**). Immunoglobulin G expression in carcinomas and cancer cell lines. ***FASEB J.*** 21(11):2931-8.

Cheng HL, Vuong BQ, Basu U, Franklin A, Schwer B, Astarita J, Phan RT, Datta A, Manis J, Alt FW (**2009**): Integrity of the AID serine-38 phosphorylation site is critical for Class Switch Recombination and Somatic Hypermutation in mice. ***Proc. Natl. Acad. Sci. USA*** 106:2717-2722.

Chiarle R, Zhang Y, Frock RL, Lewis SM, Molinie B, Ho YJ, Myers DR, Choi VW, Compagno M, Malkin DJ (**2011**). Genome-wide translocation sequencing reveals mechanisms of chromosome breaks and rearrangements in B cells. ***Cell*** 147: 107-119.

Cole MD (**1986**). Activation of the c-myc oncogene. ***Annu Rev Genet.*** 20:361-84.

Conticello SG, Harris RS, Neuberger MS (**2003**). The Vif protein of HIV triggers degradation of the human antiretroviral DNA deaminase APOBEC3G. ***Curr Biol.*** 13(22):2009-13.

Conticello, **SG**, Thomas CJ, Petersen-Mahrt, SK and Neuberger, MS (**2005**). Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. ***Mol Biol Evol*** 22: 367-377.

Conticello SG, Langlois MA, Yang Z and Neuberger MS (**2007**). DNA Deamination in Immunity: AID in the Context of Its APOBEC Relatives. ***Adv Immunol*** 94: 37-73.

Conticello SG, Ganesh K, Xue K, Lu M, Rada C, Neuberger MS (**2008**) Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNBL1. ***Mol. Cell*** 31: 474-484.

Cobb RM, Oestreich KJ, Osipovich OA, Oltz EM (**2006**). Accessibility Control of V(D)J Recombination. ***Adv Immunol.*** 91:45-109.

Correia C, Schneider PA, Dai H, Dogan A, Maurer MJ, Church AK, Novak AJ, Feldman AL, Wu X, Ding H, Meng XW, Cerhan JR, Slager SL, Macon WR, Habermann TM, Karp JE, Gore SD, Kay NE, Jelinek DF, Witzig TE, Nowakowski GS, Kaufmann SH (**2015**). BCL2 mutations are associated with increased risk of transformation and shortened survival in follicular lymphoma. ***Blood*** 125(4):658-67.

Cortizas EM, Zahn A, Hajjar ME, Patenaude AM, Di Noia JM, Verdun RE (**2013**). Alternative end-joining and classical nonhomologous end-joining pathways repair different types of double-strand breaks during class-switch recombination. ***J Immunol.*** 191(11):5751-63.

Dalla-Favera R, Bregni M, Erikson J, Patterson D, Gallo RC, Croce CM (**1982**). Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. ***Proc. Natl. Acad. Sci. USA*** 79: 7824–7827.

Dang CV, Resar LM, Emison E, Kim S, Li Q, Prescott JE, Wonsey D, Zeller K (**1999**). Function of the c-Myc oncogenic transcription factor. ***Exp Cell Res.*** 253(1):63-77.

Delbos F, De Smet A, Faili A, Aoufouchi S, Weill JC and Reynaud CA (**2005**). Contribution of DNA polymerase eta to immunoglobulin gene hypermutation in the mouse. ***J Exp Med*** 201, 1191-96.

Deutsch, AJ, Frühwirth M, Aigelsreiter A, Cerroni L and Neumeister P (**2009**). Primary cutaneous marginal zone B-cell lymphomas are targeted by aberrant Somatic Hypermutation. J ***Invest Dermatol*** 129: 476-79.

Di Noia J and Neuberger MS (**2002**). Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. ***Nature*** 419: 43-48.

Di Noia JM and Neuberger MS (**2007**). Molecular Mechanisms of Antibody Somatic Hypermutation. ***Annu Rev Biochem*** 76: 1-22.

Doetsch PW, Cunningham RP (**1990**).The enzymology of apurinic/apyrimidinic endonucleases. ***Mutat Res.*** 236(2-3):173-201.

Dorsett Y, Robbiani DF, Jankovic M, Reina-San-Martin B, Eisenreich TR, Nussenzweig MC (**2007**) A role for AID in chromosome translocations between c-myc and the IgH variable region. ***J Exp Med*** 204: 2225-2232.

Dorsett Y, McBride KM, Jankovic M, Gazumyan A,Thai TH, Robbiani DF, Di Virgilio M, San-Martin BR, Heidkamp G, Schwickert TA (**2008**): MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. ***Immunity*** 28:630-638.

DuBridge RB, Tang P, Hsia HC, Leong PM, Miller JH, Calos MP (**1987**) Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. ***Mol. Cell Biol*** 7: 379-387.

Duquette ML, Huber MD, Maizels N (**2007**). G-rich proto-oncogenes are targeted for genomic instability in B-cell lymphomas. ***Cancer Res.*** 67(6):2586-94.

Ellyard JI, Benk AS, Taylor B, Rada C, Neuberger MS (**2011**) The dependence of Ig class-switching on the nuclear export sequence of AID likely reflects interaction with factors additional to Crm1 exportin. ***Eur J Immunol*** 41: 485-490.

Endo Y, Marusawa H, Kinoshita K, Morisawa T, Sakurai T, Okazaki IM, Watashi K, Shimotohno K, Honjo T, Chiba T (**2007**). Expression of activation-induced cytidine deaminase in human hepatocytes via NF-kappaB signaling. ***Oncogene*** 26(38):5587-95.

Erikson J, Finan J, Nowell PC, Croce CM (**1982**). Translocation of immunoglobulin VH genes in Burkitt lymphoma. ***Proc. Natl. Acad. Sci. USA*** 79: 5611–5615.

Escot C, Theillet C, Lidereau R, Spyratos F, Champeme MH, Gest J and Callahan R (**1986**). Genetic alteration of the c-myc protooncogene (MYC) in human primary breast carcinomas. ***Proc. Natl. Acad. Sci. USA*** 83: 4834–4838.

Doi T, Kato L, Ito S, Shinkura R, Wei M, Nagaoka H, Wang J, Honjo T (**2009**). The C-terminal region of activation-induced cytidine deaminase is responsible for a recombination function other than DNA cleavage in Class Switch Recombination. ***Proc. Natl. Acad. Sci. USA*** 106(8): 2758-63.

Forconi F, Sahota SS, Raspadori D, Ippoliti M, Babbage G (**2004**). Hairy cell leukemia: at the crossroad of somatic mutation and isotype switch. Blood 104: 3312-3317.

Frey S, Bertocci B, Delbos F, Quint L, Weill JC and Reynaud CA (**1998**). Mismatch repair deficiency interferes with the accumulation of mutations in chronically stimulated B cells and not with the hypermutation process. ***Immunity*** 9, 127-134.

Gaidano G, Pasqualucci L, Capello D., Berra E, Deambrogi C., Rossi D., Maria Larocca L, Gloghini A, Carbone A and Dalla-Favera R (**2003**). Aberrant Somatic Hypermutation in multiple subtypes of AIDS-associated non-Hodgkin lymphoma. ***Blood*** 102: 1833-841.

Geisberger R, Rada C and Neuberger MS (**2009**). The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. ***Proceedings of the National Academy of Sciences of the United States of America*** 106: 6736-6741.

Greeve J, Philipsen A, Krause K, Klapper W, Heidorn K, Castle BE, Janda J, Marcu KB and Parwaresch R (**2003**). Expression of activation-induced cytidine deaminase in human B-cell non-Hodgkin lymphomas. ***Blood*** 101: 3574-580.

Greisman HA, Lu Z, Tsai AG, Greiner TC, Yi HS, Lieber MR. (**2012**). IgH partner breakpoint sequences provide evidence that AID initiates t(11;14) and t(8;14) chromosomal breaks in mantle cell and Burkitt lymphomas. **Blood** 120(14):2864-7.

Guirouilh-Barbat J, Rass E, Plo I, Bertrand P, Lopez BS (**2007**). Defects in XRCC4 and KU80 differentially affect the joining of distal nonhomologous ends. ***Proc. Natl. Acad. Sci. USA***. 104(52):20902-7.

John P Guilinger, David B Thompson & David R Liu (**2004**). Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. ***Nature Biotechnology*** 32: 577–582.

Hakim O, Resch W, Yamane A, Klein I, Kieffer-Kwon KR, Jankovic M, Oliveira T, Bothmer A, Voss TC, Ansarah-Sobrinho C, Mathe E, Liang G, Cobell J, Nakahashi H, Robbiani DF, Nussenzweig A, Hager GL, Nussenzweig MC, Casellas R (**2012**). DNA damage defines sites of recurrent chromosomal translocations in B lymphocytes. ***Nature*** 484(7392):69-74.

Hanahan D and Weinberg RA (**2000**). The hallmarks of cancer. ***Cell*** 100: 57-70.

Harris RS, Petersen-Mahrt SK and Neuberger MS (**2002**). RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. ***Mol. Cell*** 10: 1247-253.

Harris R and Liddament MT (**2004**). Retroviral restriction by APOBEC proteins. ***Nature Reviews Immunology*** 4: 868-877.

Hasham MG, Donghia NM, Coffey E, Maynard J, Snow KJ, Ames J, Wilpan RY, He Y, King BL and Mills KD (**2010**). Widespread genomic breaks generated by activation-induced cytidine deaminase are prevented by homologous recombination. ***Nat. Immunol.*** 11: 820-26.

Heintel D, Kroemer E, Kienle D, Schwarzinger I, Gleiss A (**2004**). High expression of activation-induced cytidine deaminase (AID) mRNA is associated with unmutated IGVH gene status and unfavourable cytogenetic aberrations in patients with chronic lymphocytic leukaemia. ***Leukemia*** 18: 756-762.

Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-André V, Sigova AA, Hoke HA, Young RA. (**2013**). Super-enhancers in the control of cell identity and disease. ***Cell*** 155(4):934-47.

Iacobucci I, Lonetti A, Messa F, Ferrari A, Cilloni D (**2010**) Different isoforms of the B-cell mutator activation-induced cytidine deaminase are aberrantly expressed in BCR-ABL1-positive acute lymphoblastic leukemia patients. **Leukemia** 24: 66-73.

Ikeda J, Mamat S, Tian T, Wang Y, Rahadiani N, Aozasa K, Morii E (**2010**). Tumorigenic potential of mononucleated small cells of Hodgkin lymphoma cell lines. ***Am. J. Pathol.*** 177(6): 3081-8.

Ireton GC, Black ME and Stoddard BL (**2003**). The 1.14 A crystal structure of yeast cytosine deaminase: evolution of nucleotide salvage enzymes and implications for genetic chemotherapy. ***Structure*** 11: 961-972.

Ito S, Nagaoka H, Shinkura R, Begum N, Muramatsu M, Nakata M, Honjo T (**2004**). Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. ***Proceedings of the National Academy of Sciences of the United States of America*** 101: 1975-80.

Karanam K, Kafri R, Loewer A, Lahav G (**2012**). Quantitative live cell imaging reveals a gradual shift between DNA repair mechanisms and a maximal use of HR in mid S phase. ***Mol. Cell*** 47(2):320-9.

Kersey JH (**1997**). Fifty years of studies of the biology and therapy of childhood leukemia. ***Blood*** 90: 4243-251.

Kim CJ, Song JH, Cho YG, Cao Z, Kim SY, Nam SW, Lee JY, Park WS (**2007**) Activation-induced cytidine deaminase expres- sion in gastric cancer. ***Tumour Biol.*** 28:333–339.

Klein IA, Resch W, Jankovic M, Oliveira T, Yamane A, Nakahashi H, Di Virgilio M, Bothmer A, Nussenzweig A, Robbiani DF (**2011**). Translocation-capture sequencing reveals the extent and nature of chromosomal rearrangements in B lymphocytes. ***Cell*** 147: 95-106.

Ko TP, Lin JJ, Hu CY, Hsu YH, Wang AH, Liaw SH (**2003**). Crystal structure of yeast cytosine deaminase. Insights into enzyme mechanism and evolution. ***J. Biol. Chem.*** 278(21):19111-7.

Komori J, Marusawa H, Machimoto T, Endo Y, Kinoshita K, Kou T, Haga H, Ikai I, Uemoto S, Chiba T (**2008**). Activation-induced cytidine deaminase links bile duct inflammation to human cholangiocarcinoma. ***Hepatology*** 47(3):888-96.

Kotani A, Kakazu N, Tsuruyama T, Okazaki IM, Muramatsu M (**2007**). Activation-induced cytidine deaminase (AID) promotes B cell lymphomagenesis in Emu-cmyc transgenic mice. ***PNAS*** 104: 1616-1620.

Kotnis A, Du L, Liu C, Popov SW and Pan-Hammarström Q (**2009**). Non-homologous end joining in Class Switch Recombination: the beginning of the end. ***Philos Trans R Soc Lond B Biol Sci*** 364: 653-665.

Krejci L, Altmannova V, Spirek M, Zhao X (**2012**). Homologous recombination and its regulation. ***Nucleic Acids Res.*** 40(13):5795-818.

Küppers R, Rajewsky K, Zhao M, Simons G, Laumann R, Fischer R, Hansmann ML (**1994**). Hodgkin disease: Hodgkin and Reed-Sternberg cells picked from histological sections show clonal immunoglobulin gene rearrangements and appear to be derived from B cells at various stages of development. ***Proc. Natl. Acad. Sci. USA*** 91(23):10962-6.

Küppers R, Dalla-Favera R (**2001**). Mechanisms of chromosomal translocations in B cell lymphomas. ***Oncogene*** 20(40):5580-94

Küppers R, Schmitz R, Distler V, Renné C, Bräuninger A, Hansmann ML (**2005**). Pathogenesis of Hodgkin's lymphoma. ***Eur J Haematol Suppl.*** (66):26-33.

Jang SK, Kräusslich HG, Nicklin MJ, Duke GM, Palmenberg AC, Wimmer E (**1988**). A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. ***J. Virol.*** 62 (8): 2636–43.

Jansen JG, Langerak P, Tsaalbi-Shtylik A, van den Berk, P, Jacobs H and de Wind N. (**2006**). Strand-biased defect in C/G transversions in hypermutating immunoglobulin genes in Rev1-deficient mice. J **Exp Med** 203: 319-323.

Jardin F, Bastard C, Contentin N, Parmentier F, Picquenot JM, Tilly H, Stevenson FK and Sahota SS (**2003**). Intronic BCL-6 mutations are preferentially targeted to the translocated allele in t(3;14)(q27;q32) non-Hodgkin B-cell lymphoma. ***Blood*** 102: 1872-76.

Jasin M (**2000**). Chromosome breaks and genomic instabilit. **Cancer Invest.** 18(1):78-86.

LeBien TW, Tedder TF (**2008**). B lymphocytes: how they develop and function. ***Blood*** 112(5): 1570-80.

Lee-Theilen M, Matthews AJ, Kelly D, Zheng S and Chaudhuri J. (**2011**). CtIP promotes microhomology-mediated alternative end joining during class- switch recombination. ***Nat. Struct. Mol. Biol.*** 18: 75–79.

Lieber MR (**2010**). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. ***Annu. Rev. Biochem.*** 79: 181–211.

Liao W, Hong SH, Chan BH, Rudolph FB, Clark SC, Chan L (**1999**). APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. ***Biochem Biophys Res Commun.*** 260(2):398-404.

Liddament MT, Brown WL, Schumacher AJ, Harris RS (**2004**). APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo. ***Curr Biol.***14(15):1385-91.

Liso A, Capello D, Marafioti T, Tiacci E, Cerri M, Distler V, Paulli M, Carbone A, Delsol G (**2006**). Aberrant Somatic Hypermutation in tumor cells of nodular-lymphocyte-predominant and classic Hodgkin lymphoma. ***Blood*** 108: 1013-020.

Little CD, Nau MM, Carney DN, Gazdar AF and Minna JD (**1983**). Amplification and expression of the c-myc oncogene in human lung cancer cell lines. ***Nature*** 306: 194– 196.

Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC, Kleinstein SH and Schatz DG (**2008**). Two levels of protection for the B cell genome during Somatic Hypermutation. ***Nature*** 451: 841-45.

Look AT (**1997**). E2A-HLF chimeric transcription factors in pro-B cell acute lymphoblastic leukemia. ***Curr Top Microbiol Immunol*** 220: 45-53.

Lossos IS, Warnke R and Levy R (**2002**). BCL-6 mRNA expression in higher grade transformation of follicle center lymphoma: correlation with somatic mutations in the 5' regulatory region of the BCL-6 gene. ***Leukemia*** 16: 1857-862.

Lu Z, Tsai AG, Akasaka T, Ohno H, Jiang Y, Melnick AM, Greisman HA, Lieber MR (**2013**). BCL6 breaks occur at different AID sequence motifs in Ig-BCL6 and non-Ig-BCL6 rearrangements. ***Blood*** 121(22):4551-4.

Malu S, Malshetty V, Francis D, Cortes P (2012). Role of non-homologous end joining in V(D)J recombination. ***Immunol Res.*** 54(1-3):233-46. doi: 10.1007/s12026-012-8329-z.

Malcolm S, Barton P, Murphy C, Ferguson-Smith MA, Bentley DL, Rabbitts TH (**1982**). Localization of human immunoglobulin kappa light chain variable region genes to the short arm of chromosome 2 by in situ hybridization. ***Proc. Natl. Acad. Sci. USA*** 79: 4957–4961.

Mangeat B, Turelli P, Caron G, Friedli M, Perrin L, Trono D (**2003**). Broad antiretroviral defence by human APOBEC3G through lethal editing of nascent reverse transcripts. ***Nature*** 424(6944):99-103.

Mansour WY, Rhein T, Dahm-Daphi J (**2010**). The alternative end-joining pathway for repair of DNA double-strand breaks requires but is not dependent upon microhomologies. ***Nucleic Acids Res.***38(18):6065-77.

Marantidou F, Dagklis A, Stalika E, Korkolopoulou P, Saetta A (**2010**) Activation-induced cytidine deaminase splicing patterns in chronic lymphocytic leukemia. ***Blood Cells Mol Dis*** 44: 262-267.

Marusawa H (**2008**) Aberrant AID expression and human cancer development. ***Int J Biochem Cell Biol*** 40:1399–1402.

Matsumoto Y, Marusawa H, Kinoshita K, Endo Y, Kou T, Morisawa T, Azuma T, Okazaki IM, Honjo T and Chiba T (**2007**). Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. ***Nat Med*** 13, 470-76.

McBride KM, Gazumyan A, Woo EM, Barreto VM, Robbiani DF (**2006**). Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. ***PNAS*** 103: 8798-8803.

McBride KM, Gazumyan A, Woo EM, Schwickert TA, Chait BT, Nussenzweig MC (**2008**) Regulation of Class Switch Recombination and somatic mutation by AID phosphorylation. ***J Exp Med***205:2585-2594.

McCarthy H, Wierda WG, Barron LL, Cromwell CC, Wang J, Coombes KR, Rangel R, Elenitoba-Johnson KS, Keating MJ and Abruzzo LV (**2003**). High expression of activation-induced cytidine deaminase (AID) and splice variants is a distinctive feature of poor-prognosis chronic lymphocytic leukemia. ***Blood*** 101: 4903-08.

Meng FL, Du Z, Federation A, Hu J, Wang Q, Kieffer-Kwon KR, Meyers RM, Amor C, Wasserman CR, Neuberg D, Casellas R, Nussenzweig MC, Bradner JE, Liu XS, Alt FW (**2014**). Convergent transcription at intragenic super-enhancers targets AID-initiate genomic instability. ***Cell*** 159(7):1538-48.

Menssen A, Hermeking H (**2002**). Characterization of the c-MYC-regulated transcriptome by SAGE: identification and analysis of c-MYC target genes. ***Proc. Natl. Acad. Sci. USA.*** 99(9): 6274-9.

Mills KD, Ferguson DO, Alt FW (**2003**). The role of DNA breaks in genomic instability and tumorigenesis. **Immunol Rev.**194:77-95.

Mitelman F, Johansson B and Mertens F (**2015**)Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer.

Mol CD, Hosfield DJ, Tainer JA (**2000**). Abasic site recognition by two apurinic/apyrimidinic endonuclease families in DNA base excision repair: the 3' ends justify the means. ***Mutat Res.*** 460(3-4):211-29.

Montesinos-Rongen M, Van Roost D, Schaller C, Wiestler OD and Deckert M (**2004**). Primary diffuse large B-cell lymphomas of the central nervous system are targeted by aberrant Somatic Hypermutation. ***Blood*** 103: 1869-875.

Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK (**2004**). Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. ***J Biol Chem.*** 279(50):52353-60.

Martomo SA, Fu, D, Yang WW, Joshi NS and Gearhart PJ (**2005**). Deoxyuridine is generated preferentially in the nontranscribed strand of DNA from cells expressing activation-induced cytidine deaminase. **J Immunol** 174, 7787-791.

Munzel P, Marx, D, Kochel H, Schauer A and Bock KW (**1991**). Genomic alterations of the c-myc protooncogene in relation to the overexpression of c-erbB2 and Ki-67 in human breast and cervix carcinomas. ***J. Cancer Res. Clin. Oncol.*** 117: 603– 607.

Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO and Honjo T (**1999**). Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. ***J. Biol.Chem.*** 274: 18470-76.

Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y and Honjo T (**2000**). Class Switch Recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. ***Cell*** 102: 553-563.

Müschen M, Re D, Bräuninger A, Wolf J, Hansmann ML, Diehl V, Küppers R, Rajewsky K. (**2000**). Somatic mutations of the CD95 gene in Hodgkin and Reed-Sternberg cells. ***Cancer Res.*** 60(20):5640-3.

Nakamura M, Kondo S, Sugai M, Nazarea M, Imamura S, Honjo T (**1996**). ***Int Immunol*** 8: 193-201.

Nambiar M, Kari V, Raghavan SC (**2008**). Chromosomal translocations in cancer. ***Biochim Biophys Acta.*** 1786(2):139-52.

Navaratnam N, Morrison JR, Bhattacharya S, Patel D, Funahashi T, Giannoni F, Teng BB, Davidson NO, Scott J (**1993**). The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. ***J. Biol. Chem.*** 268(28):20709-12.

Nelson JR, Lawrence CW and Hinkle DC (**1996**). Deoxycytidyl transferase activity of yeast REV1 protein. ***Nature*** 382: 729-731.

Neuberger MS (**2008**) Antibody diversification by somatic mutation: from Burnet onwards. ***Immunol Cell Biol*** 86: 124-132.

Noguchi E, Shibasaki M, Inudou M, Kamioka M, Yokouchi Y (**2001**). Association between a new polymorphism in the activation-induced cytidine deaminase gene and atopic asthma and the regulation of total serum IgE levels. ***J Allergy Clin Immunol*** 108: 382-386.

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

Oettinger MA, Schatz DG, Gorka C and Baltimore D (**1990**). RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. ***Science*** 248, 1517-523.

Okazaki IM, Kotani A, Honjo T (**2007**) Role of AID in Tumorigenesis. ***Adv Immunol*** 94: 245-273.

Ono M and Nose M (**2007**). Persistent expression of an unproductive immunoglobulin heavy chain allele with DH-JH-gamma configuration in peripheral tissues. ***APMIS*** 115: 1350-56.

Oppezzo P, Vuillier F, Vasconcelos Y, Dumas G, Magnac C, Payelle-Brogard B, Pritsch, O and Dighiero G (**2003**). Chronic lymphocytic leukemia B cells expressing AID display dissociation between Class Switch Recombination and Somatic Hypermutation. ***Blood*** 101: 4029-032.

Orthwein A, Patenaude AM, Affar EB, Lamarre A, Young JC, Di Noia JM (**2010**) Regulation of activation-induced deaminase stability and antibody gene diversification by Hsp90. ***J. Exp. Med*** 207: 2751-2765.

Palacios F, Moreno P, Morande P, Abreu C, Correa A (**2010**). High expression of AID and active Class Switch Recombination might account for a more aggressive disease in unmutated CLL patients: link with an activated microenvironment in CLL disease. ***Blood*** 115: 4488-4496.

Pan-Hammarström Q, Jones AM, Lahdesmaki A, Zhou W, Gatti RA, Hammarstrom L, Gennery, AR and Ehrenstein MR (**2005**). Impact of DNA ligase IV on nonhomologous end joining pathways during Class Switch Recombination in human cells. ***The Journal of experimental medicine*** 201: 189-194.

Parker SC, Stitzel ML, Taylor DL, Orozco JM, Erdos MR, Akiyama JA, van Bueren KL, Chines PS, Narisu N, Black BL, Visel A, Pennacchio LA, Collins FS (**2013**). Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. ***Proc. Natl. Acad. Sci. USA.*** 110(44):17921-6.

Parsa JY, Basit W, Wang CL, Gommerman JL, Carlyle JR, Martin A (**2007**). AID mutates a non-immunoglobulin transgene independent of chromosomal position. ***Mol Immunol*** 44: 567-575.

Pasqualucci L, Migliazza A, Fracchiolla N, William C, Neri A, Baldini L, Chaganti RS, Klein U, Küppers R (**1998**). BCL-6 mutations in normal germinal center B cells: evidence of Somatic Hypermutation acting outside Ig loci. ***PNAS*** 95: 11816-821.

Pasqualucci L, Neri A, Baldini L, Dalla-Favera R and Migliazza A (**2000**). BCL-6 mutations are associated with immunoglobulin variable heavy chain mutations in B-cell chronic lymphocytic leukemia. ***Cancer Res*** 60: 5644-48.

Pasqualucci L, Neumeister P, Goossens T, Nanjangud G, Chaganti RS (**2001**) Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. ***Nature*** 412: 341-346.

Pasqualucci L, Kitaura Y, Gu H, Dalla-Favera R (**2006**), PKA-mediated phosphorylation regulates the function of activation-induced deaminase (AID) in B cells. ***PNAS*** 103: 395-400.

Patenaude AM, Orthwein A, Hu Y, Campo VA, Kavli B, Buschiazzo A, Di Noia JM (**2009**). Active nuclear import and cytoplasmic retention of activation-induced deaminase. ***Nat Struct Mol Biol*** 16:517-527.

Pauklin S, Sernández IV, Bachmann G, Ramiro AR, Petersen-Mahrt SK (**2009**). Estrogen directly activates AID transcription and function. ***J Exp Med*** 16;206(1):99-111.

Paul K, Wang M, Mladenov E, Bencsik-Theilen A, Bednar T, Wu W, Arakawa H, Iliakis G (**2013**). DNA ligases I and III cooperate in alternative non-homologous end-joining in vertebrates. ***PLoS One*** 8(3):e59505.

Pavri R, Gazumyan A, Jankovic M, Di Virgilio M, Klein I, Ansarah-Sobrinho C, Resch W, Yamane A, Reina San-Martin B, Barreto V, Nieland TJ, Root DE, Casellas R, Nussenzweig MC (**2010**). Activation-Induced Cytidine Deaminase Targets DNA at Sites of RNA Polymerase II Stalling by Interaction with Spt5. ***Cell*** 143: 122-33.

Peitz M, Pfannkuche K, Rajewsky K, Edenhofer F (**2002**). Ability of the hydrophobic FGF and basic TAT peptides to promote cellular uptake of recombinant Cre recombinase: a tool for efficient genetic engineering of mammalian genomes. ***Proc Natl Acad Sci USA*** 99(7):4489-94.

Peled JU, Kuang FL, Iglesias-Ussel MD., Roa S, Kalis SL, Goodman MF and Scharff MD (**2008**). The biochemistry of Somatic Hypermutation. ***Annu Rev Immunol*** 26: 481-511.

Pelletier J, Sonenberg N (**1988**). "Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA". ***Nature*** 334 (6180): 320–5

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

Petersen-Mahrt SK and Neuberger MS (**2003**). In vitro deamination of cytosine to uracil in single-stranded DNA by apolipoprotein B editing complex catalytic subunit 1 (APOBEC1). ***J. Biol. Chem.*** 278: 19583-86.

Pham P, Bransteitter R, Petruska J and Goodman MF (**2003**). Processive AID-catalysed cytosine deamination on single-stranded DNA simulates Somatic Hypermutation. ***Nature*** 424: 103-07.

Phan RT, Dalla-Favera R (**2004**). The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. ***Nature*** 2: 432(7017).

Phung QH, Winter DB, Cranston A, Tarone RE, Bohr VA, Fishel R and Gearhart PJ (**1998**). Increased hypermutation at G and C nucleotides in immunoglobulin variable genes from mice deficient in the MSH2 mismatch repair protein. ***J Exp Med*** 187, 1745-751.

Qian J, Wang Q, Dose M, Pruett N, Kieffer-Kwon KR, Resch W, Liang G, Tang Z, Mathé E, Benner C, Dubois W, Nelson S, Vian L, Oliveira TY, Jankovic M, Hakim O, Gazumyan A, Pavri R, Awasthi P, Song B, Liu G, Chen L, Zhu S, Feigenbaum L, Staudt L, Murre C, Ruan Y, Robbiani DF, Pan-Hammarström Q, Nussenzweig MC, Casellas R (**2014**). B Cell Super-Enhancers and Regulatory Clusters Recruit AID Tumorigenic Activity. ***Cell*** 159(7):1524-37

Rada C, Ehrenstein MR, Neuberger MS and Milstein C (**1998**). Hot spot focusing of Somatic Hypermutation in MSH2-deficient mice suggests two stages of mutational targeting. ***Immunity*** 9: 135-141.

Rada C, Williams GT, Nilsen H, Barnes DE, Lindahl T and Neuberger MS (**2002**). Immunoglobulin isotype switching is inhibited and Somatic Hypermutation perturbed in UNG-deficient mice. ***Curr. Biol.*** 12: 1748-755.

Raimondi SC, Chang MN, Ravindranath Y, Behm FG, Gresik MV, Steuber CP, Weinstein HJ and Carroll AJ (**1999**). Chromosomal abnormalities in 478 children with acute myeloid leukemia: clinical characteristics and treatment outcome in a cooperative pediatric oncology group study-POG 8821. ***Blood*** 94: 3707-716.

Ramiro AR, Jankovic M, Eisenreich T, Difilippantonio S, Chen-Kiang S, Muramatsu M, Honjo T, Nussenzweig A and Nussenzweig MC (**2004**). AID is required for MYC/IgH chromosome translocations in vivo. ***Cell*** 118: 431-38.

Ramiro AR, Jankovic M, Callen E, Difilippantonio S, Chen HT, McBride KM, Eisenreich TR, Chen J, Dickins RA (**2006**). Role of genomic instability and p53 in AID-induced c-myc-Igh translocations. ***Nature*** 440: 105-09.

Rebhandl S, Huemer M, Zaborsky N, Gassner FJ, Catakovic K (**2014**). Alternative splice variants of AID are not stoichiometrically present at the protein level in chronic lymphocytic leukemia. ***Eur J Immunol*** 44: 2175-2187.

Reiniger L, Bödör C, Bognár A, Balogh Z, Csomor J, Szepesi A, Kopper L and Matolcsy A (**2006**). Richter's and prolymphocytic transformation of chronic lymphocytic leukemia are associated with high mRNA expression of activation-induced cytidine deaminase and aberrant Somatic Hypermutation. ***Leukemia*** 20: 1089-095.

Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Labelouse R (**2000**). Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). ***Cell*** 102: 565-575.

Robbiani DF, Bothmer A, Callen E, Reina-San-Martin B, Dorsett Y, Difilippantonio S, Bolland DJ, Chen HT, Corcoran AE (**2008**). AID Is Required for the Chromosomal Breaks in c-myc that Lead to MYC/IgH Translocations. ***Cell*** 135: 1028-038.

Robbiani DF, Bunting S, Feldhahn N, Bothmer A, Camps J, Deroubaix S, McBride KM, Klein IA, Stone G (**2009**). AID produces DNA double-strand breaks in non- Ig genes and mature B cell lymphomas with reciprocal chromosome translocations. ***Mol. Cell*** 36: 631-641.

Robbiani DF, Nussenzweig MC (**2013**). Chromosome translocation, B cell lymphoma, and activation-induced cytidine deaminase. ***Annu Rev Pathol.*** 8:79-103

Rogozin IB and Diaz M (**2004**). Cutting edge: DGYW/WRCH is a better predictor of mutability at G:C bases in Ig hypermutation than the widely accepted RGYW/WRCY motif and probably reflects a two-step activation-induced cytidine deaminase-triggered process. ***J Immunol*** 172: 3382-84

Rosenberg BR and Papavasiliou, FN (**2007**). Beyond SHM and CSR: AID and related cytidine deaminases in the host response to viral infection. ***Advances in immunology*** 94: 215-244.

Rothkamm K, Krüger I, Thompson LH and Löbrich M (**2003**). Pathways of DNA double-strand break repair during the mammalian cell cycle. ***Mol. Cell. Biol.*** 23: 5706–5715.

Sabouri S, Kobayashi M, Begum NA, Xu J, Hirota K, Honjo T (**2014**). C-terminal region of activation-induced cytidine deaminase (AID) is required for efficient Class Switch Recombination and gene conversion. ***Proc Natl Acad Sci USA*** 111(6): 2253–2258.

Sagawa Y, Nishi H, Isaka K, Fujito A, Takayama M (**2001**). The correlation of TERT expression with c-myc expression in cervical cancer. ***Cancer Lett.***;168(1):45-50.

Said J, Lones M, Yea S (**2014**). Burkitt lymphoma and MYC: what else is new? ***Adv Anat Pathol.*** 21(3): 160-5.

Sakano H, Maki R, Kurosawa Y, Roeder W, Tonegawa S (**1980**). Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy-chain genes. ***Nature*** 286(5774):676-83.

Saribasak H, Maul RW, Cao Z, Yang WW, Schenten D, Kracker S, Gearhart PJ (**2012**).DNA polymerase ζ generates tandem mutations in immunoglobulin variable regions. ***J Exp Med.*** 209(6):1075-81.

Sato Y, Probst HC, Tatsumi R, Ikeuchi Y, Neuberger MS, Rada C (**2010**). Deficiency in APOBEC2 leads to a shift in muscle fiber type, diminished body mass, and myopathy. J***. Biol. Chem.***285(10):7111-8.

Schreck S, Buettner M, Kremmer E, Bogdan M, Herbst H, Niedobitek G (**2006**). Activation-induced cytidine deaminase (AID) is expressed in normal spermatogenesis but only infrequently in testicular germ cell tumours. ***J Pathol.*** 210(1):26-31.

Seki M, Gearhart PJ and Wood RD (**2005**). DNA polymerases and Somatic Hypermutation of immunoglobulin genes. ***EMBO Rep.*** 6: 1143-48.

Sernández IV, de Yébenes VG, Dorsett Y, Ramiro AR (**2008**). Haploinsufficiency of Activation-Induced Deaminase for Antibody Diversification and Chromosome Translocations both *In Vitro* and *In Vivo*. ***PLoS One*** 3(12):e3927

Severi F, Chicca A, Conticello SG (**2011**). Analysis of reptilian APOBEC1 suggests that RNA editing may not be its ancestral function. ***Mol Biol Evol.*** 28(3):1125-9.

Shenanigans HM, Peters A, Baron B, Zhu X, Storb, U (**1998**). Mutation of BCL-6 gene in normal B cells by the process of Somatic Hypermutation of Ig genes. ***Science*** 280, 1750-52.

Shimizu T, Marusawa H, Matsumoto Y, Inuzuka T, Ikeda A, Fujii Y, Minamiguchi S, Miyamoto S, Kou T, Sakai Y, Crabtree JE, Chiba T (**2014**). Accumulation of somatic mutations in TP53 in gastric epithelium with Helicobacter pylori infection. ***Gastroenterology*** 147(2):407-17.e3.

Shinkura R, Ito S, Begum NA, Nagaoka H, Muramatsu M, Kinoshita K, Sakakibara Y, Hijikata H, Honjo T (**2004**). Separate domains of AID are required for Somatic Hypermutation and class-switch recombination. ***Nat. Immunol.*** 5: 707-12.

Shiramizu B and Ian Magrath I (**1990**). Localization of Breakpoints by Polymerase Chain Reactions in Burkitt's Lymphoma With 8;14 Translocations. ***Blood*** 75(9):1848-52.

Simsek D, Brunet E, Wong SY, Katyal S, Gao Y, McKinnon PJ, Lou J, Zhang L, Li J, Rebar EJ, Gregory PD, Holmes MC, Jasin M (**2011**). DNA ligase III promotes alternative nonhomologous end-joining during chromosomal translocation formation. ***PLoS Genet.*** 7(6):e1002080.

Soni A, Siemann M, Grabos M, Murmann T, Pantelias GE, Iliakis G (**2014**). Requirement for Parp-1 and DNA ligases 1 or 3 but not of Xrcc1 in chromosomal translocation formation by backup end joining. ***Nucleic Acids Res.*** 42(10):6380-92.

Soulas-Sprauel P, Le Guyader G, Rivera-Munoz P, Abramowski V, Olivier-Martin C, Goujet- Zalc C, Charneau P and de Villartay JP (**2007**). Role for DNA repair factor XRCC4 in immunoglobulin Class Switch Recombination. ***The Journal of experimental medicine*** 204: 1717-1727.

Soverini S, Branford S, Nicolini FE, Talpaz M, Deininger MW, Martinelli G, Müller MC, Radich JP, Shah NP (**2014**). Implications of BCR-ABL1 kinase domain-mediated resistance in chronic myeloid leukemia.***Leuk Res.*** 38(1):10-20.

Stavnezer J, Guikema JE and Schrader CE (**2008**). Mechanism and regulation of Class Switch Recombination. ***Annu Rev Immunol*** 26: 261-292.

Sungalee S, Mamessier E, Morgado E, Grégoire E, Brohawn PZ, Morehouse CA, Jouve N, Monvoisin C, Menard C, Debroas G, Faroudi M, Mechin V, Navarro JM, Drevet C, Eberle FC, Chasson L, Baudimont F, Mancini SJ, Tellier J, Picquenot JM, Kelly R, Vineis P, Ruminy P, Chetaille B, Jaffe ES, Schiff C, Hardwigsen J, Tice DA, Higgs BW, Tarte K, Nadel B, Roulland S (**2014**). Germinal center reentries of BCL2-overexpressing B cells drive follicular lymphoma progression.***J Clin Invest.*** 124(12):5337-51.

Swanson PC, Desiderio S (**1998**). V(D)J recombination signal recognition: distinct, overlapping DNA-protein contacts in complexes containing RAG1 with and without RAG2. ***Immunity*** 9(1):115-25.

Ta VT, Nagaoka H, Catalan N, Durandy A, Fischer A, Imai K, Nonoyama S, Tashiro J, Ikegawa M, Ito S, Kinoshita K, Muramatsu M, Honjo T (**2003**). AID mutant analyses indicate requirement for class-switch-specific cofactors. ***Nat. Immunol.*** 4: 843-8,.

Taub R, Kirsch I, Morton C, Lenoir G, Swan D, Tronick S (**1982**). Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. ***Proc. Natl. Acad. Sci. USA*** 79: 7837–7841.

Teng B, Burant CF and Davidson NO (**1993**). Molecular cloning of an apolipoprotein B messenger RNA editing protein. ***Science*** 260: 1816-19.

Teng G, Hakimpour P, Landgraf P, Rice A, Tuschl T, Casellas R, Papavasiliou FN (**2008**). MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. ***Immunity*** 28:621-629.

Tran TH, Nakata M, Suzuki K, Begum NA, Shinkura R (**2010**). B cell-specific and stimulation-responsive enhancers derepress Aicda by overcoming the effects of silencers. ***Nat. Immunol.*** 11: 148-154.

Uchimura Y, Barton LF, Rada C, Neuberger MS (**2011**), REG-γ associates with and modulates the abundance of nuclear activation-induced deaminase. **J. Exp. Med** 208: 2385-2391.

van Maldegem F, Scheeren FA, Aarti Jibodh R, Bende RJ, Jacobs H, van Noesel CJ (**2009**). AID splice variants lack deaminase activity. ***Blood*** 113(8):1862-4.

van Maldegem F, Jibodh RA, van Dijk R, Bende RJ, van Noesel CJ (**2010**). Activation-Induced Cytidine Deaminase Splice Variants Are Defective Because of the Lack of Structural Support for the Catalytic Site. ***J Immunol*** 184: 2487-2491.

Vennstrom B, Sheiness D, Zabielski J, Bishop JM (**1982**). Isolation and characterization of c-myc, a cellular homolog of the oncogene (v-myc) of avian myelocytomatosis virus strain 29. ***J.Virol.*** 42(3):773-9.

Wagner SD, Milstein C and Neuberger MS (**1995**). Codon bias targets mutation. ***Nature*** 376: 732.

Wang CL, Harper RA and Wabl M (**2004**). Genome-wide Somatic Hypermutation. ***PNAS*** 101: 7352-56.

Wang L, Wuerffel R, Feldman S, Khamlichi AA and Kenter AL (**2009**). S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during Class Switch Recombination. ***The Journal of experimental medicine*** 206: 1817-1830.

Wang ZR, Liu W, Smith ST, Parrish RS and Young SR (**1999**). c-myc and chromosome 8 centromere studies of ovarian cancer by interphase FISH. ***Exp. Mol. Pathol.*** 66: 140 –14.

Wenjian Ma, Vijayalakshmi Panduri, Joan F. Sterling, Bennett Van Houten, Dmitry A. Gordenin and Michael A. Resnick (**2009**). The Transition of Closely Opposed Lesions to Double-Strand Breaks during Long-Patch Base Excision Repair Is Prevented by the Coordinated Action of DNA Polymerase δ and Rad27/Fen1. ***Mol. Cell Biol*** 29(5):1212-21.

White CA, Seth Hawkins J, Pone EJ, Yu ES, Al-Qahtani A, Mai T, Zan H, Casali P (**2011**). AID dysregulation in lupus-prone MRL/Fas(lpr/lpr) mice increases class switch DNA recombination and promotes interchromosomal MYC/IgH loci translocations: modulation by HoxC4. ***Autoimmunity*** 44(8):585-98.

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA. (**2013**). Master transcription factors and mediator establish super-enhancers at key cell identity genes. ***Cell*** 153(2).

Wood C, Tonegawa S (**1983**). Diversity and joining segments of mouse immunoglobulin heavy chain genes are closely linked and in the same orientation: implications for the joining mechanism. ***Proc. Natl. Acad. Sci. USA*** 80(10):3030-4.

Wray J, Williamson EA, Singh SB, Wu Y, Cogle CR, Weinstock DM, Zhang Y, Lee SH, Zhou D, Shao L, Hauer-Jensen M, Pathak R, Klimek V, Nickoloff JA, Hromas R (**2003**). PARP1 is required for chromosomal translocations. **Blood** 121(21):4359-65.

Wu X, Darce JR, Chang SK, Nowakowski GS and Jelinek DF (**2008**). Alternative splicing regulates activation-induced cytidine deaminase (AID): Implications for suppression of AID mutagenic activity in normal and malignant B-cells. ***Blood*** 112: 4675- 682.

Xiang S, Short SA, Wolfenden R and Carter CW (**1996**). Cytidine deaminase complexed to 3-deazacytidine: a "valence buffer" in zinc enzyme catalysis. ***Biochemistry*** 35: 1335-341.

Xiang S, Short SA, Wolfenden R and Carter CW (**1997**). The structure of the cytidine deaminase-product complex provides evidence for efficient proton transfer and ground-state destabilization. ***Biochemistry*** 36: 4768-774.

Xu J, Husain A, Hu W, Honjo T, Kobayashi M (**2014**). APE1 is dispensable for S-region cleavage but required for its repair in Class Switch Recombination. ***Proc. Natl. Acad. Sci. USA*** 111(48):17242-7.

Yadav A, Olaru A, Saltis M, Setren A, Cerny J, Livák F (**2006**). Identification of a ubiquitously active promoter of the murine activation-induced cytidine deaminase (AICDA) gene. ***Mol Immunol*** 43: 529-541.

Yan CT, Boboila C, Souza EK, Franco S, Hickernell TR, Murphy M, Gumaste S, Geyer M, Zarrin AA, Manis JP (**2007**). IgH class switching and translocations use a robust non-classical end-joining pathway. ***Nature*** 449: 478–482.

Yu Q, Chen D, König R, Mariani R, Unutmaz D, Landau NR (**2004**). APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. ***J. Biol. Chem.*** 279(51): 53379-86.

Yun MH and Hiom K (**2009**). CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. ***Nature*** 459: 460– 463.

Zan H, White CA, Thomas LM, Mai T, Li G, Xu Z, Zhang J and Casali P (**2012**). Rev1 recruits ung to switch regions and enhances du glycosylation for immunoglobulin class switch DNA recombination. ***Cell Rep*** 2: 1220-1232.
Zhenming Xu, Zan H, Pone EJ, Mai T and Casali P (**2012**). Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. ***Nature Reviews Immunology*** 12: 517-531.

Zu Y, Tong X, Wang Z, Liu D, Pan R, Li Z, Hu Y, Luo Z, Huang P, Wu Q, Zhu Z, Zhang B, Lin S (**2013**). TALEN-mediated precise genome modification by homologous recombination in zebrafish. ***Nat Methods*** (4):329-31.