universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

A robust and high-throughput Assay to study Somatic Hypermutation unveiling the role of 14-3-3 Adaptor Proteins

verfasst von / submitted by

Dott.ssa.mag. Marialaura Mastrovito

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2018 / Vienna, 2018

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	A 794 685 490
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Molekulare Biologie
Betreut von / Supervisor:	Rushad Pavri, PhD

TABLE OF CONTENTS

Table of contents

Acknowledgments

I would like to start by thanking Rushad, who gave me the possibility to work on such an interesting and challenging project, and all the present and past members of the Pavri group.

I thank my PhD advisory committee for valuable scientific inputs and Ines for all the support throughout these four years. I'm thankful to the Busslinger lab and to Meinrad, for our fruitful weekly discussion.

I thank the whole VBC community, the way you can do science here is simply amazing!

A big thanks goes to Anna, Svend, and Kerstin, for believing in me and supporting me, sometimes even against their interest. Thank you for your "You are smarter than what you think".

I thank Lara and Vale, my Ifom-mums, for their precious tips that made me stronger.

I'm thankful to my friends Deniz and Hagar, this journey without you would have not been the same. We shared excitements and frustrations and if I managed till this point is for sure thanks to you two.

I thank Adri for always been there for me, even in busy and stressful times.

I thank the "Italians" and all my friends at the campus for the nice time we spent together and I thank Sarah, best thing of the new IMP desk area.

I thank Sabri and Giuli, for always being there besides the distance. I thank Federica and Giulia, *"per capire ogni situazione senza che io dicessi una parola"*.

I'm deeply and sincerely thankful to Tobi, *"per supportarmi e sopportarmi"*. This four years journey was only possible thanks to your love.

Last but not least a huge thanks goes to my family, Mamma, Papà and Pippi. Thank you for believing in me, for always thinking I could do it and for making all of this possible.

Abstract

A functional acquired immune response is achieved by the generation of an enormous number of serum immunoglobulins (*Ig*). The coding capability of the *Ig* locus is not enough to guarantee such a diverse repertoire of antibodies, therefore antibody diversification processes are needed. After rearrangement of the *Ig* locus, the variable region is diversified by somatic hypermutation (SHM).

SHM is the process by which mutations are introduced by activation-induced cytidine deaminase (AID) into the variable regions of antibodies in germinal center B cells. These mutations are crucial during the immune response for generation, maturation and clonal expansion of high-affinity antibody-expressing B cells. Although transcription has been clearly implicated, the mechanisms underlying the highly discrete mutation spectra and high mutation rates are unknown.

Studies on SHM were so far very limited due to the lack of an *in vitro* experimental system that can mimic the *in vivo* mechanism. To overcome the traditional limitations of studying SHM, here we establish an *in vitro* approach to screen for SHM co-factors.

We generated a modified Ramos B cell line, which expresses an inducible and hyperactive AID version, allowing an enhanced SHM. This line has been fully characterized using MutPE-seq (mutational analysis by paired-end sequencing) and it recapitulates the physiological mutation patterns and rates observed *in vivo*.

Using this model line, we performed a targeted RNAi screen to investigate on SHM co-factors, which highlighted the role for the 14-3-3 adaptor proteins family. Indeed, MutPE-seq on 14-3-3-depleted Ramos showed a defective hypermutation compared to the control. Preliminary results also showed that 14-3-3 interacts with Spt6, which is known to have an important role in AID targeting. All together, these findings shed light on the importance of 14-3-3 in the molecular mechanism of SHM.

Zusammenfassung

Um eine funktionsfähige adaptive Immunantwort zu bilden, werden enorme Mengen an Serumimmunglobulinen (*Ig*) erzeugt. Da die Kodierungskapazität es *Ig* Lokus jedoch nicht ausreicht um solch eine immense Antikörperproduktion sicherzustellen, werden Diversifizierungsprozesse benötigt. Daher wird die variable Region, nach der Umlagerung des *Ig* Lokus, durch somatische Hypermutation (SHM) diversifiziert.

SHM bezeichnet die Einführung von Mutationen in die variablen Antikörperregionen von B-Zellen des Keimzentrums durch die aktivierungsinduzierte Cytidindeaminase (AID). Diese Mutationen sind während einer Immunantwort wesentlich für die Erzeugung, Reifung und klonale Expansion von B-Zellen, die hochaffine Antiköper exprimieren. Obwohl die Transkription in diesem Prozess zweifellos eine Rolle spielt, ist der genaue Mechanismus, der diesem hochdiskretem Mutationsspektrum und den hohen Mutationsraten zugrunde liegt, unbekannt.

SHM-Studien waren bisher nur begrenzt möglich, da ein experimentelles System, welches den *in vivo* Mechanismus imitieren kann, nicht zur Verfügung stand. Um diese klassischen Einschränkungen von SHM-Studien zu überwinden, führen wir hier einen *in vitro* Ansatz zum Screening von SHM-Kofaktoren ein.

Wir erzeugten eine modifizierte Ramos B-Zelllinie, die eine induzierbare und hyperaktive AID-Version exprimiert, was zu einer erhöhten SHM-Rate führt. Mit Hilfe von MutPE-seq (mutative Analyse durch *Paired-end*-Sequenzierung) wurde diese Linie vollständig charakterisiert. Unsere Ergebnisse zeigen, dass dieses System die physiologischen Mutationsmuster nachbildet und die *in vivo* Mutationsraten des Keimzentrums imitiert.

Wir optimierten daher ein gezieltes RNAi-Screening um SHM-Kofaktoren zu untersuchen.

Das Screening zeigte, dass die 14-3-3 Adaptorproteinfamilie eine Rolle in der Regulierung der somatischen Hypermutation spielt. MutPE-seq von Ramoszellen nach einem 14-3-3 *Knockdown*, bestätigte einen Hypermutationsdefekt im Vergleich zu Kotrollzellen. Des Weiteren zeigten RT-qPCR Experimente, dass der *Knockdown* von 14-3-3 die Transkription des IgH Lokus nicht beeinflusst. Vorläufige *Pulldown*

Experimente bestätigten die Interaktion von 14-3-3 mit dem Transkriptionselongationsfaktor Spt6. Diese Erkenntnis gibt Aufschluss über einen potentiellen molekularen Mechanismus, der einer weiteren Untersuchung bedarf.

List of abbreviations

3'RR	3' Regulatory region
3C	Chromatin Conformation Capture
4-HT	(Z)-4-Hydroxytamoxifen
Å	Angrstrom
BCR	B cell receptor
CBE	CTCF-binding element
CDR	Complementarity determining region
ChIP	Chromatin immunoprecipitation
Co-IP	Co- Immunoprecipitation
CSR	Class switch recombination
DMEM	Dulbecco's Modified Eagle's Medium
DTT	Dithiothreitol
DZ	Dark zone
FC	Fold-change
FDC	Follicular dendritic cell
FT	Flow-through
HS	High salt
Ig	Immunoglobulin
IP	Immunoprecipitation
KDa	Kilodalton
LB	Lysogeny broth
LS	Low salt
LZ	Light zone
MHC	Major histocompatibility complex
MutPE-seq	Mutational profiling by paired-end sequencing
NES	Nuclear export signal
NF	Nuclear fraction
NLS	Nuclear localization signal
NP-CGG	4-Hydroxy-3-nitrophenylacetyl-Chicken Gamma Globulin
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PIC	Protein inhibitor cocktail
PMSF	Phenylmethane sulfonyl fluoride
R26	Rosa 26
RNAi	RNA interference
RPKM	Reads Per Kilobase Million
RT-qPCR	Reverse transcription quantitative PCR
rtTA	Reverse tetracycline-controlled transactivator
SDS	Sodium dodecyl sulfate
SHM	Somatic hypermutation

List of abbreviations

ShRNA	Short hairpins RNA
ssDNA	Single-stranded DNA
TCR	T cell receptor
TE	Tris-Ethylenediaminetetraacetic acid
Ugi	UNG inhibitor
VDJ	Variable, Diversity and Joining segments
WCE	Whole cell extract
WT	Wild-type

1. INTRODUCTION

1.1 Antibody diversification

The number of different antibodies that the human body can produce greatly exceeds the information content of the genome. To face this challenge, B lymphocytes manage to achieve a highly diverse antibody spectrum by combinatorial alterations of the immunoglobulin (*Ig*) gene. The immunoglobulin locus consists of Variable (V), Joining (J) and Diversity (D) elements encoding the variable region of the antibodies, and Switch (S) and Constant (C) fragments, encoding the constant region. The *IgH* locus contains additionally important cis-regulatory elements, such as the Eμ intronic enhancer and the 3' regulatory region (3'RR) (Fig. 1A). The primary repertoire of antibody specificities is created by a process of DNA rearrangement, called VDJ recombination, occurring in the bone marrow: the *Ig* locus undergoes rearrangement of Variable (V), Diversity (D) and Joining (J) segments, to create the VDJ region, enabling the production of the broad range IgM immunoglobulins (reviewed in Teng and Papavasiliou, 2007) (Fig. 1 A).

After maturation in the bone marrow, mature B cells migrate to secondary lymphoid organs, such us lymph nodes, spleen or Peyer's patches, where a secondary diversification may take place. Following antigen encounter, the rearranged VDJ region undergoes class switch recombination (CSR) and somatic hypermutation (SHM). One common feature of CSR and SHM is that they are both initiated by an event of DNA deamination, occurring in germinal centers and mediated by the enzyme Activation-induced cytidine deaminase, AID, expressed only in activated B cells (Fig. 1 B). AID introduces mutations in the switch regions, characteristic of CSR, but also in the Variable region, mediating SHM (Fig. 1 B).

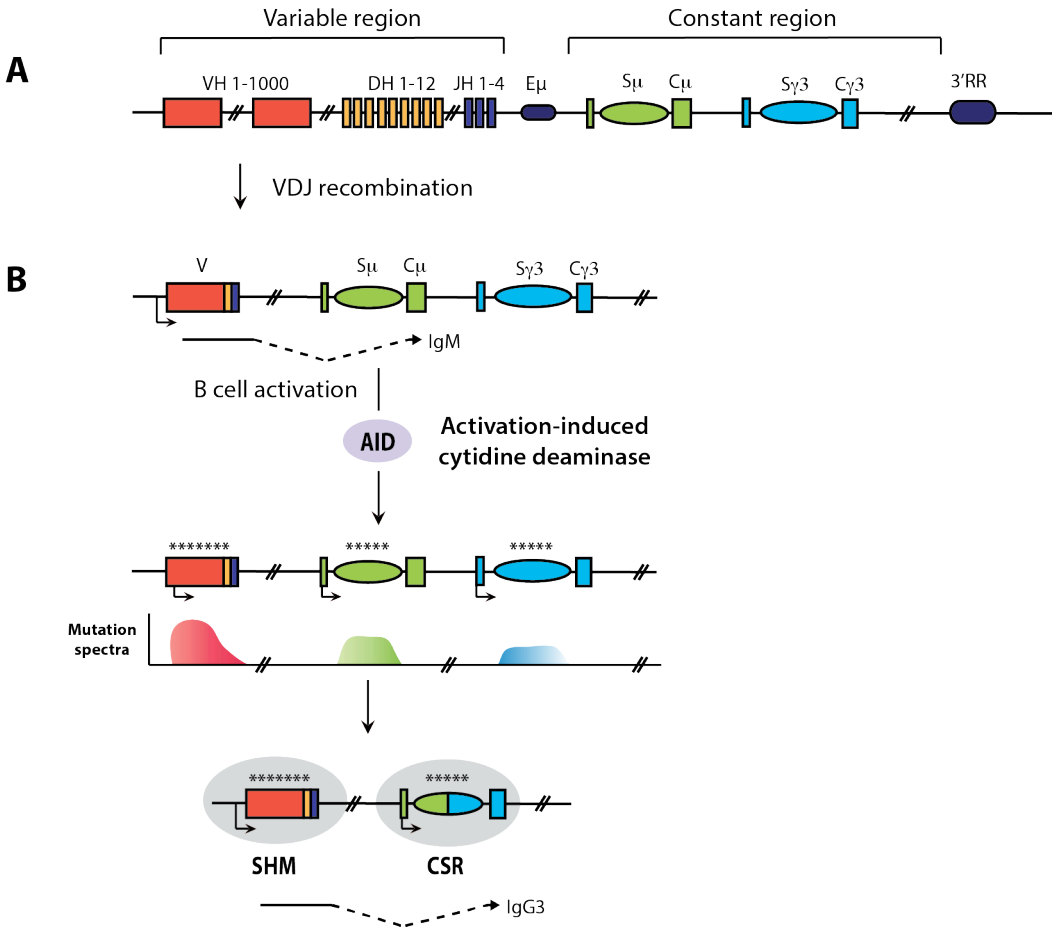


Fig. 1: DNA rearrangements at the *IgH* locus. (A) In the bone marrow the *Ig* locus undergoes rearrangement of Variable (V), Diversity (D) and Joining (J) segments, to create the VDJ region, allowing the production of the broad range IgM immunoglobulins. V, D and J segments encode for the Variable region of Ig, whereas Switch (S) and Constant (C) regions are part of the constant region of Ig. (B) After B cell activation followed by antigen challenge, AID is expressed and mediates the incorporation of mutations (shown as asterisk) in the V region mediating SHM. Mutations introduced in the S region generate DSBs, leading to CSR between the switch (Sμ) and acceptor (Sγ1) sequences. Transcription from the *IgH* promoter sustains the production of IgG3.

Germinal centers are micro-anatomical structures located in secondary lymphoid organs, and consist in sites of active proliferation and selection of B cells. When B cells in the lymph node cortical area get activated by the antigen, they undergo extensive proliferation and form germinal centers (Fig. 2 A). After B cells encounter the antigen through the B cell receptor (BCR), they process and express it onto the Major histocompatibility complex (MHC). T cells via the T cell receptor (TCR) recognize it and

provide a survival signal to the B cells, only if they express high affinity immunoglobulin.

This process is sustained also by important signalling pathways initiated by cytokines as well as the interaction between CD40, expressed by the B cell, and CD40 ligand (CD40L), coming from T cells (Fig. 2 B). High affinity B cells are later on selected by Follicular dendritic cells (FDCs) to become mature plasma cells, namely antibody secreting cells (Victora and Nussenzweig, 2012). In each Germinal center it is possible to distinguish a so called Dark Zone (DZ) where B cells actively proliferate and undergo SHM and a Light Zone (LZ), site of selection, operated by T cells (Fig. 2 A).

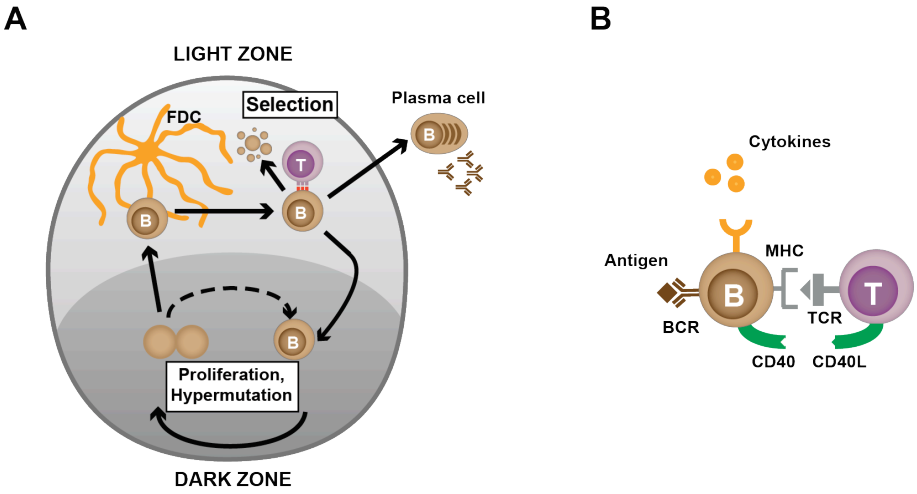


Fig. 2: Germinal center reaction. (A) The specificity for the antigen is achieved by a co-action of hypermutation and selection. Schematic representation of the germinal center reaction: in the dark zone B cells undergo introduction of mutations via SHM while actively proliferating; in the light zone T cells help to select the high affinity B cells. Those can exit the germinal center and become antibody-secreting plasma cells (Adapted from Victora & Nussenzweig, Annu. Rev. Immunol., 2012). (B) B-T cell synapse: After B cells encounter the antigen through the B cell receptor (BCR), in presence of cytokines signals, they process and express it onto the Major histocompatibility complex (MHC). T cells via the T cell receptor (TCR) recognize it and provide a survival signal to the B cells, only if they express high affinity immunoglobulin.

1.2 Molecular mechanism of SHM and CSR

CSR is a process of recombinatorial deletions, resulting from mutated switch (S) regions, that remove the exons of the IgM constant region, bringing the functional VDJ segment into proximity with the exons of downstream constant regions; this allows the switching to different Ig isotypes, with a diverse effector function (Gazumyan et al., 2012) (Fig. 1 B). SHM instead consists in a non-random scattering of single nucleotide substitutions into the variable domain (Rajewsky et al., 1987) (Fig. 1 B, indicated by asterisks). Previous findings suggest a focus of mutations in the complementarity determining regions (CDRs) of the variable regions, within characteristic overlapping AGCT hotspots (Wei et al., 2015). Mutations focus within the V region with a precise 5' boundary delimitated by the promoter region (Gearhart and Bogenhagen, 1983; Lebecque, 1990; Rothenfluh et al., 1993; Steele et al., 1992; Weber et al., 1991), and go till the non-coding intronic region at the 3' of the rearranged V(D)J region (Lebecque, 1990; Steele et al., 1992; Weber et al., 1991). Importantly mutations in the downstream C fragments are very rare (Crews et al., 1981; Kim et al., 1981) (Fig. 1 B).

In vivo, the substitutions accumulate with a rate of incorporation into the V gene of 10^{-4} and 10^{-3} bp^{-1} $generation^{-1}$ (McKean et al., 1984; Berek and Milstein, 1987), several orders of magnitude higher than the rate of spontaneous mutations.

CSR and SHM are both initiated by AID that targets single-stranded DNA (ssDNA) and deaminates the dC residues in WRC (W=A/T, R=A/G) hotspot motifs (Odegard and Schatz, 2006). Both the mechanisms result from a combination of AID mutagenic activity and error-prone DNA repair pathways. AID triggers the deamination of Cytidine to Uridine, thereby converting C:G pairs in U:G pairs (Neuberger and Rada, 2007). The resulting Uridine is repaired via different pathways (Fig. 3): (i) if DNA replication is occurring, the output is a transition at C:G pairs; (ii) the Uracil DNA Glycosylase (UNG) creates an abasic site, causing transversion and transition (Di Noia and Neuberger, 2002); (iii) the Mismatch Recognition Heterodimer MSH2-MSH6 induces mutation at A:T pairs (Wilson et al., 2005). In the absence of repair the AID-induced lesion can also lead to c-Myc/IgH translocations and lymphoid malignancies (Jankovic et al., 2010; Ramiro et al., 2004; Ramiro et al., 2006; Robbiani et al., 2008; Robbiani et al., 2009).

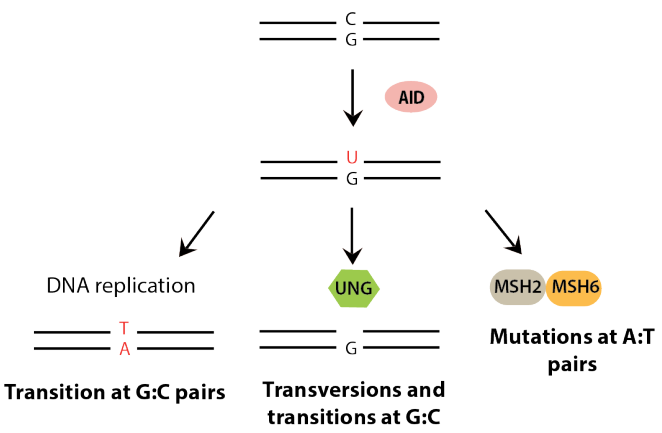


Fig. 3: Error-prone pathways in SHM. AID induces the deamination of Cytidine to Uridine. The resulting repair uses different pathways. If DNA replication is occurring, the output is a transition at C:G pairs; the Uracil DNA Glycosylase (UNG) creates an abasic site, causing transversion and transition; the Mismatch Recognition Heterodimer MSH2-MSH6 induces mutation at A:T pairs.

AID acts in a transcription-dependent manner by mutating V regions during SHM, and S regions in CSR. The two fragments differ especially in their sequence length and the nature of the AID targeting domain. V regions are relatively short (500 bp – 1 kb) and do not contain repetitive sequences or G-rich elements. On the contrary, S regions are very long (3 – 12 kb), repetitive and enriched for Gs. This allows S regions to form DNA secondary structures, previously shown to facilitate AID targeting (reviewed in Honjo et al., 2002). Thus, besides many shared features, SHM and CSR have numerous differences that let us infer a potential differential mechanism of regulation (summarized in Table 1).

Table 1: Similarities and differences between SHM and CSR.

Feature	SHM	CSR
ssDNA targeting by AID	✔	✔
Transcription-dependency	✔	✔
High mutation rates	✔	✔
Mutation domain	500 bp – 1 kb	3 -12 kb
G-rich and repetitive	no	yes
DNA secondary structure	no	yes

1.3. AID-mediated deamination

AID was discovered by Honjo and colleagues when they stimulated the mouse CH12F3 B cell lymphoma line to undergo CSR and compared the resulting cDNAs with resting, non-stimulated B cells. (Muramatsu at al., 1999). Genetic deletion of AID in B cells was soon shown to impair both CSR and SHM (Muramatsu et al., 2000; Revy, 2000): a fundamental discovery in the molecular mechanism of antibody diversification was just been made. AID is part of the apolipoprotein B mRNA-editing catalytic component, APOBEC family, and shares many similarities with APOBEC1 (Muramatsu at al., 1999). Given its high homology with APOBEC1, AID was originally thought to be an enzyme able to interact with RNA. However subsequent studies confirmed single-stranded DNA (ssDNA) as the only AID substrate (Dickerson et al., 2003; Larijani at al., 2007; Pham at al., 2003).

1.3.1 AID Biochemistry

AID is a small protein of only 198 amino acids (Muramatsu at al., 1999). As for the APOBEC proteins, AID's catalytic activity is driven by an active site containing two cysteine and a hystidine residues that, together with a glutamate serving as a proton donor, coordinates a Zinc (Zn) atom and mediates the deamination reaction (Conticello et al., 2007).

After decades of attempts, a recent study resolved the AID crystal structure, confirming AID motifs preferences as well as some crucial discrepancies from the APOBEC proteins (Pham et al., 2016). Pham and colleagues report the crystal structure of a soluble AID variant lacking 15 amino acids at the C-terminus (AID Δ15). They show that AID Δ15 retains the same biochemical features as the native AID and confirm its ability to mediate SHM in Ramos B cells. They unrevealed the structural basis of the AID preference for WRC motif showing how the substrate specificity loop is larger than the one in APOBECs and spans outside of the active site. Therefore AID is able to accommodate two purines adjacent to the targeted 5'-C (Pham et al., 2016). If on one hand WRC motifs are the AID hotspots (Odegard and Schatz, 2006), on the other hand SYC motifs (S = (G/C) and Y = (T/C)) are the least likely to be targeted by the enzyme (Pham, 2003;

Bransteitter at al., 2004; Yu et al., 2003; Wang et al., 2010).

Besides giving insights on the catalytic pocket of AID, the recently resolved crystal structure left some unresolved questions regarding the quaternary structure of the enzyme. It is in fact yet unknown if AID is enzymatically active when present in a monomeric form. Some evidences suggest that AID is present in the cell as a dimer or tetramer. Purification of AID-tagged versions and measurement of their activity after separation on size exclusion columns confirmed that the monomeric fraction of AID is lacking enzymatic activity (Larijani et al., 2007). Nonetheless the fact that an AID fraction might be present in the cells as a monomer cannot be excluded.

The analysis of AID biochemical hallmarks show that the enzyme has a very slow deamination rate, accounting for one reaction in several minutes (Larijani et al., 2007). This feature might be related to the tight regulation, which AID undergoes in order to avoid its off-target activity.

Although being AID such a small protein that could passively diffuse into the nucleus, biochemical studies identified a Nuclear Localization Signal (NLS) and a Nuclear Export Signal (NES) in the N and C-terminus respectively, as schematically represented in Fig. 4. (Durandy et al., 2006). Additional mutation studies identified amino acid residues at the C or N-terminus, essential for SHM and CSR (Brar et al., 2004; McBride et al., 2004; Patenaude et al., 2009).

Fig. 4: Functional domains of AID. Schematic representation of AID domains and their respective positions (numbers indicate the amino acid positioning). The most common phosphorylation sites are also indicated (S3, S38, Y138 and T140). NLS: nuclear localization signal; NES: nuclear export signal.

1.3.2 AID regulation

Although AID targets preferentially the *Ig* locus, it can also deaminate other genes, resulting in point mutations or oncogenic translocations (Shen et al., 1998; Pasqualucci et al., 2001; Ramiro et al., 2006; Robbiani et al., 2008; Robbiani et al., 2009). For this reasons, its expression, localization and activity are tightly regulated by several mechanisms. Besides carrying out its function on ssDNA, the steady-state localization of the enzyme is mostly in the cytoplasm (Rada et al., 2002). An active nuclear import, mediated by the NLS, counteracts the cytoplasmic retention of AID (Patenaude et al., 2009). In fact once AID is in the nucleus it targets and mutates the *Ig* locus. Its stability is controlled by a mechanism of ubiquitin-independent REG-γ-mediated degradation. REG-γ deficiency in B cells leads to less degradation of AID, subsequently more AID in the nucleus and higher levels of CSR. (Uchimura et al., 2011). In addition to the REG-γ mediated degradation, AID amount in the nucleus is controlled by a canonical ubiquitin-dependent degradation (Uchimura et al., 2011; Aoufouchi et al., 2008).

In order to actively translocate back in the cytoplasm AID takes advantage of the Crm1/Exportin-1-dependent export system, mediated by the NES (Brar et al., 2004; Ito et al., 2004; McBride et al., 2004).

Besides a tight regulation of the subcellular localization of AID, mechanisms controlling the protein stability are also essential. AID half-life is very different when nucleus and cytoplasmic protein fractions are compared. In fact AID in the nucleus is stable only for 2.5 hours. On the contrary, it can be stable in the cytoplasm up to 18-20 hours (Aoufouchi et al., 2008). This discrepancy in the half-life of the protein can be explained by the role of the Heat shock protein 90 (HSP90) that in the cytoplasm interacts with AID and prevents its proteasomal degradation (Orthwein et al., 2010). Accordingly, HSP90 inhibition leads to a reduction in immunoglobulin gene diversification and off-target mutations (Orthwein et al., 2010).

AID ubiquitous overexpression in mice leads to both T cell lymphoma and adenomas (Okazaki et al., 2003), but in order to develop B-cell derived cancers the additional deletion of p53 is needed (Ramiro et al., 2006; Robbiani et al., 2008). AID expression is cell-type restricted and can be induced *in vitro* by adding CD40 ligand, LPS, IL4 and

TGFβ. (Dedeoglu et al., 2004; Muramatsu et al., 1999; Zhou et al., 2003). This accounts for a tight transcriptional regulation of AID that requires transcription factors like Pax5 and E47 (Crouch et al., 2007) and inhibitors of differentiation proteins, Id2 and Id3 (Gonda et al., 2003; Sayegh et al., 2003; Tran et al., 2010). In addition, signaling proteins triggered by CD40 ligand, IL4 and TGFβ include the NfkB pathway, STAT6 and Smad3/4 (Tran et al., 2010).

Once AID is transcribed, the stability of its messenger RNA is controlled by miR-155 and miR-181b (de Yébenes et al., 2008; Dorsett et al., 2008; Teng et al., 2008). MiR-155 specifically accounts as a tumor suppressor gene, given that its depletion causes enhanced AID levels and c-Myc/IgH translocations (Dorsett et al., 2008).

Further step of AID regulation consists in the phosphorylation of specific residues (Fig. 4). Mass spectrometry studies on switching B cells revealed numerous phosphorylated residues, such as Ser3, Ser38, Thr140 and Tyr184 (Basu et al., 2005; Gazumyan et al., 2011; McBride et al., 2006). Although very little is known about signalling pathways controlling AID phosphorylation, it is well understood that each phosphorylated amino acid has a different output on AID activity during SHM and CSR (Basu et al., 2005; Chatterji et al., 2007; Gazumyan et al., 2011; McBride et al., 2006). Proteins known to introduce this modification on AID specific residues are (i) protein kinase A (PKA), that phosphorylates at Ser38 (Basu et al., 2005) and specifically recruits AID to S-regions (Vuong et al., 2009); (ii) protein kinase C (PKC), phosphorylating Thr140; (iii) and protein phosphatase 2A (PP2A), specific for AID-Ser3 phosphorylation (Gazumyan et al., 2011; McBride et al., 2008). AID phosphorylation at Ser38 by PKA allows AID to interact with replication protein A (RPA) and target ssDNA at *Ig* locus (Basu et al., 2005; Chaudhuri et al., 2004).

Taken together, this information remark the importance of AID activity restriction, in order to keep a fine equilibrium between the necessity of a broad-range antibody repertoire and the oncogenic nature of AID.

1.4. AID targeting to the *IgH* locus

Studies on mutational patterns in B cells undergoing SHM have shown that mutations cluster in transcribed regions of 1-2 kb, at the 3' of the V-region or S-region promoters (Gearhart and Bogenhagen, 1983; Lebecque, 1990; Steele et al., 1992; Weber et al., 1991). Replacing V-promoters with unrelated non-*Ig* promoters, such as β-globin or B29, leaves the level of SHM unchanged (Betz et al., 1994; Tumas-Brundage and Manser, 1997). Therefore, long time before AID discovery, it was already clear that rates of SHM correlate with rates of transcription. The definitive link between somatic hypermutation and the transcription machinery came with the study of Storb and colleagues (Peters and Storb, 1996). They placed a non-mutated exon under a promoter control and showed it is undergoing SHM. Thus proving that locating a sequence within a transcriptional unit promotes hypermutation (Peters and Storb, 1996). Their model accounts for a "mutator factor", later on discovered to be AID, expressed only in activated B cells, that is linked to the transcription machinery, specifically with the initiation complex at the promoter and stays associated with the elongation complex as well (Peters and Storb, 1996). After the discovery of AID and studies showing its *in vitro* interaction with ssDNA, the connection with the transcription machinery became clear: transcription is providing ssDNA as template for AID-mediated mutagenesis (Petersen-Mahrt et al., 2002; Ramiro et al., 2003; Sohail et al., 2003). Nambu and colleagues show that AID is present at transcriptionally active *Ig* sites and associates with RNA Polymerase II (RNA Pol II) (Nambu et al., 2003). The physical association was also confirmed biochemically by Willmann and colleagues. They characterized chromatin binding protein complexes containing AID and found that most of the proteins binding AID are part of the RNA polymerase II elongation complex (Willmann et al., 2012). Following studies discovered that AID physically interacts with several molecular players of transcription elongation. The transcription elongation factor, SPT5 has been shown to stabilize AID at stalled Pol II, in order to guarantee that B cells undergo efficient CSR (Pavri et al., 2010). Another evidence of AID targeting linked to transcription elongation comes from the observation that the DNA Topoisomerase I (TOPI) knock-down promotes SHM and CSR (Kobayashi et al., 2011; Maul et al., 2015). RNA Pol II stalling resulting from unsolved supercoiling fosters in fact AID targeting at the *Ig* locus (Maul et al., 2015). AID interacts also with

the RNA Exosome complex. The AID targeting model proposed by Basu and colleagues considers that the Exosome-mediated degradation of R-loops, often arising at S-regions, helps recruiting AID at ssDNA targets (Basu et al., 2011; Pefanis et al., 2015).

A very recent study from Di Noia and colleagues have shown that AID mutants at specific Arginine residues (R171, R174, R178), so called R-mutants fail to interact with the elongation factor SPT6, nonetheless keeping the biochemical ability to bind it (Methot et al., 2018). Their model of AID targeting coupled with transcription elongation postulates that when the amount of AID in the nucleus is low, it accumulates at promoter-proximally paused RNA Pol II, via SPT5 (Pavri et al., 2010). AID is targeting the *Ig* locus only when it has a permissive transcriptional landscape, therefore when AID is coupled with transcription elongation, it is licensed to deaminate the downstream regions. The fact that R-mutants still have the biochemical capability to bind SPT6 suggests that the interaction might not be direct. The interactome map derived by Bio-ID studies reveals that R-mutants lose interaction also with two of the 14-3-3 adaptor subunits, scaffold proteins known to mediate several cellular processes (reviewed in Fu et al., 2000). This evidence suggests that the 14-3-3 adaptor protein family may be the linking factor of AID and SPT6.

1.5. 14-3-3 adaptor proteins

The 14-3-3 adaptor protein discovery resulted from a screening for abundant brain proteins in mammals (Moore, 1967). Further studies revealed that these proteins are ubiquitously expressed in eukaryotic organisms and very abundant in almost all cell types (reviewed in Fu et al., 2000). In Human, seven isoforms have been described (α/β, ε, η, δ/ζ, τ, γ and σ), each encoded by different genes (YWHAB, YWHAE, YWHAH, YWHAZ, YWHAQ, YWHAG, SFN1) and highly homologues among each other (Fig. 5). They form homo- or heterodimers through the N-terminal portion (Liu et al., 1995; Xiao et al., 1995) and they mainly localize in the cytoplasm (Garcia-Guzman et al., 1999; Freed et al., 1994; reviewed in Fu et al., 2000). 14-3-3 proteins form preferentially some dimer combinations. For example, 14-3-3 γ forms heterodimers mainly with 14-3-3 ε, and ε with 14-3-3 β, η, γ and ζ, but does not form homodimers (Aitken, 2018). 14-3-3 σ preferentially forms homodimers (Xu et al., 2010). The 14-3-3 protein monomer has a

molecular weight of approximately 30 KDa and consists in nine alpha-helices with an antiparallel orientation. Crystal structures of 14-3-3 proteins showed that each monomer has an inner amphipathic concave surface that is very conserved among the several protein isoforms (Yaffe et al., 1997; Petosa et al., 1998). This suggests that, since most of the 14-3-3 ligands are shared among different subunits, the inner surface might act as a binding domain.

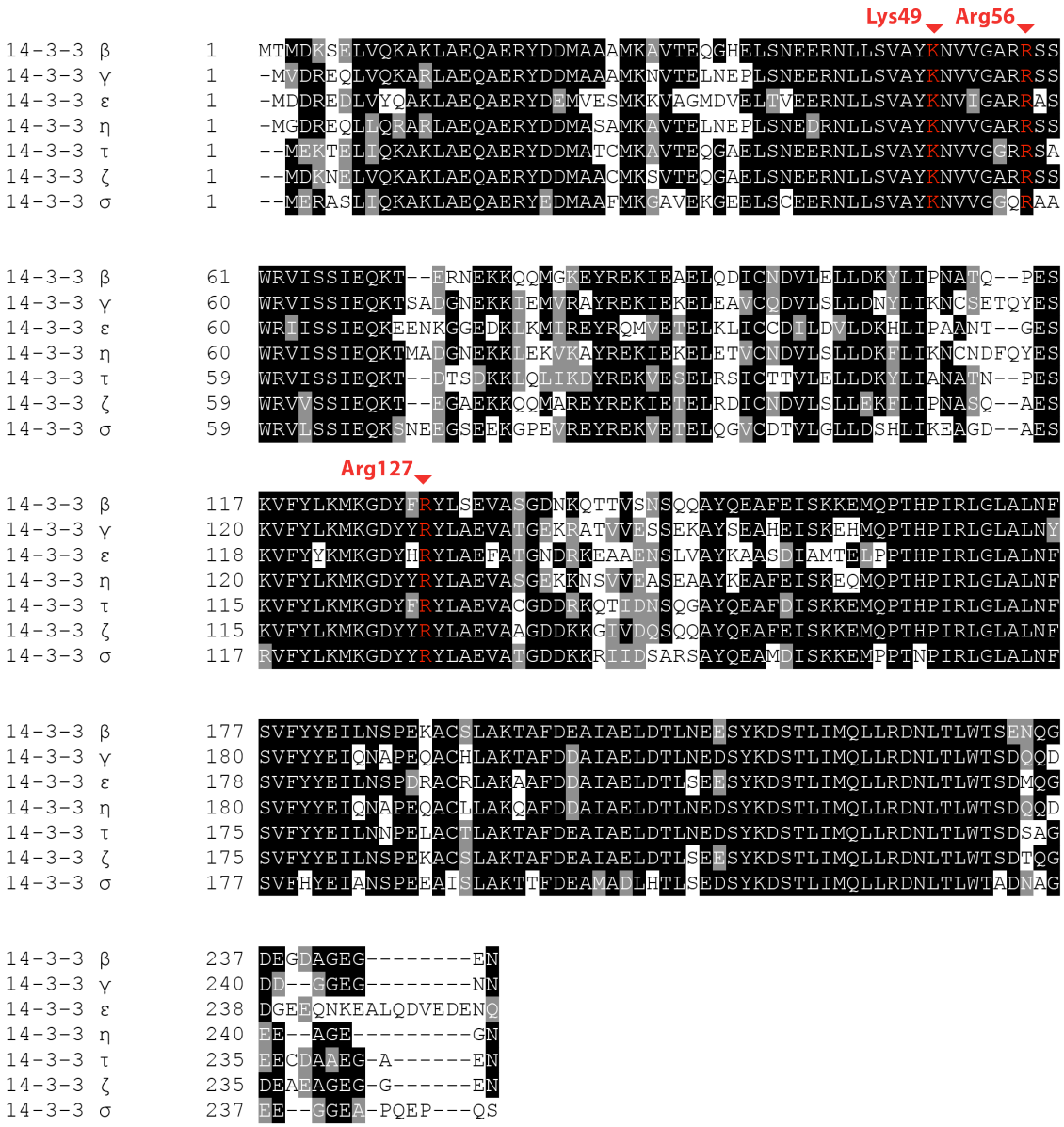


Fig. 5: 14-3-3 isoforms alignment. The protein alignment was realized using the tool tcoffee (http://tcoffee.crg.cat) and is based on data generated by Xu Z. and colleagues (Xu et al., 2010).

14-3-3 proteins regulate numerous cellular processes, including cell cycle (Chan et al., 1999; Conklin et al., 1995; Jiang et al., 2003; Lee et al., 2001; Wang et al., 2000; Waterman et al., 1998), development (Lau and Muslin, 2009) and growth factors signalling (Thomas et al., 2005) as well as apoptosis and cell metabolism (Aitken, 2002; Fu et al., 2000; Muslin and Xing, 2000; reviewed in Yaffe, 2002). It is therefore not surprising that proteomic affinity-based purifications identified around hundred 14-3-3 binding factors, and *in silico* prediction studies account for more than two thousands ligands (reviewed in Aitken et al., 2002), phosphatases (Cdc25 (Dalal et al., 1999; Peng et al., 1998) and kinases (Raf1 (Michaud et al., 1995), Breakpoint cluster region protein, Bcr (Braselmann and McCormick, 1995)), receptors (insulin-like growth factor I receptor, IGFIR (Craparo et al., 1997), and Nuclear receptor, Nur77 (Masuyama et al., 2001)) and nuclear proteins and transcription factors (p53 (Waterman et al., 1998), Histone deacetylase 4, HDAC4 (Wang et al., 2000), Histone deacetylase 5, HDAC5 (McKinsey et al., 2000), transcription co-activator with PDZ-binding domain, TAZ (Kanai et al., 2000)).

All the 14-3-3 isoforms contain a phosphoserine- and phosphothreonine-dependent binding motif (reviewed in Fu et al., 2000), suggesting that the phosphorylation of the ligand is a determinant for 14-3-3 to interact with co-factors. Specifically Muslin and colleagues defined the motif RSxpSxP (x represents any amino acids and pS a phosphorylated Serine) as conserved in several 14-3-3 ligands and able to interact with the amphipathic groove of 14-3-3 proteins. Proteins like Raf1 (Muslin et al., 1996; Morrison and Cutler, 1997), Bad (Zha et al., 1996; Hsu et al., 1997), Cdc25 (Peng et al., 1997; Kumagai et al., 1998; Yang et al., 1999; Lopez-Girona et al., 1999; Dalal et al., 1999), PTPH1 (Zhang et al., 1997) and others, have the conserved RSxpSxP motif and this mediates their binding to 14-3-3 proteins.

Three residues in particular are highly conserved in the amphipathic groove among the seven isoforms, namely Lys49, Arg56 and Arg127 (highlighted in red in Fig. 5). Both, mutational and crystallization studies have shown how these three amino acids are essential for 14-3-3 proteins to interact with their phosphorylated substrates (Zhang et al., 1997; Wang et al., 1998; Thorson et al., 1998; Yaffe et al., 1997; Rittinger et al., 1999; Petosa et al., 1998).

In spite of a preferential binding for phosphorylated proteins, some studies showed that 14-3-3 proteins can also bind unphosphorylated ligands (reviewed in Obsilová et al., 2008). Examples of proteins binding 14-3-3 in this unconventional way are the ExoS ADP-rybosyl-transferase, 5-phosphatase and Raf-1. The latter in fact contains in addition to a RSxpSxP motifs, a 14-3-3 binding site consisting in a cysteine-rich domain (CRD) (Braselmann and McCormick, 1995). In order to achieve a high affinity binding, negatively charged Glutamic acid residues are needed (reviewed in Fu et al., 2000; Pozuelo Rubio et al., 2004).

When 14-3-3 proteins form dimers, the amphipathic groove of both monomers forms a hole of 6-8 Å that allows the simultaneous binding of two ligands. This finding was confirmed by co-crystallization experiments that show how the concave surface formed by the dimer is occupied on each site. An important implication of this finding is that 14-3-3 proteins can act as bridging factors for two different ligands (reviewed in Fu et al., 2000).

1.5.1 Role of 14-3-3 adaptor proteins in Antibody diversification

Being 5'-AGCT-3' motifs specifically dense in S-regions and preferentially targeted by AID, Xu and colleagues performed affinity purification assays to discover possible binding partners (Xu et al., 2010). Using 5'-AGCT- 3' oligonucleotides as bait they have been able to identify that all seven 14-3-3 adaptor proteins isoforms can bind to the motif. 14-3-3 proteins are found to be highly expressed in germinal center B cells, as well as in human and mouse B cells stimulated to undergo CSR by either T-dependent (CD40 engagement plus IL4) and T-independent (lipopolysaccharide, LPS, Resiquimod, R-848, and CpG) stimuli (Mai et al., 2013). The kinetic of induction is pretty rapid. In fact, when B cells are activated with LPS, 14-3-3γ transcript levels peak after only 3 hours and sustain the process of switching by staying steadily expressed after 48 hours (Mai et al., 2013). ChIP experiments showed that 14-3-3 proteins bind those S-regions that are involved in undergoing CSR, depending on the activation stimulus and that they co-localize with AID and the catalytic subunit α of protein kinase A (PKA Cα) (Xu et al., 2010). Blocking 14-3-3 using difopein or genetic deletion impair AID recruitment and therefore CSR, showing that 14-3-3 adaptor proteins are indispensable for switching (Xu

et al., 2010). Several attempts in trying to understand the role of 14-3-3 proteins at the S-regions led to the conclusion that they function as scaffold protein necessary to build the CSR machinery (Lam et al., 2013). Bimolecular fluorescence complementation (BiFC) assays show that the 14-3-3 protein family directly interacts with AID and targets it to the correspondent switch region (Xu et al., 2010). Very interestingly, it cannot bind the AID C-terminal truncation variant AIDΔ (190-198) (Xu et al., 2010) as well as the AIDF193A and AIDL196A mutants, known to be CSR-defective (Lam et al., 2013). Toghether with AID, 14-3-3 are also able to bind both the cathalitical and regulatory subunits of PKA (PKA Cα and PKA RIα), as well as UNG, in a direct manner (Lam et al., 2013). This shows that 14-3-3 proteins act as scaffold, with no enzimatic function, in order to mediate the assembly and nucleation of the CSR machinery (Lam et al., 2013).

A detailed analysis of the 14-3-3 γ promoter in mouse shows that it only has two Transcription Start Sites (TSSs), but instead of a TATA-box element, it is dense in CpG islands. Those motifs interact with the CXXC finger protein 1 (CFP1) structural subunit of histone methyltransferase complex COMPASS. This complex guarantees an accumulation of histone 3 lysine 4 trimethylation (H3K4me3), marker for an open chromatin conformation. Mai and colleagues show that both the canonical and non-canonical NF-κB pathways, which are also one of the transcriptional activator of AID expression (Tran et al., 2010), enrich for CFP1 at the 14-3-3 γ promoter, also promoting the binding of the B cell lineage specific factor, E2A (Mai et al., 2013). Therefore the expression of 14-3-3 γ in B cells undergoing CSR is guaranteed by an open chromatin state at the promoter, mediated by CFP1 and induced by NF-κB and the transcription factor E2A (Mai et al., 2013). It was still poorly understood how 14-3-3 proteins can dock onto the 5'-AGCT-3' motifs, until the role of the co-transcriptional modification H3K9acS10ph was investigated. Thanks to RNA Polimerase II stalling, also known to be important for AID targeting (Pavri et al., 2010), histone modifications occur and favor the binding of 14-3-3 to the ssDNA. Thanks to the highly conserved residues in the amphipathic groove, Lys49 (K49), Arg56 (R56) and Arg127 (R127), 14-3-3 proteins can interact as a dimer with both the ssDNA directly and the combinatorial H3K9acS10ph modification, functioning as "readers" of the epigenetic information. R56 and R127 on one 14-3-3 monomer bind to the phosphorylated Ser10 of the H3. The steady docking of

14-3-3 onto ssDNA is ensured by K49 on the other 14-3-3 monomer, which binds directly the 5'-AGCT-3' motif (Li et al., 2013). All the studies on the 14-3-3 regulation and function in antibody diversification were so far focused only on class switching. Understanding if 14-3-3 is able to regulate somatic hypermutation through a similar mechanism still needs to be investigated.

1.6. Aim of the study

A big effort has been put so far in order to elucidate the molecular mechanisms of antibody diversification. Nonetheless various aspects remain still unresolved: (1) What is the mechanism of AID targeting to the *Ig* variable regions? (2) What determines the discrete mutation spectrum? (3) Why are mutation rates so much higher than in non Ig-genes? (4) How much divergent SHM and CSR are?

Previous studies show the involvement of numerous transacting factors in the AID targeting to the *Ig* locus. During CSR, looping of the locus due to formation of secondary DNA structures has been shown to have a decisive role in AID targeting (Kinoshita et al., 1998) (Tashiro et al., 2001). Nonetheless the profound different nature of V regions compared to S regions suggests a potential different mechanism underlying SHM. Several studies have been focused in understanding if the IgH V region context favours mutations *per se*. Overlapping AGCT motifs in V-regions represent an entry point for AID to initiate the mutation cascade (Wei et al., 2015). It is well known that CDR regions are preferentially mutated (Wei et al., 2015), but being them the actual antigen's binding site, it has been hard to differentiate mutations coming from intrinsic AID activity and diversification due to antigen selection. The use of cell lines and *ex vivo* activated B cells is limited due to the very low mutation rates acquired over time (Nagaoka et al., 2002; reviewed in Liu and Schatz, 2009; Maul et al., 2014). Moreover, the attempt of boosting SHM using hyperactive AID variants often results in a non-physiological AID mutation signature and targeting specificity (Wang et al., 2009; Ito et al., 2004). Studies in SHM have been revolutionized after the introduction of non-productive *Ig* transgenes (Betz et al., 1993) that allow accumulation of mutations without B cells undergoing selection. These systems gave crucial insights on the SHM mutation spectrum and the hotpots of SHM (reviewed in (Di Noia and Neuberger, 2007) but they fail in recapitulating the

genomic context. A breakthrough came from Alt and colleagues who generated a mouse model where a "passenger" unrelated sequence is placed on one endogenous IgH allele, whereas the second allele stays productive and favours normal selection, therefore allowing B cell development and germinal center formation (Yeap et al., 2015). They conclude that the genomic context influences high rates of SHM, that mutation load is sequence-dependent and confirmed that AID hotspots serve as entry point to the locus (Yeap et al., 2015).

In this study we develop a tool platform to screen for putative SHM cofactors. We took advantage of the Ramos B cell line, an IgM positive human Burkitt Lymphoma line that constitutively expresses AID and undergoes spontaneous SHM (Sale and Neuberger, 1998). Ramos cells have been traditionally used to study SHM also because they allow a semi-quantitaive read-out of hypermutation by losing IgM expression when mutating (Sale and Neuberger, 1998). The line has been first engineered to have an inducible and enhanced hypermutation by expressing the JP8Bdel AID variant, an AID deleted mutant lacking 16 aa at the C-terminal, corresponding to the nuclear export signal (NES) (Ito et al., 2004), see Fig. 4). We then optimized a straightforward flow cytometry-based SHM assay, where the IgM loss rates correlate with rates of hypermutation. The JP8Bdel Ramos line recapitulates the physiological SHM spectrum and mutation rates, without altering the AID targeting specificity. This robust *in vitro* system has been finally used as a platform for an RNAi targeted screen, which revealed the 14-3-3 adaptor proteins, as well as other interesting hits, as putative AID co-factors in targeting the V-regions during SHM. Some initial pull-down experiments have been performed in order to define the molecular mechanism through which 14-3-3 proteins regulate SHM. Preliminary results show that 14-3-3 β interacts with the transcription elongation factor SPT6 and this binding only happens in the nucleus.

3. RESULTS

3.1 Ramos B cell line as an *in vitro* model to study Somatic hypermutation

Historically most of the studies on antibody diversification focused on Class switch recombination only, because of the absence of a robust model to investigate SHM. First goal of the project was to develop a SHM Assay where we can finely control the mutation patterns and mimic the *in vivo* germinal center mutation rates.

3.1.1 JP8Bdel-AID efficiently boosts IgM loss in Ramos B cells

We took advantage of the Ramos cell line, traditionally one of the models to study somatic hypermutation. SHM in Ramos often results in stop codons, leading to surface IgM negative variants (Fig. 7 A). Michael Neuberger's laboratory showed how Ramos cells constitutively mutate in culture and, as *in vivo*, mutations are preferentially nucleotides substitutions, biased towards G/C pairs (Sale and Neuberger, 1998). However poor IgM loss rates make Ramos not suitable for our screening. In fact, after two weeks of culturing only 2,85% of cells become negative for IgM (Fig. 7 B).

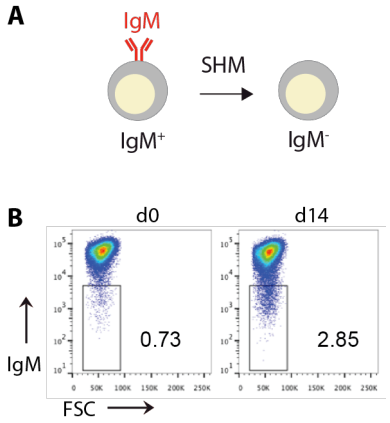


Fig. 7: SHM in Ramos cells leads to loss of surface IgM. (A) Schematic representation of Ramos cells forming IgM- variants when they undergo SHM. (B) Representative FACS plots of wt Ramos stained for IgM, after two weeks (d14) of culturing. Numbers indicate percentages of IgM- cells.

Being a genome mutator enzyme, AID, under physiological conditions, is primarily located in the cytoplasm (Fig 8 A). We aimed to generate a line where, for experimental purposes, AID can be induced to translocate in the nucleus (Fig. 8 B). Therefore we exploited the Estrogen receptor (ER) fusion system: we transduced Ramos cells with a fusion protein AID-ER, in a way that 4-Hydroxytamoxifen (4-HT) treatment will force AID to translocate into the nucleus (Fig. 8 C).

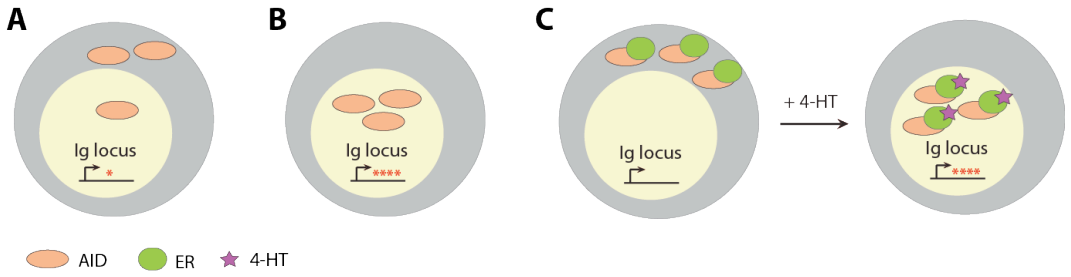


Fig. 8: AID can be induced to translocate to the nucleus by the ER-system. (A) Cartoon showing the physiological subcellular localization of AID. (B) Our goal is to achieve a massive AID gathering in the nucleus, where it can heavily mutate the Ig locus (red asterisks indicate mutations). (C) Schematic of the implemented model, where AID is fused with the Estrogen receptor (ER) and its translocation into the nucleus triggered by 4-Hydroxytamoxifen (4-HT) (C)

Preliminary experiments aimed firstly to make Ramos cells proficient for retroviral infection. The wild-type line in fact gave us very poor infection efficiency (1.64 % infection rate) when transduced with pMX-IRES-GFP (or mCherry) vectors, using pCL-Ampho retrovirus packaging vector as helper (Fig. 9 A). We therefore engineered a Ramos cell line stably expressing the ecotropic receptor (referred to as "Ramos RIEP"). This is an established method to effectively transduce human cells with ecotropic virus that otherwise could only infect murine cells. When pCL-Eco retrovirus packaging vector is used as helper, infection rates rise to 24.6% (Fig. 9 B).

Ramos RIEP cells are not only more competent for infection, but they also constitutively express rtTA (Reverse Tetracycline-controlled Transactivator). As an element of the tetracycline ON system (Gossen et al., 1995), this will be needed later on for the screening procedure, to be used in combination with a plasmid containing a tetracycline-dependent promoter, upstream of the gene of interest.

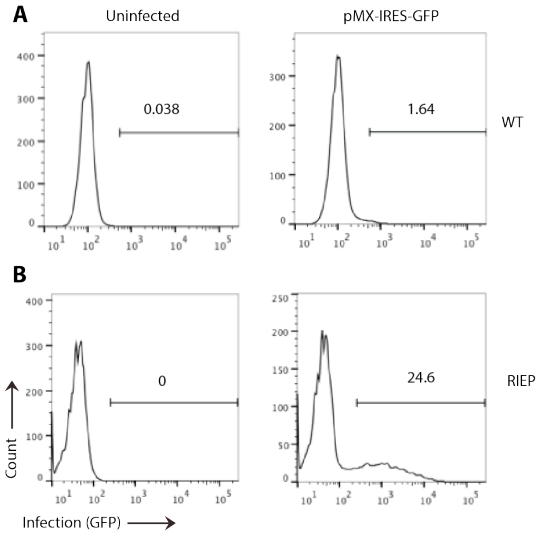


Fig 9: The ecotropic receptor boosts infection rates in Ramos cells. (A) FACS plots showing the percentage of infected wt Ramos cells. (B) Transduction efficiency in RIEP Ramos. Infection efficiency with pMX-IRES-GFP is measured by FACS two days after transduction, by measuring GFP levels.

In the germinal center mutations accumulate with a rate of incorporation of nucleotide substitutions into the V exon at between 10^{-4} and 10^{-3} per base pair (bp), at each cell division (McKean et al., 1984; Berek and Milstein, 1988). Trying to mimic the *in vivo* scenario, we tested two different constructs, with the final aim of enhancing SHM: AID-ER-Ugi and JP8Bdel-ER-Ugi (referred to as JP8Bdel-AID). In both AID is fused with ER and Uracil DNA Glycosilase (UNG) inhibitor, Ugi. UNG is the main player for the base excision repair pathway (BER) acting downstream of the AID-induced damage to repair the DNA, in an error-free manner (Di Noia and Neuberger, 2002) (Fig. 10).

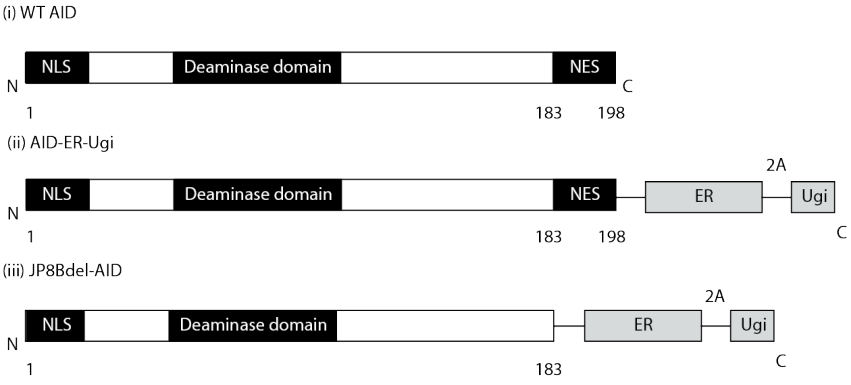


Fig. 10: Domain representation of the AID constructs in use. (i) wt AID; (ii) AID-ER-Ugi; (iii) JP8Bdel-ER-Ugi. Numbers indicate the amino acid positions.

JP8Bdel-ER is a truncated variant of AID (1-183 amino acids) lacking the nuclear export signal (NES) (Ito et al., 2004). Once 4-HT is added to the culture, AID will get sequestered into the nucleus, where it massively mutates the *IgH* locus (Fig. 8 C).

The constructs are tested in Ramos RIEP cells for IgM Loss after 6 days of 4-HT treatment. To fully characterize the AID variants efficiency, after 2 days of infection, cells were sorted for GFP highly (GFP^{hi}) and intermediate (GFP^{int}) expressing cells, with GFP as marker of infection. An empty plasmid and a wild-type AID construct are used as controls. JP8Bdel-ER is strongly boosting reversion to IgM negative variants, whereas the UNG inhibitor Ugi does not cause a more potent effect. IgM loss is more prominent in GFP^{hi} cells, therefore high expression of JP8Bdel-AID guarantees high levels of SHM. Nonetheless, it also causes a notable leakiness (Fig. 11).

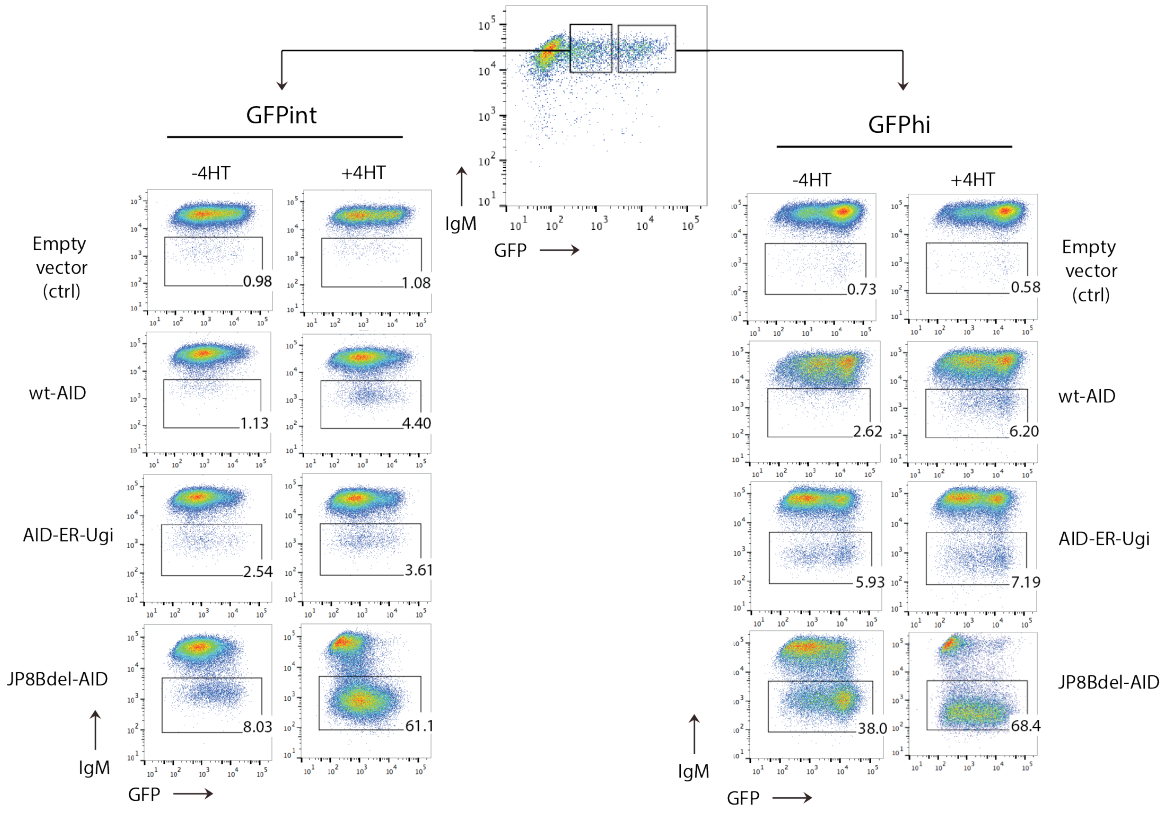


Fig. 11: JP8Bdel-AID efficiently boosts the IgM loss. IgM loss assay is performed in Ramos RIEP cells. Cells are infected with the indicated set of constructs and sorted for GFP highly and GFP intermediately-expressing populations, two days after transduction. After six days of 4-HT

treatment, they are stained for IgM and percentages of IgM loss are assessed by FACS measurement. Results from cells 4-HT untreated are also plotted.

To overcome this limitation, the JP8Bdel-AID construct was cloned under a mutated Kozak sequence. Lowering transcription of the gene of interest, we managed to drastically reduce the leakiness by 4-fold. Nevertheless IgM loss levels are still strikingly high (40.4%) (Fig. 12).

We generated a line stably expressing JP8Bdel-AID under the control of a mutated Kozak sequence (mk JP8Bdel-AID line). Both an antibiotic resistance (Puromycin) and a fluorescent marker (mCherry) allowed an efficient selection of the desired clone. Altogether these data show that the mk JP8Bdel-AID line has an enhanced and inducible somatic hypermutation and it will therefore be used for the following experiments.

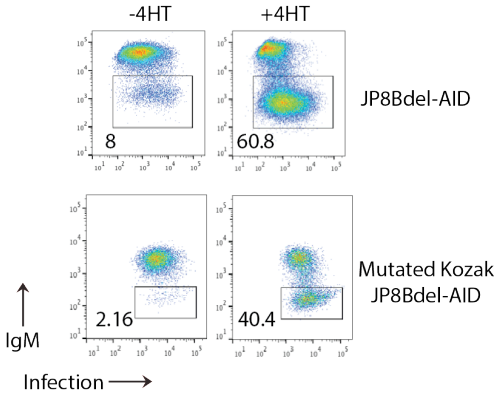


Fig. 12: A mutated Kozak sequence ensures less leakiness of JP8Bdel-AID. FACS profile of IgM loss after 6 days of 4-HT treatment. Cells transduced with JP8Bdel-AID and Mutated Kozak JP8Bdel-AID are compared. Results from cells 4-HT untreated are also plotted.

3.1.2 IgM loss in the JP8Bdel-AID line correlates with rates of hypermutation

In addition to test the line for efficient IgM loss, a full characterization considering a detailed analysis of the mutation levels and the mutation spectrum is needed. For this purpose, a published sequencing approach is adapted (Robbiani et al., 2015). MutPE-seq consists in a mutational analysis by sequencing that measures SHM at specific genomic sites. 250 bp fragments and a paired-end sequencing approach guarantee a full coverage of the amplicon (Fig. 13 A, left). After four days of 4-HT treatment, sorting of IgM positive (IgM^+) and IgM negative (IgM^-) cells is performed. The two samples, together with the bulk population, were treated independently for library preparation and sequencing. Five regions in total in the IgH locus are analysed: three of them spanning

the whole VDJ region and two mapping the Switch and Constant regions respectively. The last is included as a control for AID off-target activity (Fig. 13 A, right).

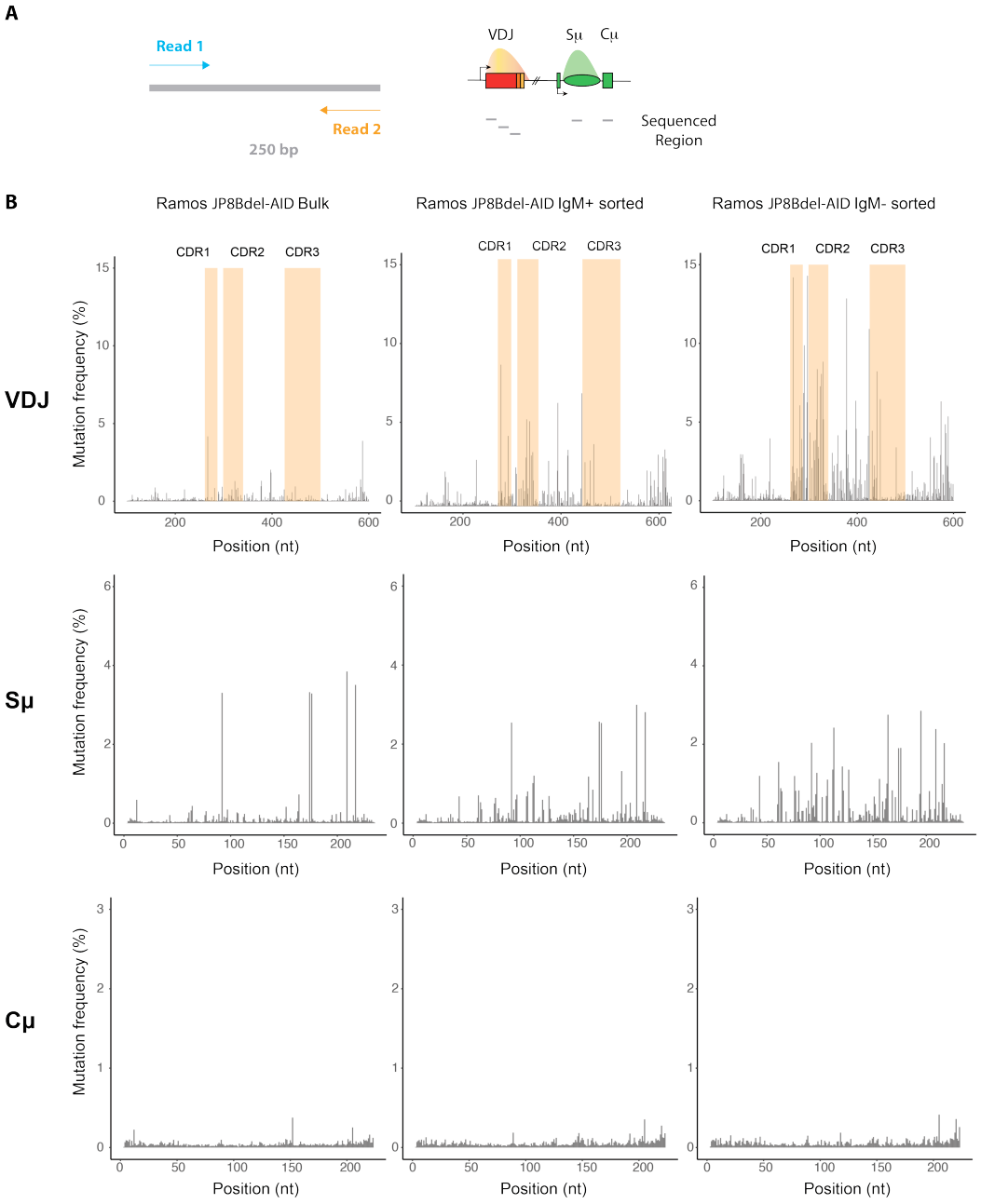


Fig. 13: Mutation profiling of Ramos JP8Bdel-AID line by MutPE-seq (I). (A) Schematic illustration of the sequencing strategy (left); representation of the sequenced regions and their mapping in the IgH locus (right). (B) MutPE-seq is performed on Ramos JP8Bdel-AID cells, sorted for IgM^+ and IgM^- populations, after six days of 4-HT treatment. Ramos JP8Bdel-AID bulk population is used as control. The mutation levels are represented as number of mutated reads for

each nucleotide position (nt). Mutations at VDJ, Sμ and Cμ regions are shown. In the VDJ region, orange boxes represent the complementary determining regions, CDR1, CDR2 and CDR3.

Bar plots show the mutation frequency for each nucleotide position. It is appreciable how IgM^- sorted cells have more mutations in the VDJ exon, when compared to the bulk population, as well as to the IgM^+-sorted cells. In particular mutations rise mostly in the CDR1 and CDR2 (Fig. 13 B). Mutations increase as well in the Sμ region of the IgM^- population, although not as dramatically as for the VDJ region (Fig. 13 B and 14 C). Very importantly, SHM levels in the Constant region are barely detectable for both samples and they do not rise in the IgM^- cells (Fig. 13 B).

Hypermutation in Ramos cells has a bias for mutations in C/G pairs. To check if the AID signature is retained in the presence of JP8Bdel-AID, we look into the distributions of mutations among different nucleotides. We found that both in VDJ and Switch region, mostly Cs and Gs are mutated, following the classical AID signature (C to T and G to A mutations). On the contrary in the constant region there is no bias for specific nucleotides (Fig. 14 A). When the overall mutation frequencies at VDJ and Sμ region are calculated, it is appreciable how only C>T and A>G mutations are increasing, whereas other mutations stay the same (Fig. 14 B and C). Altogether, these findings show that the JP8Bdel-AID line recapitulates a physiological somatic hypermutation spectrum, AID targeting specificity and mutation signature.

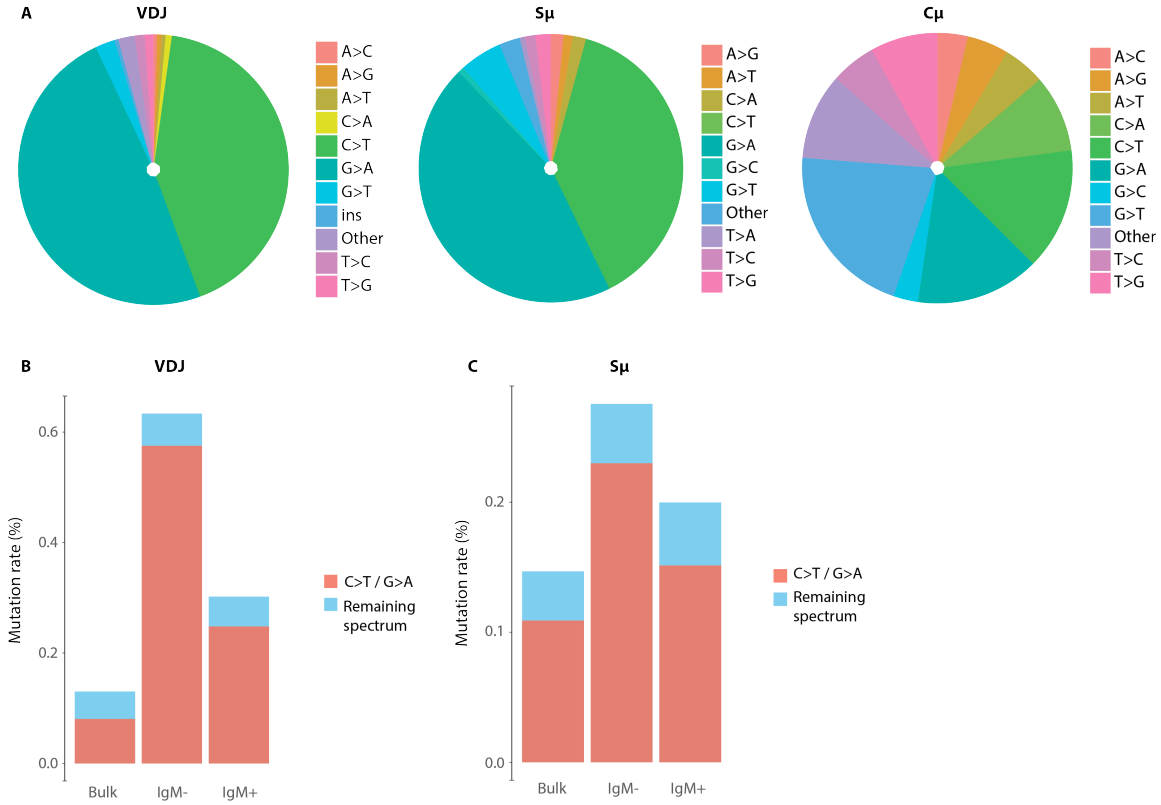


Fig. 14: Mutation profiling of Ramos JP8Bdel-AID line by MutPE-seq (II). (A) Pie charts depicting the proportion of clones of Ramos JP8Bdel-AID IgM$^-$ cells that contain the indicated nucleotide substitution. The distribution of mutations is shown for the VDJ, Sμ and Cμ amplicons. (B) Mutation rates plotted as percentages of number of mutations, normalized to the number of bases in the VDJ region and (C) in the Sμ region.

3.2 A targeted RNAi screening to investigate on SHM molecular players

Having optimized a powerful tool to assess somatic hypermutation in Ramos cells and a cell line recapitulating physiological SHM pattern, we performed a targeted RNAi screen to investigate on SHM co-factors.

3.2.1 Screening design and optimization

One of the first steps of optimization included the choice of the plasmid backbone to use. This was a close collaboration with Johannes Zuber laboratory, which provided us with different constructs and short hairpins-RNAs (shRNAs) sequences. Hairpins targeting essential genes (referred to as killer hairpins) like RPA3 (RPA3.1401), RRM1 (RRM1.1119), c-Myc (Myc.1891 and Myc.1888), together with a hairpin targeting Renilla (Ren.713), a gene not expressed in mammalian cells, are cloned in two different backbones, namely LENC and RT3GEN. LENC has a stable promoter driving the expression of the shRNA, followed by a PGK promoter encoding for the Neomycin resistance gene and mCherry fluorescent protein expressed using an IRES sequence (Fig. 15 A). RT3GEN contains a Doxycycline inducible promoter (T3G), used in the Ramos RIEP (rtTA+, see text above and Fig. 9 B) cell line, which will drive the expression of both GFP and the hairpin. RT3GEN has similarly the Neomycin gene, allowing cell selection using G418. (vector schematics in Fig. 15 A and 15 B). Frequency of transduction was measured after two, four, six and eight days of infection. The percentage of cells positively infected with the killer hairpins should logically decrease overtime, when compared to the Renilla control.

Data show a more drastic drop of positively infected cell percentages when transduced with killer shRNAs cloned into the RT3GEN backbone, compared to LENC backbone, while Ren.713 infected cells have a steady GFP expression overtime (Fig. 15 B). In particular, hairpins targeting RPA3 and RRM1 are the ones showing the highest defect. Therefore, we decided to clone the screening library into the RT3GEN plasmid and to use hairpins against RPA3 and RRM1 as quality control of the screening. Very importantly, the percentage of RPA3 positively transduced cells decreases even under selection with

G418, excluding that the positive pressure of selection overcomes the effect of the killer hairpin (Fig. 15 C).

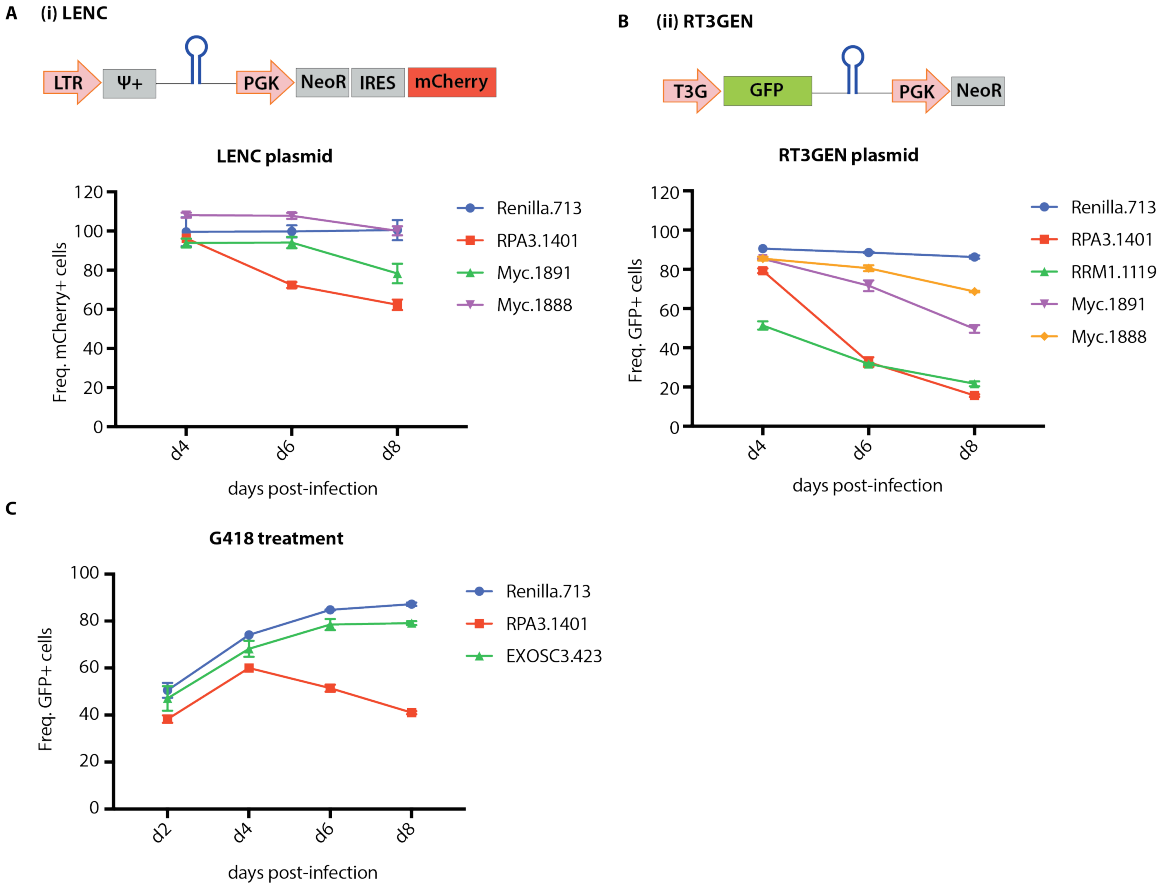


Fig. 15: Comparison of plasmid's efficiency. (A) Scheme of LENC vector. Infection efficiency of killer hairpins and Renilla.713 is measured by FACS. Data are plotted as frequency of positively (mCherry+) transduced cells (bottom). Values at time-points d4, d6 and d8 are normalized on d2 $mCherry^{+}$ cells levels. (B) Scheme of RT3GEN (top) vector. Infection efficiency of killer hairpins and Renilla.713 is measured by FACS. Data are shown as frequency of positively (GFP+) transduced cells (bottom). Values at time-points d4, d6 and d8 are normalized on d2 GFP^{+} cells levels. Arrows boxes in (i) and (ii) show the promoters. (C) Frequency of positively infected cells after transduction with Renilla.713, RPA3.1401 and EXOSC3.412, all clone into RT3GEN backbone vectors, measured after two, four, six and eight days of selection with G418.

After setting up the lethal shRNAs controls, a very important step was to choose positive controls that will make the screen reliable. Studies on SHM co-factors, at the step of AID targeting, were so far very limited. Therefore the choice of a good positive control for the screening was very restricted.

It has been already shown how AID is acting in a transcription dependent manner and that DNA targeting is occurring at sites where Polymerase II is stalling by interacting with the transcription elongation factor SPT5 (Pavri et al., 2010).

Moreover depletion of the Exosome complex subunits 3 (EXOSC3) and 10 (EXOSC10) were shown to cause a mild defect in SHM (Pefanis et al., 2015). We therefore focus our interest in knocking down those three factors and test them as putative positive controls. Four hairpins are tested for SPT5 and EXOSC3 and five for EXOSC10. Ramos mk JP8Bdel-AID line is infected with RT3GEN-shRNA and on the same day of the infection the culture medium is supplemented with Doxicycline and G418. Six days after infection one million cells are harvested to confirm protein depletion after knock-down (Fig. 16 A). Immunoblot results allowed us to conclude that EXOSC3.412, EXOSC3.423, EXOSC10.1509 and EXOSC10.2066 cause a severe protein depletion. In the case of SPT5-targeting hairpins, the knock-down of the protein is very poor. Nonetheless, we followed up in assessing IgM loss levels in SUPT5H.534 and SUPT5H.3479 (Fig. 16 B).

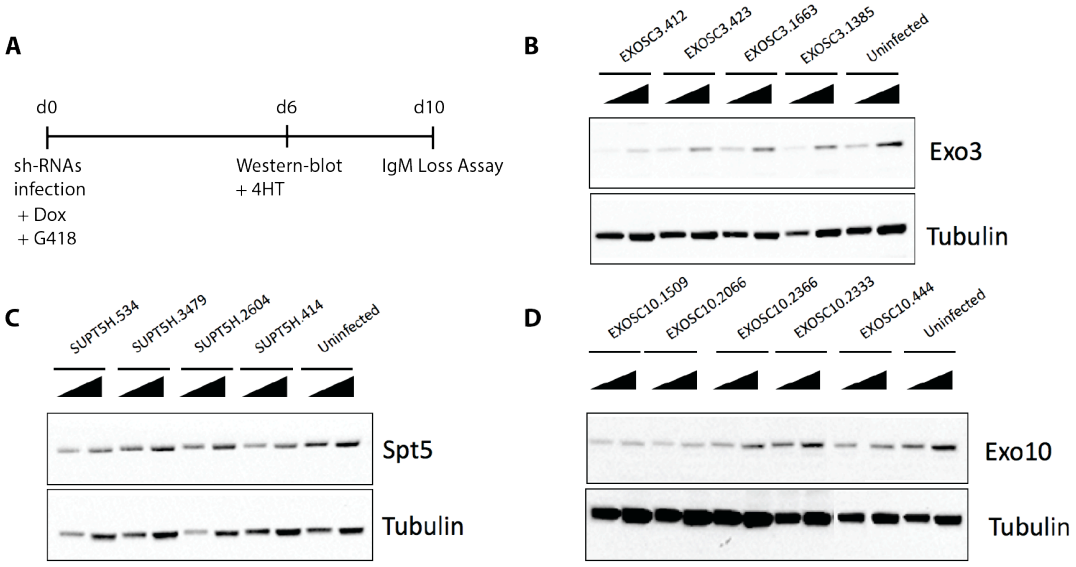


Fig. 16: Optimization of positive controls for the screening. (A) Experiment outline for testing positive controls. (B) Whole cell lysates are immunoblotted using antibodies for SPT5, EXO3 and EXO10. Triangles indicate crescent amount of protein loaded on the gel (10 µg and 30 µg). Uninfected samples are used as controls. Tubulin is probed as loading control.

We also assessed protein levels after twelve days of knock-down to monitor the expression depletion over a longer time, but no stronger expression defect was observed (data not shown). Cells transduced with the best scoring hairpins are treated with 4-HT, in order to induce SHM. After 4 days, cells undergo flow cytometry to assess IgM loss. Results show how SPT5 depletion is not sufficient to cause an IgM loss defect. In contrast, both shRNAs targeting EXO3 and one targeting EXO10 reveal less cells undergoing hypermutation, when compared to the Renilla control (Fig. 17 A and B). Very importantly, targeting EXO3 doesn't show any lethal effects (Fig. 15 C).

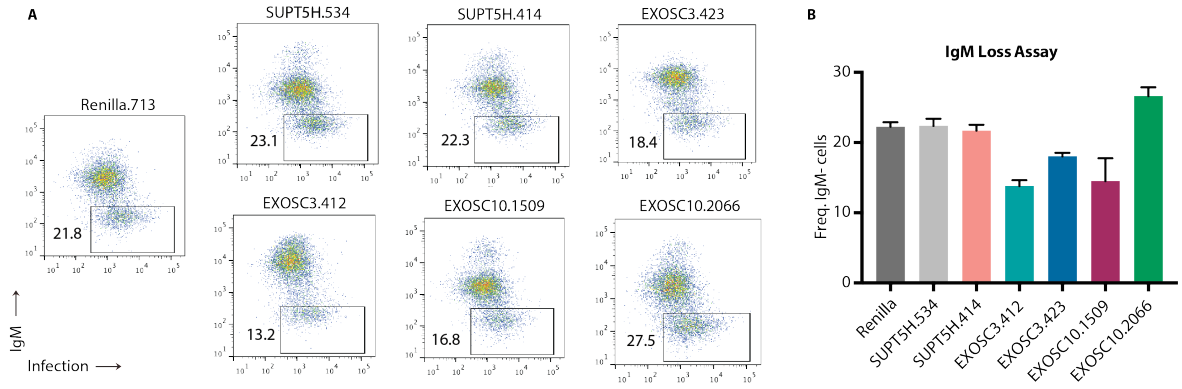


Fig. 17: EXO3 and EXO10-targeting shRNAs dampen IgM loss in JP8Bdel Ramos cells. (A) Representative FACS profile of IgM loss after 6 days of knock-down and 4 days of 4-HT treatment. (B) IgM Loss Assays from three independent experiments. Percentages of IgM^- cells are plotted. Error bars indicate Standard Deviations (SD).

4-HT dosage and length of treatment was a crucial point to investigate. We addressed the IgM loss defect in Ramos EXO3 and EXO10-depleted cells, after 2, 4, 6 and 8 days of SHM and found that a more pronounced defect appears after four days of 4-HT treatment (Fig. 18 A). This evidence, together with MutPE-seq results (Fig. 13 and Fig. 14), suggested us to perform four days of treatment in the actual screening. In addition, to rule out that the standard 4-HT dosage (0.4 µg/ml) is too high and results in massive translocation of JP8Bdel-AID in the nucleus, that could mask the knock-down effect, we titrated it down. We could confirm that 4-HT at 0.4 µg/ml results in the most appreciable IgM loss defect, after depletion of EXO3 and EXO10 subunits (Fig. 18 B).

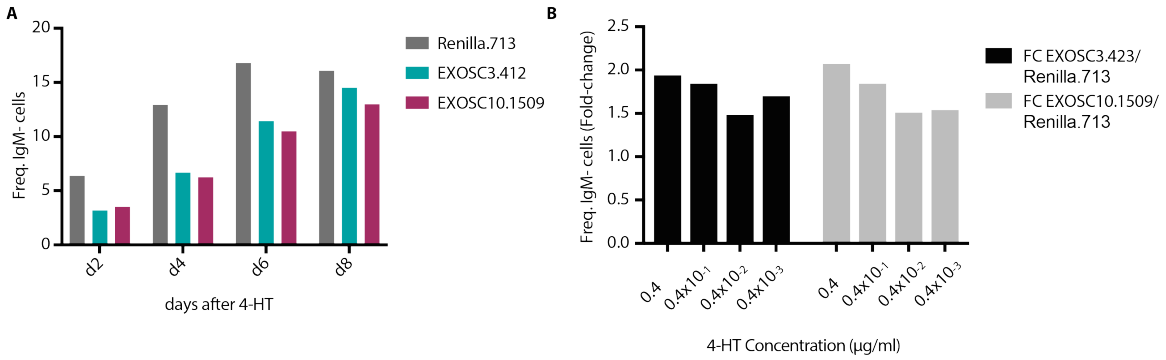


Fig. 18: Optimization of the treatment with 4-HT. (A) Frequency of IgM$^-$ cells, measured by FACS after 2, 4, 6 and 8 days of 4-HT treatment. (B) IgM Loss Assay after titration of 4-HT. 10^{-1}, 10^{-2} and 10^{-3} are serial dilution from the starting concentration of 0.4 µg/ml. Results are plotted as the fold-change of frequency of IgM- cells after infection with EXOSC3.412 and EXOSC10.1509 calculated on Renilla.713.

All together, these optimization experiments define important specifics for the RNAi screening. We were able to define that the RT3GEN backbone vector is the most suitable for the RNAi screen; Renilla (Ren.713) and RPA3 (RPA3.1401) together with RRM1 (RRM1.1119) will respectively be negative and killer controls, whereas EXO3 (EXOSC3.412) and EXO10 (EXOSC10.1509) the positive ones;
4-HT treatment will be performed at a concentration of 0.4µg/ml and it will last 4 days.

3.2.2 14-3-3 adaptor proteins are required for somatic hypermutation

The shRNA library we designed consists of 96 genes, each targeted by four different hairpins (384 shRNAs in total) (Table S3). The pool of genes includes a negative control (Ren.713) as well as killer hairpins (RPA3.1401 and RRM1.1119). In order to decide which genes to include, among genes expressed in Ramos cells (derived from RNA-seq data, published by Qian at al., 2014), we made an educated guess choosing factors known to play a role in AID targeting in the context of CSR, as well as transcription factors and chromatin regulators. The most efficient procedure considered cloning all the hairpins as a pool, we could not reach full library coverage, but a final coverage was 75%.

The final screening outline considers transducing the JP8Bdel-AID Ramos line with shRNAs cloned into the RT3GEN plasmid, using a one to one approach. On the same day of the infection, 8 hours later, both Doxicycline, to induce the expression of the hairpin as

well as GFP, and G418, to select for positively infected cells, are added in culture. At day 6 we add 4-HT to let JP8Bdel-AID translocate into the nucleus and initiate SHM. After 4 days of treatment (day 10) the IgM Loss Assay is performed (Fig. 19).

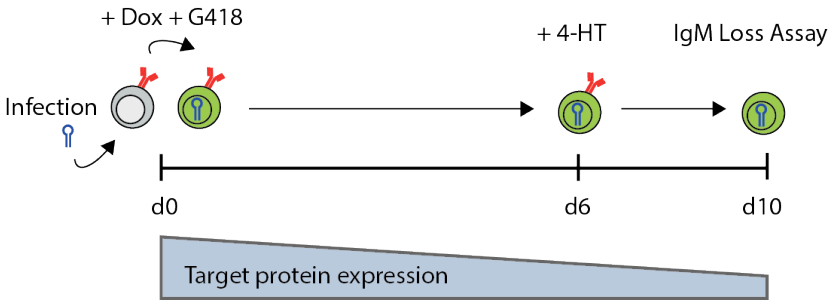


Fig. 19: Finalized experimental outline of the screening. After infection with single hairpins cloned into the RT3GEN backbone, the expression of both GFP and the shRNA is induced by treatment with Doxycyxline (Dox). After six days of knock-down, while the cells undergo selection with G418, SHM is induced by adding 4-HT and the IgM loss assay is performed after four days of 4-HT treatment (d0, d6 and d10 are zero, six and ten days post-infection).

For technical limitations the screening was performed in six rounds of infection (each using roughly 47 shRNAs). We made sure to include Ren.713, RPA3.1401, RRM1.1119 and EXOSC.423 in each round, as a proof of quality for the screen (Table S4). Each hairpin was tested in triplicates. GFP levels are measured over time (data not shown) and in the case of killer hairpins, for each round we checked they are consistently decreasing when compared to Renilla (Fig. 20).

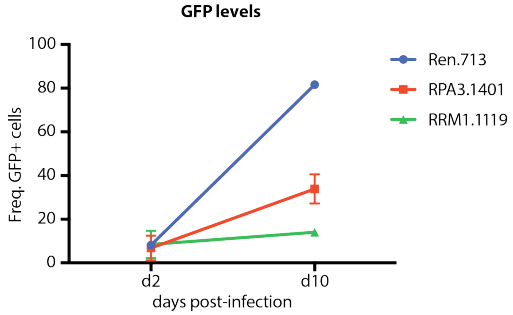


Fig. 20: RT3GEN-killer hairpins do not enrich for GFP throughput the selection process. GFP levels are measured at day two and day ten of selection, after infection with Renilla.713, RPA3.1401 and RRM1.1119.

Ren.713 is used as a reference in each of the experiments. In fact, after gating for successfully transduced cells ($mCherry^{+}GFP^{+}$), we looked at the IgM expression, calculating the ratio between IgM^{+} and IgM^{-} cells and expressing it as fold-change on Renilla. We considered a hairpin as scoring only when it was increased by two-fold (Fig.

21 A). Detailed results are shown as tables (Table S4), as well as bar-plots (Fig. S6).

As depicted in the summary figure (Fig. 21 B), many hairpins score above the cut-off. Among the best scoring shRNAs the 14-3-3 adaptor proteins stand out and they are the ones we decided to follow-up on. 14-3-3 adaptor protein family's subunits β, γ and τ, encoded respectively by YWHAB, YWHAG and YWHAQ genes, interact with AID during class switch recombination (Xu et al., 2010). Although a full validation was performed only for the 14-3-3 adaptors (next chapter), preliminary data on another set of proteins scoring in the screen, Lim domain-binding protein 1 (LDB1) and CCCTC-binding factor (CTCF), are shown in the Supplementary section. All together these findings show that the targeted SHM RNAi screening successfully identified potential AID co-factors.

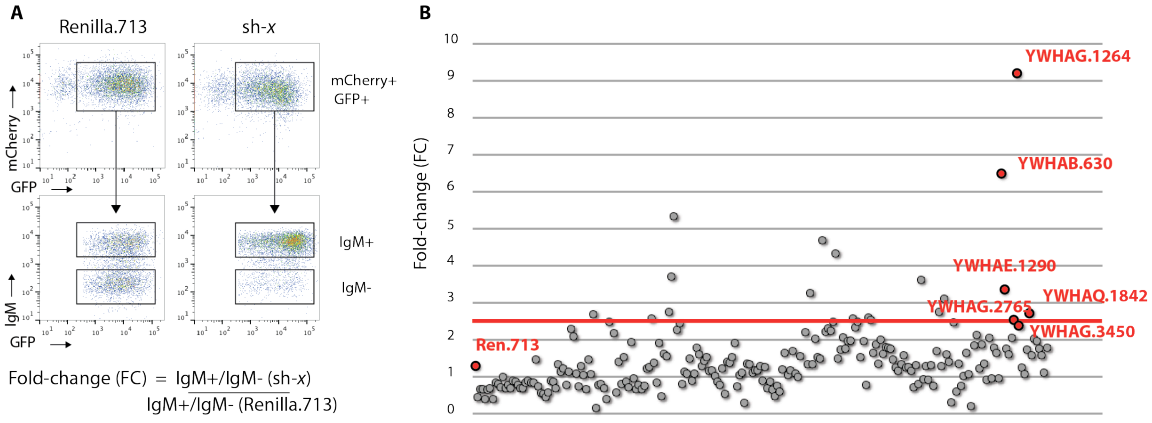


Fig. 21: A targeted RNAi screening reveals 14-3-3 proteins as putative regulators of SHM. (A) Gating strategy of the screening. In the process of analysis of the flow cytometry data the gates are set by showing the IgM^+ and IgM^- populations, within the $mCherry^+GFP^+$ cells. The plots show an example case (sh-*x*) compared to Renilla.713. Screening data are presented as fold-change of the ratio of IgM+ on IgM- cells, over Renilla.713 values. (B) Dot-plot summarizing results of the six rounds of screening. The average of Renilla.713 ratios in each round is set as reference. The cut-off is set to 2-fold enrichment, compared to Renilla.713. shRNAs chosen for follow-up validation are labeled with red dots.

3.3 Hit validation

The 14-3-3 adaptor proteins scoring in the screen underwent further steps of validation. The knock-down experiment was performed in small scale, aiming to reduce any potential technical error. We made sure that G418 selection of the cells is not resulting in an artefact (false positive) by performing the experiment on selected and unselected Ramos, in parallel. Results confirmed a defect in IgM loss, even when the culture is not pushed towards positively infected cells (Fig. 22 A). Moreover, we confirmed significant protein depletion compared to control, after 6 days of knock-down (Fig. 22 B).

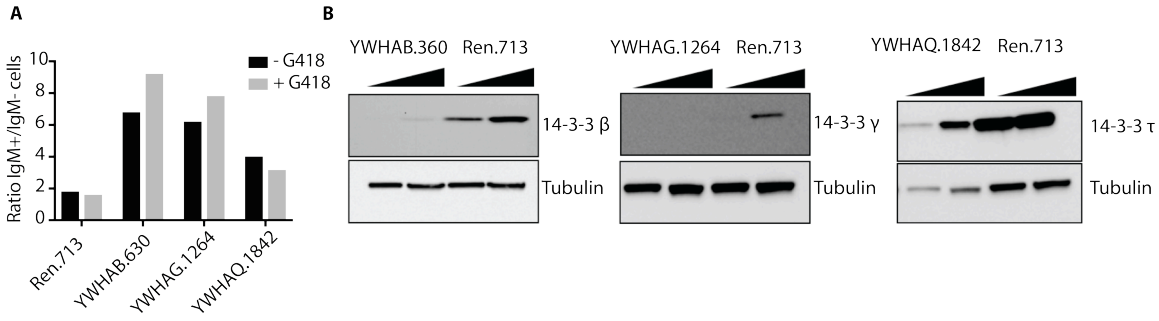


Fig. 22: Validation of the screening results. (A) FACS measurement of the ratio of IgM+ on IgM- cells, after infection with RT3GEN-shRNAs, in presence or absence of selection with G418. (B) Whole cell lysates are immunoblotted using antibodies for 14-3-3 β, 14-3-3 γ and 14-3-3 τ. Triangles indicate crescent amount of protein loaded on the gel (10 µg and 30 µg). Ren.713 infected samples are the controls. Immunoblot for Tubulin is used as loading control.

3.3.1 MutPE-seq reveals defective SHM levels in cells depleted for the hits

As for the JP8Bdel-AID line, Ramos infected with 14-3-3 shRNAs underwent genomic DNA purification after 4 days of 4-HT treatment and mutational analysis by sequencing (MutPE-seq). We included three amplicons in total, spanning the VDJ region. The mutational profiling shows defective SHM throughout the whole VDJ region when 14-3-3 β, 14-3-3 γ and 14-3-3 τ proteins are depleted, compared to the Renilla control. There is no particular bias for the complementary determining regions, CDRs (Fig. 23 A). The overall mutational rate, calculated as the number of mutations normalized on the total numbers of sequenced nucleotides, reveals a 1.7, 1.8 and 1.5-fold drop in 14-3-3β, 14-3-3γ and 14-3-3τ-depleted cells, respectively. Very interestingly defective SHM accounts

only for mutations generated by AID (C>T and G>A), the remaining mutation spectrum does not show any changes after the depletion of the hits (Fig. 23 B).

MutPE-seq has been also performed on CTCF and LDB1-depleted cells. The mutation rates per each position decrease overall after depletion of both CTCF and LDB1. As for 14-3-3 proteins, the defect in mutations does not have a bias for CDRs (Fig. S7 A). The C>T and G>A mutation frequencies are reduced in both samples of almost 2-fold, compare to the control (Fig. S7 B).

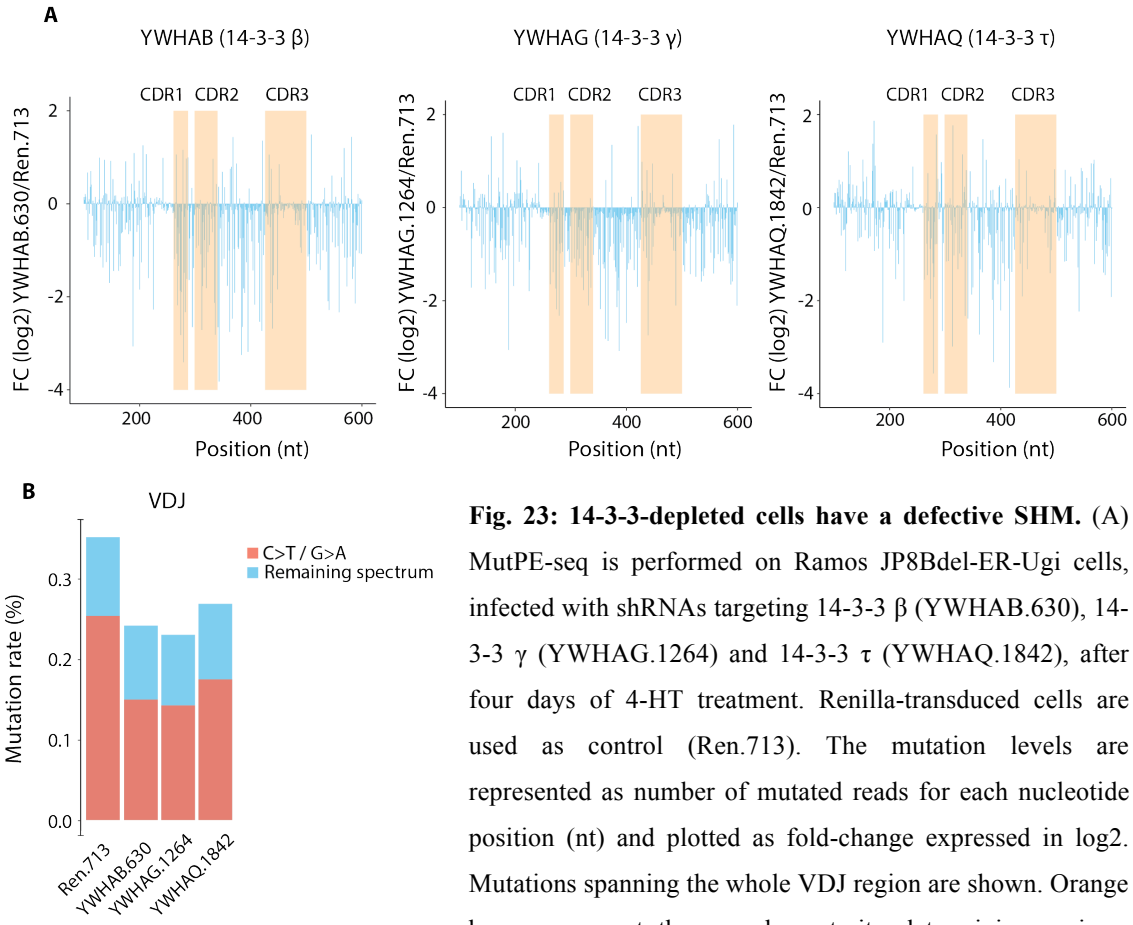


Fig. 23: 14-3-3-depleted cells have a defective SHM. (A) MutPE-seq is performed on Ramos JP8Bdel-ER-Ugi cells, infected with shRNAs targeting 14-3-3 β (YWHAB.630), 14-3-3 γ (YWHAG.1264) and 14-3-3 τ (YWHAQ.1842), after four days of 4-HT treatment. Renilla-transduced cells are used as control (Ren.713). The mutation levels are represented as number of mutated reads for each nucleotide position (nt) and plotted as fold-change expressed in log2. Mutations spanning the whole VDJ region are shown. Orange boxes represent the complementarity determining regions, CDR1, CDR2 and CDR3. (B) Mutation rate in the VDJ region plotted as percentage of number of mutations, normalized on the number of sequenced bases.

3.3.2 Knock-down of 14-3-3 does not impair transcription at the IgH locus

Very important was to rule out that transcription at the IgH locus is not impaired after knock-down of the 14-3-3 subunits. Therefore RT-qPCR is performed using primers amplifying transcripts at the promoter region (Fig. 24 left) and at the Jh6 region (Fig. 24 right). Results show that besides depletion of 14-3-3 adaptor proteins, transcription levels at the IgH locus stay unchanged. Therefore we can conclude that the defect in IgM loss is resulting from an intrinsic defect in somatic hypermutation (Fig. 24).

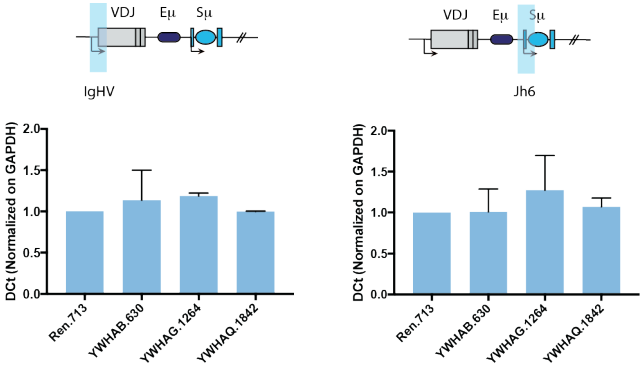


Fig. 24: 14-3-3 depletion does not impair transcription at the IgH locus. RT-qPCR analysis of samples transduced with Ren.713, YWHAB.630, YWHAG.1264, YWHAQ.1842. Values corresponding to Ren.713 are set to 1. Results shown come from two independent experiments and error bars indicate SD.

3.3.3 Preliminary co-immunoprecipitations show Spt6 as 14-3-3 binding partner

In order to investigate on putative 14-3-3 binding partners, we put effort in optimizing 14-3-3 immunoprecipitation (IP) assays. An antibody targeting 14-3-3 β is used to perform IP from total cell lysates and it was tested in three different amount (2 µg, 1 µg and 0.5 µg). All three samples show a very efficient 14-3-3 β pull-down (Fig. 25 A). In this particular experiment phosphate inhibitors are not used, although 14-3-3 proteins are known to bind preferentially phosphorylated ligands (Fu et al., 2000). When immunoprecipitated samples are subjected to Mass Spectrometry analysis, we were able to detect all 14-3-3 isoforms, except 14-3-3 σ. Very interestingly known 14-3-3 ligands, like M-Phase inducer phosphatase 2, CDC25 (Conklin et al., 1995; Dalal et al., 1999; Peng et al., 1998) and Histone deacetylase 7, HDAC7 (Kanai et al., 2000) are detected as well. Although not scoring very high and therefore needing further confirmation, the transcription elongation factor SPT6 is part of the detected binding partners as well (Table 2).

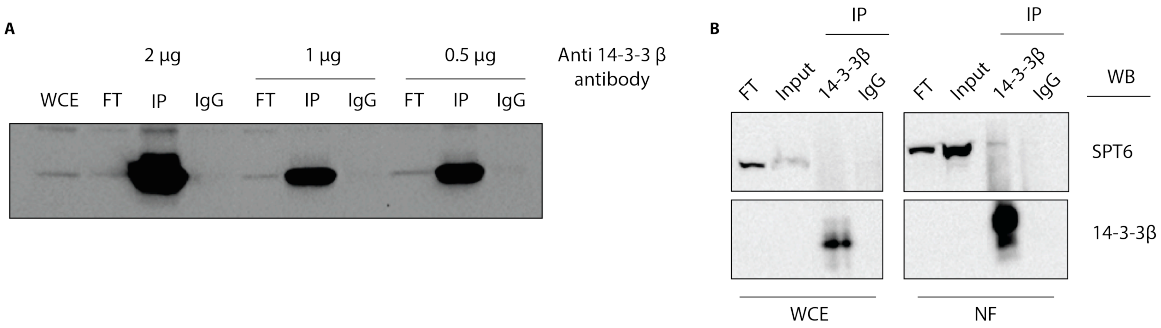


Fig. 25: Interaction partners of 14-3-3 β. (A) Immunoprecipitation assay of endogenous 14-3-3 β. Whole cell lysates from Ramos RIEP cells, were subjected to immunoprecipitation with sepharose beads coupled to 14-3-3 β antibody (IP) or to IgG only (IgG). Immunoprecipitates were analyzed by immunobotting with anti-14-3-3 β and anti-SPT6 antibodies. A representative experiment out of three is shown. (B) Co-immunoprecipitation assay of endogenous 14-3-3 β and SPT6. Lysates from Ramos RIEP cells, either whole cell extracts (WCE) and nuclear fractions (NF), were subjected to immunoprecipitation with 14-3-3 β antibody (or IgG, as control). Immunoprecipitates were analyzed by immunobotting with anti-14-3-3 β and anti-SPT6 antibodies.

Table 2: Immunoprecipitated samples are subjected to Mass Spectrometry analysis. The table lists the top-scoring proteins, quantified as Normalized Area data (Norm. Area) and shown as fold-change on the IgG control. 14-3-3 protein beta/alpha in bold is the bait protein.

Description	Fold-change IP/IgG (Norm. Area)
14-3-3 protein β/α	387084.28
14-3-3 protein η	79955.77
14-3-3 protein θ	75937.75
14-3-3 protein γ	64393.87
14-3-3 protein ζ/δ	30810.62
14-3-3 protein ε	25169.02
Branched-chain-amino-acid aminotransferase, cytosolic	885.65
E3 ubiquitin-protein ligase TRIM21	699.76
RAS protein activator like-3	354.78
Regulator of microtubule dynamics protein 3	259.61
M-phase inducer phosphatase 2	136.43
Actin-binding LIM protein 1	89.52
Histone deacetylase 7	75.22
Collagen alpha-1(I) chain	70.48
Putative RNA polymerase II subunit B1 CTD phosphatase RPAP2	40.35
GTP-binding nuclear protein Ran	3.12
Transcription elongation factor SPT6	2.60

14-3-3 adaptors are mainly located in the cytoplasm (Garcia-Guzman et al., 1999; Freed et al., 1994; reviewed in Fu et al., 2000), nonetheless hypermutation is an event occurring

in the nucleus. We therefore optimized IP conditions, including inhibitors of phosphatases, purifying only nuclear fractions and testing them in parallel with whole cell extracts. Co-IP of endogenous 14-3-3 β and SPT6 shows that their interaction can occur in the nucleus only (Fig. 25 B). We are aware that this experiment needs to be repeated, however these preliminary findings give us a hint that the transcription elongation factor SPT6 is one of the 14-3-3 binding partners, in the context of SHM.

5. DISCUSSION

Studies on somatic hypermutation were so far limited due to the absence of a suitable approach to investigate its underlying mechanism. *Ex vivo* stimulated B cells, as well as B cell lines, accumulate very poor levels of SHM (Nagaoka et al., 2002; reviewed in Liu and Schatz, 2009; Maul et al., 2014). Moreover, established cell lines undergo constitutively hypermutation without the need of activation, therefore generating a robust clonal variability. A significant problem with using enzymatically hyperactive AID proteins is that, in spite of enhanced mutation rates, they do not recapitulate the physiological AID mutation signature and targeting specificity (Wang et al., 2009; Ito et al., 2004).

Here we show that we can overcome these limitations by generating Ramos cells expressing JP8Bdel, AID deleted mutant lacking the C-terminal 16 amino acids, corresponding to the nuclear export signal (NES) (Ito et al., 2004). This increases the mutation frequency of almost 6-fold, resembling the germinal center mutations rate. The mutation profiling generated by MutPE-seq shows that the JP8Bdel-AID Ramos line has an unchanged mutation spectrum and AID targeting signature, when compared to WT Ramos cells. Mutations are in fact mainly occurring in the variable and switch regions and JP8Bdel-AID contributes to enhance mutation rates massively in the variable region and in the switch region to a lesser extent. Very importantly the constant region, which in physiological conditions is not targeted by AID, remains unmutated.

The ER-system allowed us to achieve an inducible hypermutation in Ramos B cells: the line undergoes SHM only when treated with 4-HT. JP8Bdel-transduced cells do not show has been any defective cell viability. If AID off-target genes acquire more mutations in the Ramos JP8Bdel-AID line remains to be investigated.

We are nonetheless aware of the fact that being the endogenous AID still active this will generate, to certain extent, some clonal variability. In our experiments we manage to overcome this limitation at the step of the bioinformatics analysis. Mutations generated by endogenous AID can be considered as background noise and therefore subtracted; this ensures a preserved defect (Fig. S1). It’s moreover worth mentioning that, in our

experimental time-frame, Ramos cells accumulate very low levels of SHM resulting from endogenous AID. In spite of introducing a mutated Kozak sequence dampening the transcription of JP8Bdel-AID, the leakiness of the construct is not completely assessed. For this reason the optimization experiments as well as the actual screening were performed using freshly sorted IgM positive cells.

One of the advantages of using Ramos cells is that they allow a straightforward semi-quantitative measurement of SHM. As the cells undergo somatic hypermutation they lose IgM expression, because of frame-shift or structural mutations occurring in the coding region of the immunoglobulin (Sale and Neuberger, 1998). Very importantly, cells negative for IgM expression do not show any viability defect. Moreover, IgM positive-sorted cells when put back in culture in presence of 4-HT continue undergoing IgM reversion, suggesting that this is a constant and iterative phenomenon (data not shown). This allowed us to build a screening platform to investigate on potential co-factors in SHM.

All together our data reveal that we optimized a novel platform for investigating on SHM co-factors, providing a powerful tool for large-scale genetic screens, including CRISPR/Cas9 screenings.

We designed a targeted library of 96 genes, all involved in AID targeting during class switching, chromatin remodeling and transcription regulation (Fig. S4). Among the top-scoring hits, the screening revealed the involvement of the 14-3-3 adaptor proteins in regulating SHM. Among the 14-3-3 isoforms, the best scoring in the screening are 14-3-3β, 14-3-3γ and 14-3-3τ.

The 7 subunits of 14-3-3 adaptors have a highly conserved amphipathic groove binding the ligands, suggesting that different subunits could bind the same ligands. Based on the screening results we still cannot rule out if only some of the subunits play a role in SHM and not all of them. It can be that some of them might not have scored because of inefficient gene targeting of the shRNA (Fig. S8). Future experiments will focus on performing combinatorial knock-down (or knock-out) experiments, where more than one isoform is depleted at the same time. In addition designing a shRNA or sgRNA targeting

possibly all the 7 subunits, when cell viability is still preserved, will give us insights on the role of all 14-3-3 subunits in SHM.

Given the fact that the sequence of 14-3-3 ligand binding pocket is highly conserved among subunits, one would not expect any particular isoform specificity in the binding between 14-3-3 and their ligands. Therefore eventual isoform-specific functions observed might derive from 14-3-3 subunits expression levels and/or subcellular localization. That can be the case in Ramos cells undergoing SHM; available RNA-seq data show that expression levels of these genes are very variable. RPKM values of the most expressed subunit in Ramos (YWHAZ) are 4-fold higher than the least transcribed (YWHAH) (data not shown). This may suggest that only some of the subunits would act as hetero- or homo-dimers to regulate SHM, maybe due to their subcellular localization, in spite of the fact that all the seven isoforms are upregulated by CSR-inducing stimuli (Xu et al., 2010).

The 14-3-3 subunits have a preference for particular heterodimer combinations and this correlates with our results. Both 14-3-3 γ and ε in fact score in the screening and they are known to preferentially form heterodimers. The next experiments will focus on generating double knock-out Ramos lines, where we will then assess the SHM levels after disrupting heterodimer formation. Importantly, when we perform Mass Spectrometry analysis after 14-3-3 β immunoprecipitation from whole cell extracts, we are able to detect all the 14-3-3 isoforms, but 14-3-3 σ, in line with published data showing that σ preferentially forms homodimers (reviewed in Aitken at al., 2002).

Our data do not support still a clear mechanism on how 14-3-3 would favor SHM. All we can do at this stage is to speculate about a possible model.

It was shown that AID get targeted to the acceptor and donor S regions through 14-3-3, which docks onto the DNA by recognizing the combinatorial histone modification H3K9acS10ph (Li et al., 2013). It is known that 14-3-3 is able to bind *in vitro* 5'-AGCT-3' motifs (Xu et al., 2010) and sequencing profiling of the V region in Ramos cells shows only three AGCT motifs (data not shown). Moreover the H3K9acS10ph modification is very likely introduced co-transcriptionally and in Ramos cells there are no activation steps that would suddenly induce transcription. In addition, Casali and colleagues have

shown that the AID C-terminus is indispensable for its direct interaction with 14-3-3. AID C-terminus mutants (AIDΔ (190-198) and AIDΔ (180-198)) comparable to the JP8Bdel-AID variant in use, fail to interact with all 14-3-3 isoforms (Xu et al., 2010). These evidences suggest that the role of 14-3-3 proteins in SHM might be divergent from CSR.

Solving the crystal structure of some of the 14-3-3 subunits showed that 14-3-3 dimers can simultaneously bind two different ligands (Yaffe et al., 1997; Petosa et al., 1998). That can act for example by binding two sites of the same protein, as for the case of Raf-1 (Braselmann and McCormick, 1995), where this will promote conformational changes able to regulate the protein function, or by bridging together two different factors. The amphipathic groove of the 14-3-3 monomer may work as stabilizer of one ligand and mediates the binding with a second molecule, docked onto the other 14-3-3 monomer. This could support the model postulated by Di Noia and colleagues, suggesting that a licensing mechanism of AID targeting is needed (Methot et al., 2018). Throughout the elongation process of RNA Pol II, AID will be able to target the IgH locus only if factors such SPT6 are recruited. In this context, 14-3-3 might work as scaffold proteins able of bridging together SPT6 and AID, therefore favoring the docking of the latter onto V-regions.

Our preliminary co-immunoprecipitations show that 14-3-3 and SPT6 can interact when localized in the nucleus. When we look at binding partners of cytoplasmic 14-3-3 proteins via Mass Spectrometry, SPT6 scores very low and AID is not eluted together. This observation might be due to the fact that their interaction is mostly nuclear and the nature of the binding is very labile, considering that phosphorylation is likely to play a role. Therefore the postulated model considers that when the transcriptional landscape is permissive, AID targeting is coupled with the transcription elongation machinery (Methot et al., 2018). Within the promoter-proximal region, the RNA Pol II and the transcription elongation factor SPT5 favors the AID targeting onto the IgH locus CSR (Pavri et al., 2010). In order to make AID able to mutate the DNA in a processive manner, another layer of regulation is needed. AID couples with SPT6 via specific R-residues (R171, R174 and R178) (Methot et al., 2018). Previously published data show that R-mutants AID variants loose the interaction with SPT6, although still retaining the biochemical

capability to interact with it (Methot et al., 2018). This suggests that the AID-SPT6 binding is not direct. Given that Di Noia and colleagues already showed that R-mutants loose interaction with some 14-3-3 isoforms too (Methot et al., 2018), we propose that the interaction might be 14-3-3-mediated. One 14-3-3 monomer can bind (possibly via the phosphorylated Arginine residues) and stabilize AID in the dimeric groove. This allows AID binding to SPT6, held by the other 14-3-3 monomer (Fig. 26).

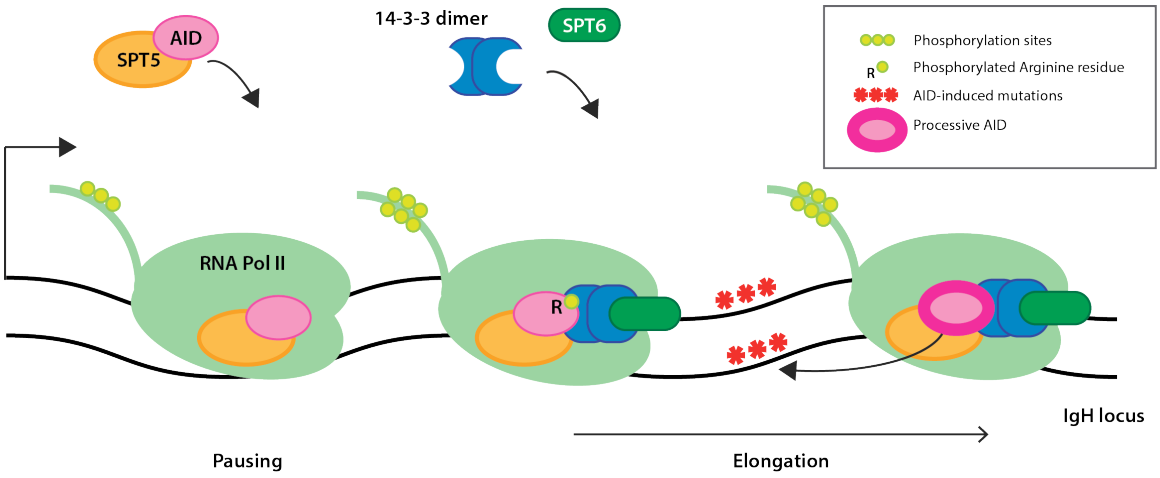


Fig. 26: Putative role of 14-3-3 proteins in AID targeting during SHM. Paused RNA Pol II co-localize with SPT5 and AID and keeps the two proteins throughout the whole elongation process. This is when the transcription elongation factor SPT6 is recruited. Nonetheless SPT6 does not interact with AID directly, but it needs the 14-3-3 adaptor dimer in order to mediate the binding. One 14-3-3 monomer contacts AID at phosphorylated Arginine residues; the other binds SPT6. Bringing together the two molecules allows AID to processively introduce mutations on ssDNA. This model is based on findings published by Javier Di Noia laboratory (Methot et al., 2018).

Future experiments will focus on confirming the 14-3-3-SPT6 interaction. The next step will be to investigate if the AID-SPT6 binding is lost in the absence of 14-3-3. AID ChIP assays will aim to show if 14-3-3 depletion leads to a defect in AID targeting and/or AID processivity. The model we postulate considers that AID is still able to target and mutate the V-region in the absence of 14-3-3, but AID coupling with SPT6 will increase its processivity. If SPT6 and 14-3-3 proteins co-immunoprecipitate with RNA Pol II will also be object of further investigation. It is crucial to take into account that the Ramos JP8Bdel line might not be the ideal system for studying AID regulation steps. In fact, the

massive amount of AID in the nucleus, due to the deletion of the nuclear export signal in the JP8Bdel-AID variant, can overrule various steps of AID regulation. For further characterizations we consider to use wild-type Ramos cells and *ex vivo* activated primary B cells.

In vivo studies considered to use 14-3-3γ knock-out mice (kind courtesy of Markus Otto). In spite of previously published data revealing the absence of phenotypic alteration in knock-out mice (Steinacker et al., 2005), we were not able to assess the SHM phenotype in homozygous knock-out mice since they were dying after birth. As it has been previously shown that the protein levels in heterozygous mice are depleted by 78 % (Steinacker et al., 2005), we considered using heterozygous mice. In parallel we will perform an adoptive transfer experiment where 24 hours anti CD40-activated B cells from B1-8hi Rosa26-Cas9 mice are retroviral infected with small guide RNAs (sgRNAs) targeting 14-3-3. Transduced cells will be adoptively transferred back into Rosa26-Cas9 mice and the analysis of SHM in germinal center B cells will follow 4-Hydroxy-3-nitrophenylacetyl-Chicken Gamma Globulin (NP-CGG) immunization (Fig. 27). Adoptive transfer conditions for B cells have been already optimized in our laboratory using sgRNAs targeting AID and showing a CSR defect (data not shown).

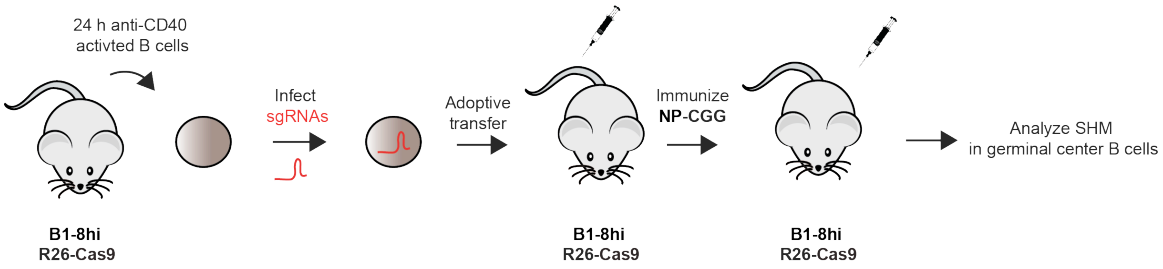


Fig. 27: Experiment outline of the adoptive transfer. B cells derived from B1-8hi Rosa26-Cas9 mice are *in vitro* activated for 24 hours with anti-CD40 and then transduced with sgRNAs. After adoptive transfer into recipient Rosa26-Cas9 mice and immunization with NP-CGG, germinal center B cells can be sorted out and evaluated for SHM.

Notably, 14-3-3 proteins were not the only factors unveiled by the screening. The two factors Lim domain-binding protein 1 (LDB1) and CCCTC-binding factor (CTCF) are

scoring among the best hits (Fig. S5, S6 and S7). In the past decades many CTCF binding elements (CBEs) have been identified throughout the IgH locus. Most of them are in the variable region (Degner et al., 2009) and in the intergenic region between V and D segments (Degner et al., 2009; Featherstone et al., 2010; Guo et al., 2011). The 3' regulatory enhancer (3'RR) consists in a cluster of DNAse I hypersensitive sites. The 10 kb at the 3'-end of the 3' regulatory enhancer (3'RR) contain around 10 CBEs and constitute the 3' boundary element of the IgH locus (Garrett et al., 2005).

When B cells undergo CSR, the 3'RR enhancer physically interacts with the Eμ enhancer and the resulting synapsis brings two recombining S-regions into proximity, promoting the efficient recombination of the AID-induced DNA breaks (Wuerffel et al., 2007). Previously published studies show that the IgH locus is intrinsically able to undergo long distance looping (Wuerffel et al., 2007) and that factors affecting chromatin structure promote this configuration (Thomas-Claudepierre et al., 2013).

CTCF is known to be an important factor for late B cell development (Pérez-García and Marina-Zárate, at al., 2017) and conditional knock-out mice have compromised B cell viability and enhanced CSR (Marina-Zárate et al., 2017). Lee and colleagues show that CTCF interacts with LDB1 and that promoters of genes expressed during erytroid development loop to LDB1- occupied enhancers via CTCF (Lee et al., 2017). We therefore speculated that the IgH locus looping in B cells could regulate somatic hypermutation via a similar mechanism. Future studies will involve chromatin immunoprecipitations (ChIP) to investigate where CTCF and LDB1 co-localize on the IgH locus, together with Chromosome Conformation Capture experiments (3-C) to assess the loop formation in presence or absence of one of the above mentioned factors.

All together our data reveal that we can overcome traditional limitations of studying somatic hypermutation by engineering Ramos cells with JP8Bdel-AID. We optimized a novel platform for investigating on SHM co-factors, providing a powerful tool for large-scale genetic screens, including CRISPR/Cas9 screenings. The targeted RNAi screen we performed gave insights on potential SHM players. 14-3-3 adaptor proteins play a putative role in SHM and future studies will focus on characterizing the underlying

molecular mechanism. Very interestingly, the enhanced and inducible SHM model we built, combined with the CRISPR/Cas9 technology for replacing the immunoglobulin V-regions with variable regions of interest, can find its application as a platform for *in vitro* generation of specific antibodies.

2. MATERIALS AND METHODS

2.1 Gibson and restriction digest-based cloning

For Gibson assembly inserts with about 25 nucleotides (nt) overhangs (o/h) on both sides that are homologous to the immediately adjacent overhangs are generated. Inserts are amplified by PCR using primers that encode the 25 nt overhang followed by the gene specific sequence (Primers are listed in Table S1). Q5 High-Fidelity DNA Polymerase (cat. no. M0491, NEB) is used for all the PCR reactions. Master mix is supplemented with Q5 High GC Enhancer. Reaction mix and termocycling conditions are performed according to manufacturer's instructions. All PCR reactions are performed using the Mastercycler pro S (cat. no. 6325000013, Eppendorf). Gibson Assembly is performed using an in-house Assembly master mix (15 µl) and 100 ng of the linearized backbone vector. Incubation at 50˚C for one hour is performed on a thermoblock (Eppendorf). 5 µl of the assembly reaction is transformed in 100 µl of in-house made competent DH5α cells. All generated plasmids are listed in Table S2.

Restriction digest-based cloning is used to generate the following constructs: pMX mk JP8Bdel-ER-Ugi and pMX AID-ER-Ugi. pMX mCherry (or GFP) backbone is linearized using BamHI (10 U/µl) (cat. no. ER0051, Thermo Fisher Scientific) and XhoI (10 U/µl) (cat. no. ER0691, Thermo Fisher Scientific) and Buffer R (10X) (cat. no. BR5, Thermo Fisher Scientific). 1 U of each of the enzymes and 3 µl of Buffer R are added to the reaction mix in a final volume of 30 µl, where 3 µg of backbone vector is digested. Same reagents are used to digest the gel-purified PCR product, which is ligated for 3 hours at room temperature using T4 DNA Ligase (cat. no. M0202M, NEB).

2.2 Generation of cell lines

Retrovirus packaging is performed in Platinum- E (PlatE) cells, cultured in DMEM, high glucose, HEPES, no phenol red (cat. no. 21063029, Thermo Fisher Scientific), supplemented with 10% Fetal Bovine Serum (cat. no. 10270106, FBS South American, Gibco) and Penicillin-Streptomycin (cat. no. P0781-100ML, Sigma-Aldrich). Ramos B cells are cultured with RPMI 1640 Phenol red-free medium (cat. no. 11835-030, Thermo Fisher Scientific), supplemented with 10% Fetal Bovine Serum (cat. no. 10270106, FBS

South American, Gibco), 1 mM Sodium Pyruvate (cat. no. 11360070, Gibco), Antibiotic Antimycotic Solution (100×) (cat. no. A5955, Sigma-Aldrich), 2.5 mM L-Glutammine (cat. no. 2530-024, Gibco), 0.5 mM Hepes pH 7.3 (produced in-house), freshly supplemented with 0.1 mM 2-Mercaptoethanol (cat. no. 60-24-2, Millipore). Ramos cells are transfected with the Fugene 6 transfection reagent (cat. no. E2691, Promega) with a ratio of Fugene:DNA of 2.5:1. The reaction mix for transfecting cells seeded in a 10 cm dish consider to use 7 µg of plasmid DNA, 7 µg of helper DNA (pCL-Eco or pCL-Ampho retrovirus packaging vectors), 35 µl of Fugene 6 reagent and up to 600 µl with Opti-MEM Reduced Serum Medium (cat. no. 31985062, Gibco). For screening purpose the reaction mix was scaled down and adapted to a six well-plate format. Spin infection is performed using virus harvested 48 hours after transfection and filtered through 0.45 µm filters to eliminate residual packaging cells from the viral suspension. Polybrene (Hexadimethrine Bromide, cat. no. H9268-104, Sigma-Aldrich) is added at the final concentration of 5 µg/ml. The virus is incubated on the cells for 4 hours at 37˚C and then substituted with fresh RPMI supplemented medium, complemented with the following reagents, accordingly to the plasmid in use: Doxycycline (cat. no. D1822 Sigma), Puromycin (Puromycin Dihydrochloride, cat. no. P8833-100MG, Sigma-Aldrich), and Geneticin G-418 Sulphate (cat. no. 11811-031, Gibco).

Infection efficiency is measured 48 hours after transduction by checking cell fluorescence via flow cytometry with the machine FACS LSR Fortessa. At the same time-point the transduced bulk population is sorted in a single-cell manner, where in a 96-well plate each well will contain only one cell. Medium is replaced accordingly to the cell confluence, which is never exciding 80%. We expand cells gradually from a 96-well plate, to a 24-well plate and finally to a 6-well plate. Single clones are then tested for infection efficiency (in the case of the RIEP Ramos line) or for IgM Loss (for the mk JP8Bdel-ER-Ugi Ramos line) and the best scoring clone is expanded to make freezing stocks.

2.3 shRNA library cloning

The shRNA library is cloned as a pool, following a cloning protocol kindly provided from Johannes Zuber laboratory. A total of 384 97-mers is ordered from IDT (Integrated DNA Technologies). Complete library specifics are indicated in Table S3. The oligomers are resuspended in a final volume of 100 µl Tris EDTA (TE). A pool of hairpins is made by mixing 5 µl of each oligomer. After few PCR optimization steps to set up the amount of DNA and the number of PCR cycles needed, the actual amplification of the pool is performed using the Platinum Pfx Kit (cat. no. 11708-013, Invitrogen). Accordingly to the protocol, the pool is PCR amplified in 24 parallel reactions, since shRNAs in the pool are less than 500. The following PCR program is run on the Mastercycler pro S: 94˚C for 2 minutes, 20 cycles of 94˚C for 14 seconds, 56˚C for 30 seconds, 68˚C for 25 seconds, 68˚C for 5 minutes and on hold at 4˚C. The PCR reaction is optimized with a DNA starting material of 0.4 ng. PCR products are purified (QIAquick PCR purification Kit, cat. no. 28106, Qiagen) and pooled. We then digested them with XhoI (NEB, R0146L; 25,000 units 20,000 units/ml) and EcoRI (NEB, cat. no. R3101L; 50,000 units 20,000 units/ml) for 4 hours at 37˚C. The whole reaction is run on a gel and gel purified (QIAquick Gel Extraction Kit, cat. no. 28706, Qiagen). RT3GEN backbone vector is digested in the same way and ligated with the pool of inserts in a molar ratio of 1:2.5 (Backbone:insert) using the T4 DNA ligase. Six ligation reactions wit 500 ng of backbone vector are set up. The pool of ligated products is purified using Phenol and PhaseLock tubes (Phase Lock Gel Light 1.5 ml - 5 PRIME, cat. no. 2302800) and precipitated with 95% Ethanol and 3M Sodium Acetate (NaAc) (pH 5.2). Transformation of electrocompent cells follows. We use MegaX DH10B T1 electrocompetent cells (cat. no. C6400-03, Invitrogen) and follow manufacturer's protocol. Serial dilutions of the bacteria culture are prepared, in a way that single colonies can be easily visible on LB/Amp Agar plates. Colonies are screened with Sanger sequencing following DNA purification. In parallel, each colony is stricked on fresh LB/Amp Agar plates to make a replica plate for DNA purification after sequencing check. Both mini preps and Sanger sequencing are performed from the in-house facility. Once the right clone is found, the corresponding colony from the replica plate is inoculated for Midi preps (NucleoBond®

Xtra Midi kit, cat. no. 740410.10, Macherey-Nagel).

2.4 IgM Loss Assay (Screening)

Ramos mk JP8Bdel-ER-Ugi cells are infected with shRNAs cloned into the RT3GEN backbone vector, using spin infection protocols (specifics are illustrated in the paragraph "Generation of cell lines"). After virus incubation, medium is supplemented with Doxycyxline and G418 (for selection of Neomycin-resistant cells). 48 hours after infection the transduction efficiency is checked via flow cytometry. In the process of screening optimization, six days after infection one million cells are harvested to confirm protein depletion by western blot. In the actual screening at this time point culture medium is supplemented with 4-Hydroxytamoxifen, at the final concentration of 0,4 µg/ml (cat. no. H7904-5MG, Sigma-Aldrich). After four days of 4-HT treatment, cells are stained with anti-IgM APC-conjugated antibody and undergo flow cytometry detection (FACS LSR Fortessa).

The screening is performed in six different rounds for handling limitations. In each of the rounds the following controls are included: Ren.713, EXOSC3.413, RPA3.1401 and RRM1.1119.

2.5 PCR for library generation and MutPE-seq

Genomic DNA is purified with a standard Phenol/Chloroform precipitation. After resuspension in TE buffer, all the samples are diluted to a final concentration of 50 ng/µl. MutPE-seq consists in a mutational analysis by sequencing that measures SHM at specific genomic sites (Robbiani et al., 2015). Libraries are generated with a two-step PCR. All the amplicons are roughly 250 bp long. Primers for the first-step PCR have a gene specific sequence and one annealing to the second-step PCR primers. The second PCR is to add Illumina P5 and P7 adapters and a 6 nt barcode for multiplexing (Fig. 6). All primers sequences used in this study are listed in Table S1 and they map in the VDJ region (three amplicons), Sµ and Cµ.

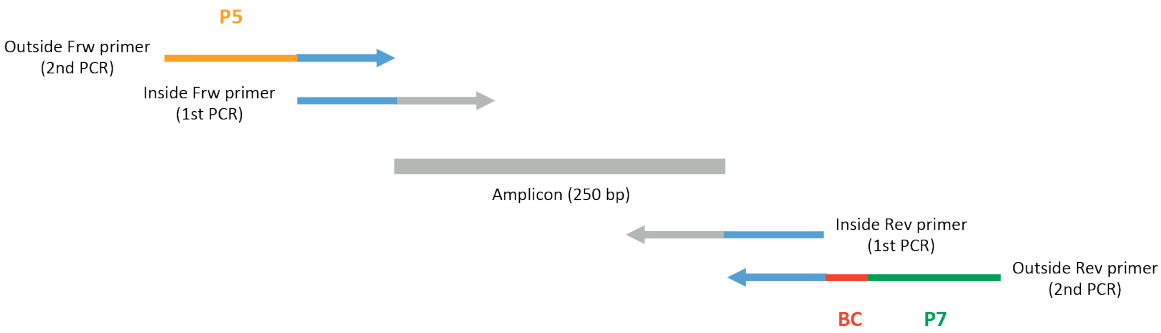


Fig. 6: Mutational analysis by Paired-End deep-sequencing (MutPE-seq). Schematic representation of the PCR strategy for library generation.

For PCR reactions we use the high-fidelity PfuUltra II Hotstart PCR Master Mix (cat. no. 600850, Agilent). The reaction mix follows manufacturer's instructions and uses 50 ng of template DNA. The following PCR program is run on the Mastercycler pro S: (First-step PCR) 95˚C for 5 minutes, 30 cycles of 95˚C for 30 seconds, 57˚C for 30 seconds, 72˚C for 30 seconds, 72˚C for 5 minutes and on hold at 4˚C; (Second-step PCR) 95˚C for 5 minutes, 10 cycles of 95˚C for 30 seconds, 57˚C for 30 seconds, 72˚C for 30 seconds, 72˚C for 5 minutes and on hold at 4˚C. Gel-purified amplicons are pooled in an equimolar manner and sequenced from both ends with HiSeq250 (Next Generation Sequencing (NGS) Core Facility, VBCF). Mutations present in only one of the paired reads are considered a sequencing artefact and therefore discarded.

2.6 Whole cell extract preparation and immunoblot analysis

Extraction of protein for Immnoblot analysis is performed using RIPA Buffer (50 mM Tris-Cl pH8, 150 mM NaCl, 0.5 mM EDTA, 0.5% Sodium Deoxycholate, 0.1% SDS, 1% Nonidet P-40 (NP-40), H20), with a starting material of at least one million cells.

After adding ice-cold RIPA buffer, lysates are incubated for 30 minutes on ice and then centrifuged at 14000 rpm for 10 minutes. Supernatant is collected and protein content is measured using the Bradford method (Bio-Rad Protein Assay Dye Reagent Concentrate, cat. no. 5000006, Bio-rad). The proteins are denatured in 6X Laemmli sample buffer (375 mM Tris-HCl, 9% SDS, 50% glycerol and 0.03% bromophenol blue, freshly supplemented with 1 mM DTT and 1X PIC-Protease Inhibitor cocktail), boiled, separated by SDS-PAGE and analyzed by immunoblot analysis.

2.7 Nuclear and whole cell extract preparation for Immunoprecipitation (IP)

For both whole cell and nuclear extracts, starting material consists in at least 100 million cells per sample. Cells are washed once with ice-cold 1X PBS buffer (Phosphate Buffered Saline), supernatant is removed entirely and cell pellets are flash-frozen in liquid Nitrogen. After thawing pellets on a heat-block at 37˚C, cells are resuspended in Low Salt (LS) Buffer (20 mM Hepes, 10 mM KCl, 1 mM MgCl2, 10% Glycerol, 1% IGEPAL) supplemented with Benzonase, 1 mM DTT, 0.5 mM PMSF and 1X PIC. Sonication is performed using a Bioruptor machine (Diagenode), with specifics of 30 sec on/off cycle, on High settings, for 5 minutes. After nutation for 30 minutes at 4˚C and spinning at full speed for 10 minutes, supernatants are transferred to fresh tubes, supplemented with High Salt (HS) Buffer (20 mM Hepes, 10 mM KCl, 1 mM MgCl2, 10% Glycerol, 1% IGEPAL, 400 mM NaCl) and undergo nutation and spinning as indicated above. Supernatants are collected and protein content is measured using the Bradford method.

For nuclear fractions cell pellets are resuspended in Sucrose Buffer IA (0.16 M sucrose, 3 mM $CaCl_2$, 2 mM Mg Acetate, 0.1 mM EDTA, 10 mM Tris.HCl pH 8.0, 0.5% NP-40, freshly supplemented with 1 mM DTT) and incubate on ice for 3 minutes. After spinning (700 x g for 5 minutes), we remove supernatant and resuspend nuclei in Nuclei Resuspension Buffer (NRB) (50 mM Tris.HCl pH 8.0, 40% Glycerol, 5 mM MgCl2, 0.1 mM EDTA, freshly supplemented with 1 mM DTT). After counting and spinning, nuclei are resuspended in Extraction buffer (EBT) (20 mM Tris-HCl pH 7.4, 10 mM KCl, 1 mM MgCl2, 10% Glycerol, 0.1% Triton-X 100, freshly supplemented with Benzonase, 1X PIC, 1 mM DTT, 5 mM Na-Orthovanadate and 25 mM NaF). Sonication is performed using a Bioruptor machine, with specifics of 30 sec on/off cycle, on High settings, for 5 minutes. After spinning at full speed for 10 minutes, supernatants are collected and transferred to a fresh tube.

Immunoprecipitations (IP) are performed using either Protein G Sepharose 4 Fast Flow (cat. no. GE17-0618-01, Sigma) or Dynabeads Protein G for Immunoprecipitation (cat. no. 10004D, Invitrogen). Whole cell or nuclear extract are pre-cleared with the beads, for

1 hour at 4˚C. Pre-cleared material is transferred to a fresh tube with 2 μg of specific antibody, over-night at 4˚C. A control run with an isotype-matched antibody is always performed in parallel (Mouse IgG1 Isotype control, cat. no. M5284, Sigma-Aldrich). Elution is performed by adding Whole Cell Extract (WCE) Buffer (LS:HS Buffers in a 5:3 ratio, supplemented with 0.5 mM PMSF) and 6x Laemmli buffer (freshly supplemented with PIC and DTT) to the beads and boiling them for 10 minutes. Immunoprecipitated samples are separated by SDS-PAGE and analyzed by immunoblot analysis.

2.8 Antibodies

For immunoblot the following antibodies are used: 14-3-3 β Antibody (A-6) (cat. no. sc-25276, Santa Cruz Biotech), 14-3-3 γ Antibody (6A1) (cat. no. sc-69955, Santa Cruz Biotech), 14-3-3 θ Antibody (3B9) (cat. no. sc-59414, Santa Cruz Biotech), β-Tubulin Antibody (cat. no. T5201, Sigma), EXOSC3 Antibody (FL-275) (cat. no. sc-98776, Santa Cruz Biotech), EXOSC10 Antibody (B-8) (cat. no. sc-374595, Santa Cruz Biotech), SPT5 Antibody (D-3) (cat. no. sc-133217, Santa Cruz Biotech), SUPT6H Antibody (cat. no. A300-801A, Bethyl Laboratories Inc.), Anti-Rabbit FC (HRP) Antibody (cat. no. 111-035-008, Dianova), Anti-Rabbit LC (HRP) Antibody (cat. no. 211-032-171, Dianova), Goat $F(ab)_2$ Anti-Mouse IgM Antibody (cat. no. SB-1022-14, Southern Biotech). For flow cytometry the following antibody is used: APC anti-human IgM Antibody (MHM-88, cat. no. B314510, Biolegend).

2.9 RT-qPCR

Total RNA is isolated from Ramos B cells by using standard Trizol purification (cat. no. 15596026, Thermo Fisher scientific). Genomic DNA is eliminated by digesting the RNA with DNase I (RNase-free) (cat. no. M0303S, NEB), following manufacturer's protocol. Reverse transcription is performed by using random primers (cat. no. 11034731001, Roche) and SuperScript II Reverse transcriptase (cat. no. 18064014, Invitrogen). IgH nascent transcripts are analyzed by quantitative PCR amplification with primers located within the promoter and the JH6 intron (Fig. 23). Data are normalized to those obtained from GAPDH transcripts.

2.10 Data analysis

Statistical significance is determined accordingly to unpaired t-test (calculated with the software Graph Prism).

2.11 MutPE-seq Analysis

Raw reads were aligned to the respective amplicon sequences with bowtie2 v2.2.1 (Langmead et al., 2012). For the JP8Bdel-AID line characterization, read-pairs were merged using bamUtil 1.0.13 (Jun et al., 2015). Positional base-pileups were computed using samtools mpileup (Li et al., 2009; Li at al., 2011) retaining anomalous read pairs, disabling per-base alignment quality and allowing a per-file depth of 100,000,000. Finally, all mutations exceeding 30% of the overall signal at a given position are filtered before plotting.

6. SUPPLEMENTARY INFORMATION

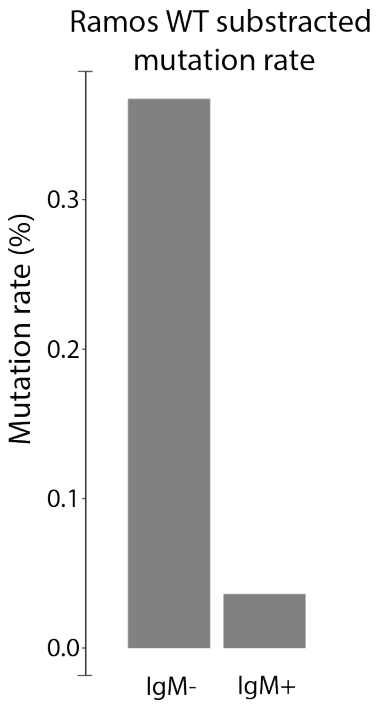
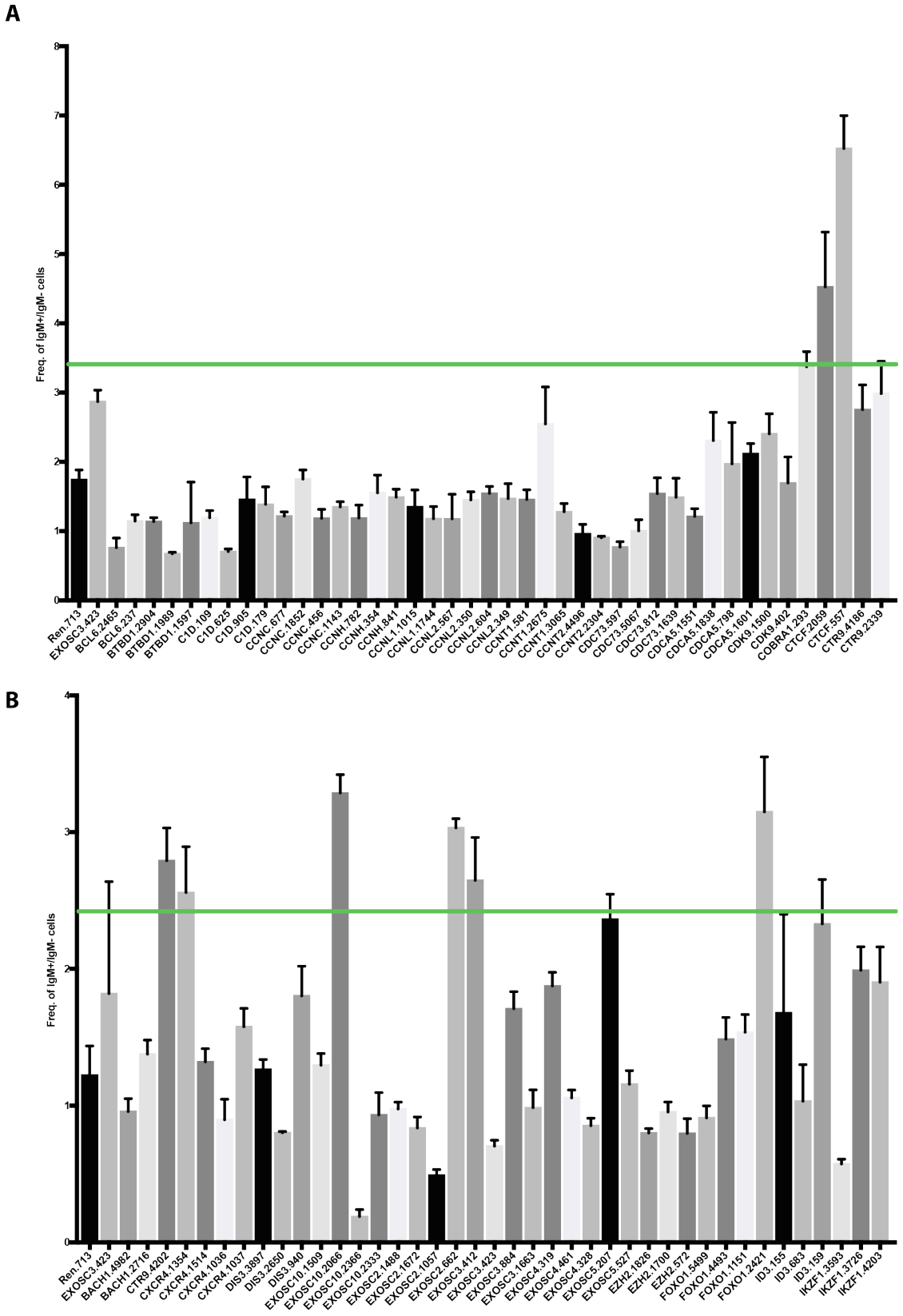
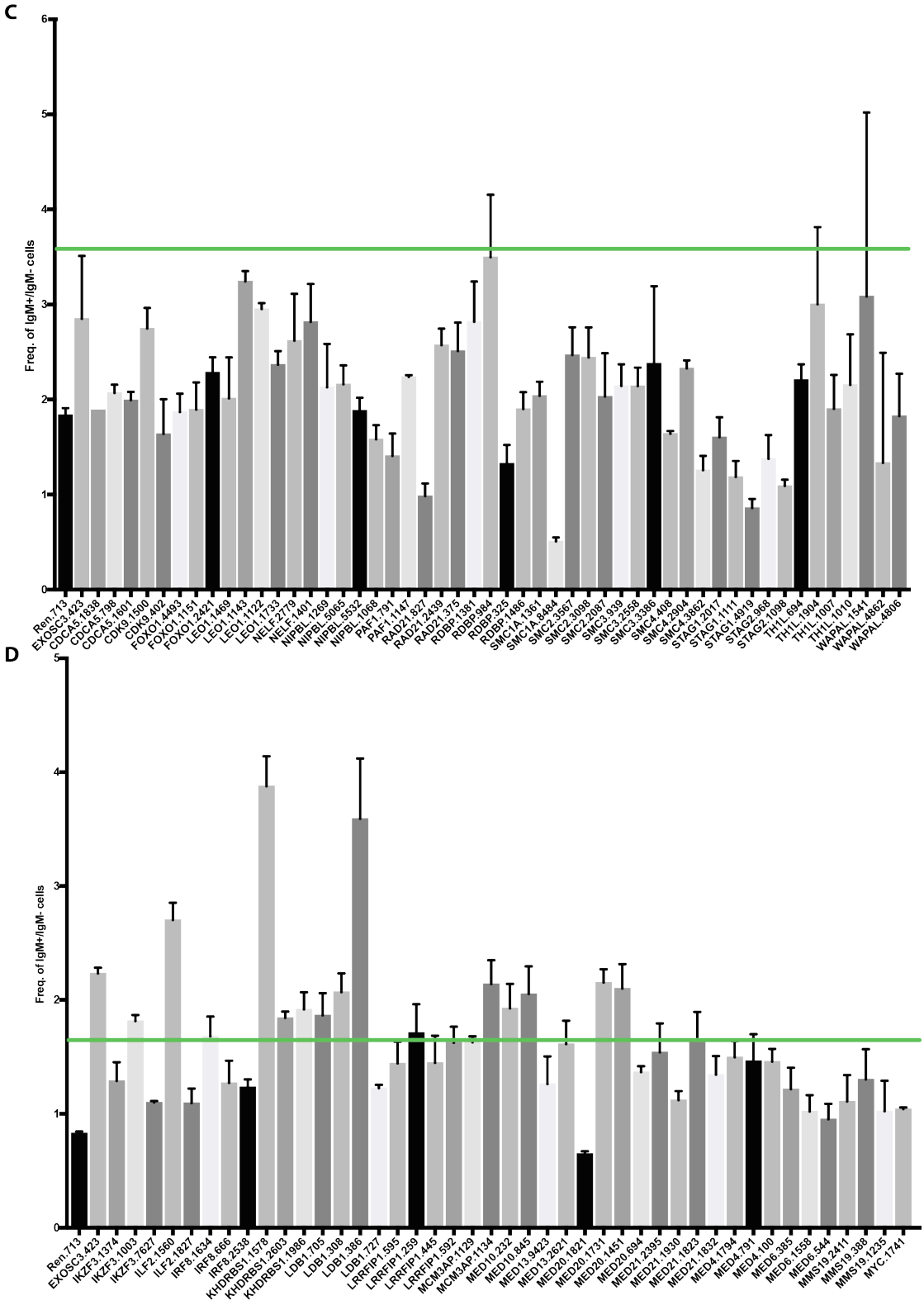


Fig. S1: Mutation rates in the VDJ region plotted as percentages of number of mutations, normalized on the total number of bases. IgM- sorted cells are compared to IgM+ cells. In both samples the background mutation noise is subtracted by normalizing on the mutation rate of WT Ramos cells.


A
Freq. of IgM+/IgM- cells
Ren.713
EXOSC3.423
BCL6.2465
BCL6.237
BTBD1.2904
BTBD1.1989
BTBD1.1597
C1D.109
C1D.625
C1D.905
C1D.179
CCNC.677
CCNC.1852
CCNC.456
CCNC.1143
CCNH.762
CCNH.354
CCNH.841
CCNL1.1015
CCNL1.1744
CCNL2.567
CCNL2.350
CCNL2.604
CCNL2.349
CCNT1.581
CCNT1.2675
CCNT1.3065
CCNT2.4496
CCNT2.2304
CDC73.597
CDC73.5067
CDC73.812
CDC73.1639
CDCA5.1551
CDCA5.1838
CDCA5.798
CDCA5.1601
CDK9.1500
CDK9.402
COBRA1.293
CTCF.2059
CTCF.557
CTR9.4186
CTR9.2339
B
Freq. of IgM+/IgM- cells
Ren.713
EXOSC3.423
BACH1.4982
BACH1.2716
CTR9.4202
CXCR4.1354
CXCR4.1514
CXCR4.1036
CXCR4.1037
DIS3.3897
DIS3.2650
DIS3.940
EXOSC10.1509
EXOSC10.2066
EXOSC10.2366
EXOSC10.2333
EXOSC2.1488
EXOSC2.1672
EXOSC2.1057
EXOSC2.662
EXOSC3.412
EXOSC3.423
EXOSC3.884
EXOSC3.1663
EXOSC4.319
EXOSC4.461
EXOSC4.328
EXOSC5.207
EXOSC5.527
EZH2.1826
EZH2.1700
EZH2.572
FOXO1.5499
FOXO1.4493
FOXO1.1151
FOXO1.2421
ID3.155
ID3.663
ID3.159
IKZF1.3593
IKZF1.3726
IKZF1.4203


C
Freq. of IgM+/IgM- cells
Ren.713
EXOSC3.423
CDCA5.1838
CDCA5.798
CDCA5.1601
CDK9.1500
CDK9.402
FOXO1.4493
FOXO1.1151
FOXO1.2421
LEO1.1469
LEO1.1143
LEO1.1122
LEO1.1733
NELF.2719
NELF.1401
NIPBL.1269
NIPBL.5065
NIPBL.5532
NIPBL.1068
PAF1.791
PAF1.1147
RAD21.827
RAD21.2439
RAD21.375
RDBP.1381
RDBP.984
RDBP.325
RDBP.1486
SMC1A.1361
SMC1A.8484
SMC2.3567
SMC2.3098
SMC2.2087
SMC3.939
SMC3.3558
SMC3.3386
SMC4.408
SMC4.2904
SMC4.3862
STAG1.2017
STAG1.1111
STAG1.4919
STAG2.968
STAG2.1098
THIL.694
THIL.1904
THIL.1007
THIL.1010
WAPAL.1541
WAPAL.4862
WAPAL.4806
D
Freq. of IgM+/IgM- cells
Ren.713
EXOSC3.423
IKZF3.1374
IKZF3.1003
IKZF3.7627
ILF2.1560
ILF2.1827
IRF8.1634
IRF8.666
IRF8.2538
KHDRBS1.1578
KHDRBS1.2603
KHDRBS1.1986
LDB1.705
LDB1.308
LDB1.386
LDB1.727
LRRFIP1.595
LRRFIP1.259
LRRFIP1.445
LRRFIP1.593
MCM3AP.1129
MCM3AP.1134
MED10.232
MED10.845
MED13.9423
MED13.2621
MED20.1821
MED20.1731
MED20.1451
MED20.694
MED21.2395
MED21.1930
MED21.1823
MED21.1832
MED4.1794
MED4.791
MED4.100
MED6.385
MED6.1558
MED6.544
MMS19.2411
MMS19.388
MMS19.1235
MYC.1741

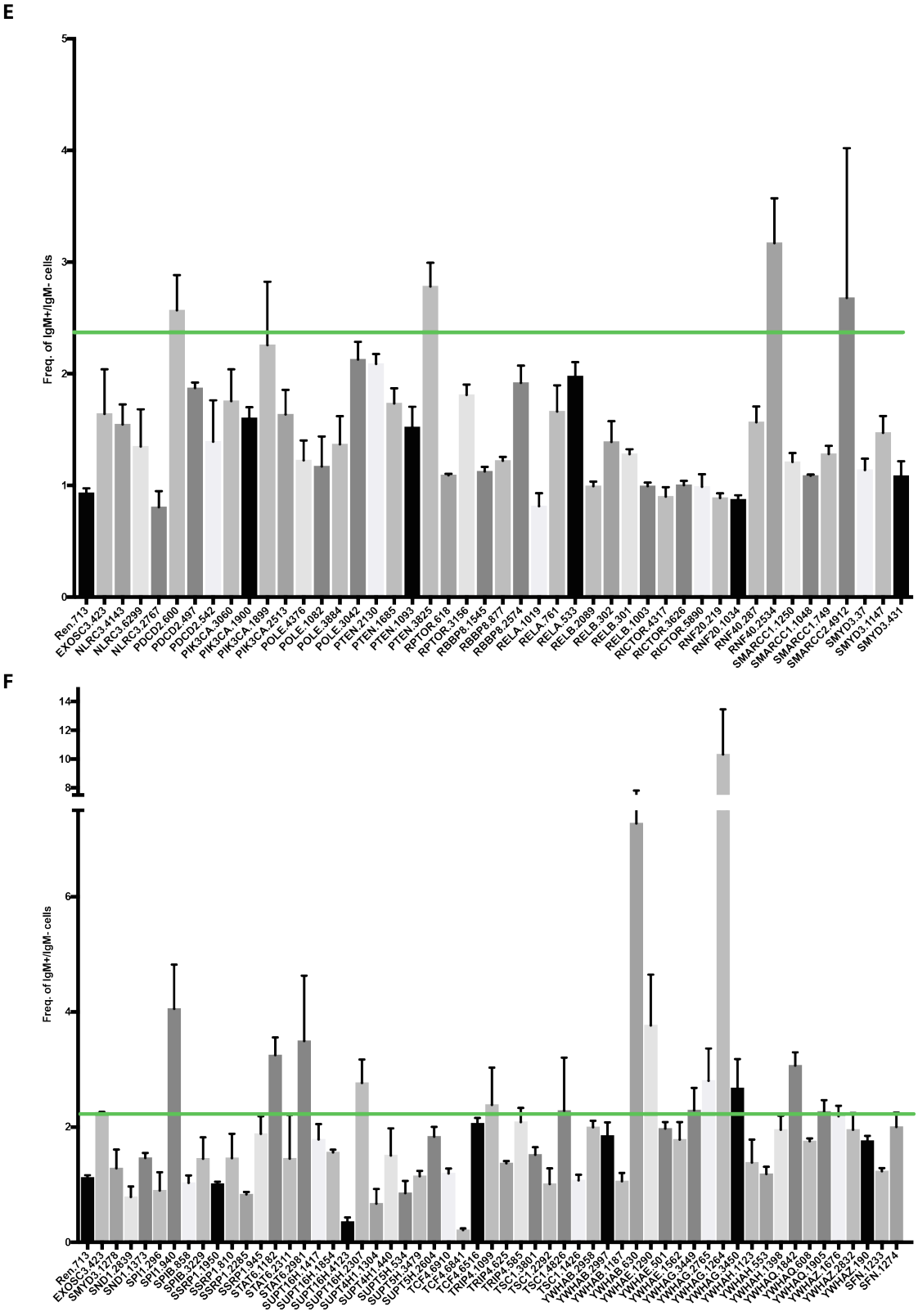


Fig. S2: Detailed bar-plots of the screening results. Each graph illustrates one single round of screening ((A) Round 1. (B) Round 2. (C) Round 3. (D) Round 4. (E) Round 5. (F) Round 6). The

ratio of IgM+ and IgM- cells frequencies is plotted. Error bars represent standard deviations (SD). The green line represents the 2-fold Renilla.713 cut-off.

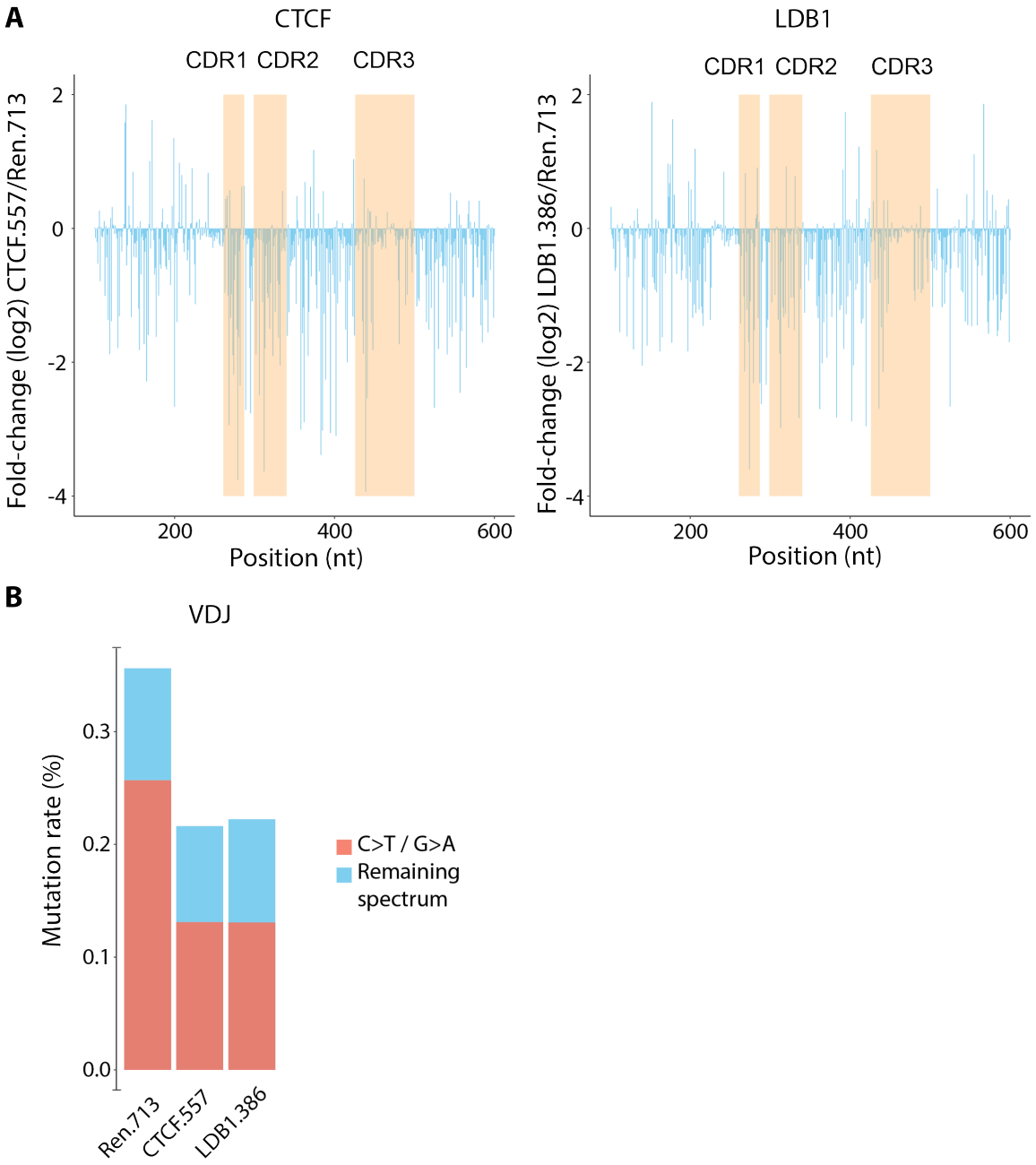


Fig. S3: CTCF and LDB1-depleted cells have a defective SHM. (A) MutPE-seq is performed on Ramos JP8Bdel-ER-Ugi cells, infected with shRNAs targeting CTCF (CTCF.557) and LDB1 (LDB1.386), after four days of 4-HT treatment. Renilla-transduced cells are used as control (Ren.713). The mutation levels are represented as number of mutated reads for each nucleotide

position (nt) and plotted as fold-change expressed in log2. Mutations spanning the whole VDJ region are shown. Orange boxes represent the complementary determining regions, CDR1, CDR2 and CDR3. (B) Mutation rate in the VDJ region plotted as percentage of number of mutations, normalized on the number of sequenced bases.

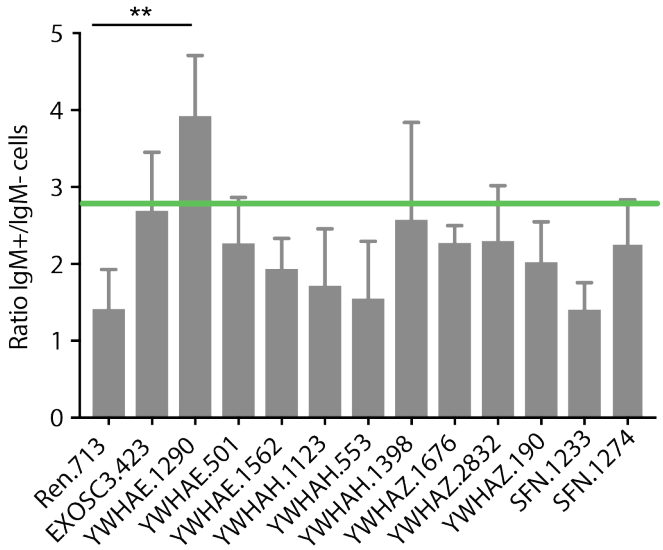


Fig. S4: IgM Loss assay after depletion of 14-3-3 ε, η, ζ and σ isoforms. The ratio of IgM+ on IgM- cells is shown. Ren.713 and EXOSC3.423 represent the negative and positive controls, respectively and Ren.713 is used as normalizers. The green line is the cut-off (two-fold Ren.713). Error bars represent standard deviations (SD) calculated from four independent experiments. (**P<0.005).

Table S1: Oligonucleotide information.

Name of Primer	Sequence	Comments
Gib_9	GACCATCCTCTAGACTGCCGGATCCACCATGGACAGCCTCTTGATGAAC	To PCR JP8Bdel with o/h for pMX mCherry and Flag-ER
Gib_10	AATCGGTACCCAGGGGCAAAAGGATGCG	
Gib_11	TTTGCCCCTGGGTACCGATTACAAGGACG	To PCR Flag-ER with o/h for JP8Bdel and 2A-Ugi
Gib_12	AGTTGGTGGCGACTGTGGCAGGGAAACC	
Gib_13	TGCCACAGTCGCCACCAACTTCAGCCTG	To PCR 2A-Ugi with o/h for Flag-ER and pMX mCherry
Gib_14	CAGCGGCCGCGTGGAATTCGGATCCTTATCACAGCATTTTGATCTTGTTTTC	
Gib_15	GACCATCCTCTAGACTGCCGGATCCACCATGGACAGCCTCTTGATG	To PCR AID with o/h for pMX mCherry and Flag-ER
Gib_16	AATCGGTACCAAGTCCCAAAGTACGAAATG	
Gib_17	TTTGGGACTTGGTACCGATTACAAGGACG	To PCR Flag-ER with o/h for AID and 2A-Ugi
Gib_18	AGTTGGTGGCGACTGTGGCAGGGAAACC	
JP8elER_mK-F	CGGGGATCCGGTTTATGGACAGCCTCTTGATGAAC	To PCR JP8Bdel-ER-Ugi with o/h with BamHI and XhoI restriction sites
JP8elERugi_mK-R	CGGCTCGAGTCACAGCATTTTGATCTTGTTTTCTCCG	
mirEXhoI5	TACAATACTCGAGAAGGTATATTGCTGTTGACAGTGAGCG	To clone shRNAs library and single hairpins for screening set up.
mirEEcoRI3	TTAGATGAATTCTAGCCCCTTGAAGTCCGAGGCAGTAGGCA	
Ramos_MutPE_1	TCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGCACAAGAACATGAAACACC	1st round MutPE-seq.To PCR the VDJ region (amplicon size 246 bp)
Ramos_MutPE_2	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTAGTAACCACTGAAGGACCCA	
Ramos_MutPE_3	TCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACTACTGGAGTTGGATCCGC	1st round MutPE-seq.To PCR the VDJ region (amplicon size 221 bp)
Ramos_MutPE_4	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCCAGGACTCGCCCTAGTAA	
Ramos_MutPE_5	TCTACACTCTTTCCCTACACGACGCTCTTCCGATCTACAGACGGGAGGTACGGT	1st round MutPE-seq.To PCR the VDJ region (amplicon size 228 bp)
Ramos_MutPE_6	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGAAAATCCACAGAGGCTCC	

Name of Primer	Sequence	Comments
Ramos_MutPE_7	TCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTCACAGCACAGGCTCCTAAAT	1st round MutPE-seq.To PCR the Sμ region (amplicon size 237 bp)
Ramos_MutPE_8	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGCTAAAGCCATCTCATTGCCG	
Ramos_MutPE_9	TCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCCTGTGAGAATTCCCCGTCG	1st round MutPE-seq.To PCR the Cμ region (amplicon size 225 bp)
Ramos_MutPE_10	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTACACCACGTGTTCGTCTGTG	
MutPE_common1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAC	2nd round MutPE-seq.To Frw primer w/o Barcode
MutPE_common2 (revIDX6)	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTG	2nd round MutPE-seq. Rev primer with TruSeq single index (GCCAAT)
MutPE_common3 (revIDX12)	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTG	2nd round MutPE-seq. Rev primer with TruSeq single index (CTTGTA)
MutPE_common4 (revIDX7)	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTG	2nd round MutPE-seq. Rev primer with TruSeq single index (CAGATC)
MutPE_common5 (revIDX5)	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTG	2nd round MutPE-seq. Rev primer with TruSeq single index (ACAGTG)
MutPE_common6 (revIDX2)	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTG	2nd round MutPE-seq. Rev primer with TruSeq single index (CGATGT)
IGHV (1-100)-F1	TGAAACACCTGTGGTTCTTCCT	PCR for IgH transcripts.
IGHV (1-100)-R1	GACCCCTGTGAACAGAGAAACC	
JH6pre-mRNA-F1	GAGGTACGGTATGGACGTCTG	
JH6pre-mRNA-R1	AGGACCAACCTGCAATGCTC	
YWHAB_7529_11_For	CACCGCATTGAGCAGAAAACAGAG	To clone small guides RNA targeting YWHAB gene into PX458
YWHAB_7529_11_Rev	AAACCTCTGTTTTCTGCTCAATGC	
YWHAB_7529_10_For	CACCGAAGGCAGTCACAGAACAG	To clone small guides RNA targeting YWHAB gene into

Name of Primer	Sequence	Comments
YWHAB_7529_10_Rev	AAACCTGTTCTGTGACTGCCTTC	PX458
YWHAG_7532_7_For	CACCGTACCGGGAGAAGATAGAGA	To clone small guides RNA targeting YWHAG gene into PX458
YWHAG_7532_7_Rev	AAACTCTCTATCTTCTCCCGGTAC	
YWHAG_7532_11_For	CACCGTTCTACCTGAAGATGAAAG	To clone small guides RNA targeting YWHAG gene into PX458
YWHAG_7532_11_Rev	AAACCTTTCATCTTCAGGTAGAAC	
YWHAB_genotyping	CATCAATATTGTTGATGGCA	To genotype 14-3-3 β KO line
	TGACACAAATTATATCATATGC	
YWHAG_genotyping	GTCTTGGCCAAGTGGCACGCTT	To genotype 14-3-3 γ KO line
	CAGACACCGTTGATCCATTG	
YWHAQ_genotyping	CAGACACCGTTGATCCATTG	To genotype 14-3-3 τ KO line
	CGTGTGCGGGAGACAAAAG	

Table S2: Plasmid list.

Name of plasmid	RE site used for cloning/Comments
pRRL.SFFV-rtTA3-IRES-EcoR-PGK-Puro (pRRL.RIEP)	Kindly provided from Zuber lab
pMX_JP8del-ER-2A-Ugi	BamHI
pMX_hAID-ER-2A-Ugi	BamHI
pMX AIDER	Previously generated in the lab
pMX-mk JP8delERUgi_mCherry	BamHI, XhoI
LENC_Ren.713	XhoI, EcoRI
LENC_Rrm1.1118	XhoI, EcoRI
LENC_RPA3.1401	XhoI, EcoRI
LENC_Myc.1891	XhoI, EcoRI
LENC_Myc1888	XhoI, EcoRI
RT3GEN_Ren.713	XhoI, EcoRI
RT3GEN_Rrm1.1118	XhoI, EcoRI
RT3GEN_RPA3.1401	XhoI, EcoRI
RT3GEN_Myc.1891	XhoI, EcoRI
RT3GEN_Myc1888	XhoI, EcoRI
RT3GEN_SUPT5H.534	XhoI, EcoRI
RT3GEN_SUPT5H. 3479	XhoI, EcoRI
RT3GEN_SUPT5H. 2604	XhoI, EcoRI
RT3GEN_SUPT5H. 414	XhoI, EcoRI
RT3GEN_EXOSC3.412	XhoI, EcoRI
RT3GEN_EXOSC3.423	XhoI, EcoRI
RT3GEN_EXOSC3.1663	XhoI, EcoRI
RT3GEN_EXOSC3.1385	XhoI, EcoRI
RT3GEN_EXOSC10.1509	XhoI, EcoRI
RT3GEN_EXOSC10.2066	XhoI, EcoRI
RT3GEN_EXOSC10.2366	XhoI, EcoRI
RT3GEN_EXOSC10.2333	XhoI, EcoRI
RT3GEN_EXOSC10.444	XhoI, EcoRI
pSpCas9(BB)-2A-GFP (PX458)	BbsI, Addgene plasmid #48138

Table S3: shRNA Library.

shRNA ID	Sequence
BACH1.287	TGCTGTTGACAGTGAGCGCTCCAGAAGAGGTGACAGTTAATAGTGAAGCCA CAGATGTATTAACTGTCACCTCTTCTGGAATGCCTACTGCCTCGGA
BACH1.4982	TGCTGTTGACAGTGAGCGCTCGTTAAATTGTCATATTCAATAGTGAAGCCAC AGATGTATTGAATATGACAATTTAACGAATGCCTACTGCCTCGGA
BACH1.1188	TGCTGTTGACAGTGAGCGAAGGGAAGATAGTAGTGTTGCATAGTGAAGCCA CAGATGTATGCAACACTACTATCTTCCCTGTGCCTACTGCCTCGGA
BACH1.2716	TGCTGTTGACAGTGAGCGCAGCAATGCTAATAGAGTTCTATAGTGAAGCCA CAGATGTATAGAACTCTATTAGCATTGCTTTGCCTACTGCCTCGGA
BCL6.2465	TGCTGTTGACAGTGAGCGCTGGCAGAGTTGTAAATATATATAGTGAAGCCA CAGATGTATATATATTTACAACTCTGCCATTGCCTACTGCCTCGGA
BCL6.1381	TGCTGTTGACAGTGAGCGACCGGCTCAATAACATCGTTAATAGTGAAGCCA CAGATGTATTAACGATGTTATTGAGCCGGCTGCCTACTGCCTCGGA
BCL6.2403	TGCTGTTGACAGTGAGCGCAGACTTCAGTATGTTGTCAAATAGTGAAGCCA CAGATGTATTTGACAACATACTGAAGTCTTTGCCTACTGCCTCGGA
BCL6.237	TGCTGTTGACAGTGAGCGAACCAGTTGAAATGCAACCTTATAGTGAAGCCA CAGATGTATAAGGTTGCATTTCAACTGGTCTGCCTACTGCCTCGGA
BTBD1.2904	TGCTGTTGACAGTGAGCGCACGGATATAAACCTCAGTTAATAGTGAAGCCA CAGATGTATTAACTGAGGTTTATATCCGTATGCCTACTGCCTCGGA
BTBD1.1989	TGCTGTTGACAGTGAGCGCTGGGTCAGTATTCCTACAGAATAGTGAAGCCA CAGATGTATTCTGTAGGAATACTGACCCATTGCCTACTGCCTCGGA
BTBD1.1597	TGCTGTTGACAGTGAGCGCACCATACAATCTAGTGTCAAATAGTGAAGCCA CAGATGTATTTGACACTAGATTGTATGGTATGCCTACTGCCTCGGA
BTBD1.1159	TGCTGTTGACAGTGAGCGAACCGACCAAGATGCTGTCTCATAGTGAAGCCA CAGATGTATGAGACAGCATCTTGGTCGGTCTGCCTACTGCCTCGGA
C1D.109	TGCTGTTGACAGTGAGCGCTCCAGTAGAAATTCACGAGTATAGTGAAGCCA CAGATGTATACTCGTGAATTTCTACTGGATTGCCTACTGCCTCGGA
C1D.625	TGCTGTTGACAGTGAGCGAACAGTAAATATGTAAAGCTAATAGTGAAGCCA CAGATGTATTAGCTTTACATATTTACTGTGTGCCTACTGCCTCGGA
C1D.905	TGCTGTTGACAGTGAGCGCTCACTATATGATATTAAGAAATAGTGAAGCCA CAGATGTATTTCTTAATATCATATAGTGAATGCCTACTGCCTCGGA
C1D.179	TGCTGTTGACAGTGAGCGAACCATGATGTCTGTTTCTAGATAGTGAAGCCAC AGATGTATCTAGAAACAGACATCATGGTCTGCCTACTGCCTCGGA
CCNC.677	TGCTGTTGACAGTGAGCGCAGCCTGTGTTGTACAGCAGAATAGTGAAGCCA CAGATGTATTCTGCTGTACAACACAGGCTATGCCTACTGCCTCGGA
CCNC.1852	TGCTGTTGACAGTGAGCGAAGAGTAGGACATCATACTAAATAGTGAAGCCA CAGATGTATTTAGTATGATGTCCTACTCTCTGCCTACTGCCTCGGA
CCNC.456	TGCTGTTGACAGTGAGCGCAGGATGAATCATATATTAGAATAGTGAAGCCA CAGATGTATTCTAATATATGATTCATCCTATGCCTACTGCCTCGGA
CCNC.1143	TGCTGTTGACAGTGAGCGCACAGACAGACATACATAGACATAGTGAAGCCA CAGATGTATGTCTATGTATGTCTGTCTGTATGCCTACTGCCTCGGA
CCNH.782	TGCTGTTGACAGTGAGCGATGGAATTACTATGGAAAGTTATAGTGAAGCCA CAGATGTATAACTTTCCATAGTAATTCCAGTGCCTACTGCCTCGGA
CCNH.354	TGCTGTTGACAGTGAGCGCACGGCTTGTATGTATTTCAAATAGTGAAGCCAC AGATGTATTTGAAATACATACAAGCCGTATGCCTACTGCCTCGGA
CCNH.848	TGCTGTTGACAGTGAGCGAACAGTTACTAGATATAATGAATAGTGAAGCCA CAGATGTATTCATTATATCTAGTAACTGTGTGCCTACTGCCTCGGA
CCNH.841	TGCTGTTGACAGTGAGCGCGCCTGTCACAGTTACTAGATATAGTGAAGCCA CAGATGTATATCTAGTAACTGTGACAGGCATGCCTACTGCCTCGGA
CCNL1.1015	TGCTGTTGACAGTGAGCGCTCGAAGAAGCACCTAGAAGAATAGTGAAGCCA CAGATGTATTCTTCTAGGTGCTTCTTCGATTGCCTACTGCCTCGGA
CCNL1.607	TGCTGTTGACAGTGAGCGCCCGGGCCTCATTCGACAGCTATAGTGAAGCCA CAGATGTATAGCTGTCGAATGAGGCCCGGATGCCTACTGCCTCGGA

shRNA ID	Sequence
CCNL1.1744	TGCTGTTGACAGTGAGCGAACAGCAAGAGAAGTAGAAATATAGTGAAGCC ACAGATGTATATTTCTACTTCTCTTGCTGTCTGCCTACTGCCTCGGA
CCNL1.8	TGCTGTTGACAGTGAGCGCCCACAGCCTTGTGTTCTTCAATAGTGAAGCCAC AGATGTATTGAAGAACACAAGGCTGTGGATGCCTACTGCCTCGGA
CCNL2.567	TGCTGTTGACAGTGAGCGAGCGGAAAGACGAGTTCTCAAATAGTGAAGCCA CAGATGTATTTGAGAACTCGTCTTTCCGCCTGCCTACTGCCTCGGA
CCNL2.350	TGCTGTTGACAGTGAGCGCCCAGCGGTTCTTTTATACCAATAGTGAAGCCAC AGATGTATTGGTATAAAAGAACCGCTGGATGCCTACTGCCTCGGA
CCNL2.604	TGCTGTTGACAGTGAGCGATCCATGTGAAGCATCCTCATATAGTGAAGCCA CAGATGTATATGAGGATGCTTCACATGGACTGCCTACTGCCTCGGA
CCNL2.349	TGCTGTTGACAGTGAGCGCTCCAGCGGTTCTTTTATACCATAGTGAAGCCAC AGATGTATGGTATAAAAGAACCGCTGGAATGCCTACTGCCTCGGA
CCNT1.581	TGCTGTTGACAGTGAGCGAAGCCTTGTTTCTAGCAGCTAATAGTGAAGCCAC AGATGTATTAGCTGCTAGAAACAAGGCTGTGCCTACTGCCTCGGA
CCNT1.2675	TGCTGTTGACAGTGAGCGAAGGGTTTCTTGTATCAGACAATAGTGAAGCCA CAGATGTATTGTCTGATACAAGAAACCCTCTGCCTACTGCCTCGGA
CCNT1.1336	TGCTGTTGACAGTGAGCGACGCGCGAAGCATGCAGAAGAATAGTGAAGCCA CAGATGTATTCTTCTGCATGCTTCGCGCGGTGCCTACTGCCTCGGA
CCNT1.3065	TGCTGTTGACAGTGAGCGCTCCAACAGTCTACTAGAGAAATAGTGAAGCCA CAGATGTATTTCTCTAGTAGACTGTTGGAATGCCTACTGCCTCGGA
CCNT2.5069	TGCTGTTGACAGTGAGCGCAGGGTGGTTTGTTCACTGAAATAGTGAAGCCA CAGATGTATTTCAGTGAACAAACCACCCTTTGCCTACTGCCTCGGA
CCNT2.513	TGCTGTTGACAGTGAGCGACCAGTTAGTAAGAGCAAGCAATAGTGAAGCCA CAGATGTATTGCTTGCTCTTACTAACTGGGTGCCTACTGCCTCGGA
CCNT2.4496	TGCTGTTGACAGTGAGCGAAGGATTGTGTGAGCTATTCAATAGTGAAGCCA CAGATGTATTGAATAGCTCACACAATCCTGTGCCTACTGCCTCGGA
CCNT2.2304	TGCTGTTGACAGTGAGCGATCCACATATGATAGTGTTATATAGTGAAGCCAC AGATGTATATAACACTATCATATGTGGAGTGCCTACTGCCTCGGA
CDC73.597	TGCTGTTGACAGTGAGCGCTCAGCGATCTACTCAAGTCAATAGTGAAGCCA CAGATGTATTGACTTGAGTAGATCGCTGAATGCCTACTGCCTCGGA
CDC73.5067	TGCTGTTGACAGTGAGCGAAGGAATATTTGTTGAACTAAATAGTGAAGCCA CAGATGTATTTAGTTCAACAAATATTCCTCTGCCTACTGCCTCGGA
CDC73.812	TGCTGTTGACAGTGAGCGCTGGCTAAGAAAAGATCTACTATAGTGAAGCCA CAGATGTATAGTAGATCTTTTCTTAGCCATTGCCTACTGCCTCGGA
CDC73.1639	TGCTGTTGACAGTGAGCGCCCAGTTGATATATTTGCTAAATAGTGAAGCCAC AGATGTATTTAGCAAATATATCAACTGGTTGCCTACTGCCTCGGA
CDCA5.1551	TGCTGTTGACAGTGAGCGATCAGACCATAAGTGTGTACTATAGTGAAGCCA CAGATGTATAGTACACACTTATGGTCTGAGTGCCTACTGCCTCGGA
CDCA5.1838	TGCTGTTGACAGTGAGCGAGCCCTAGAAGAGAAAGTTGAATAGTGAAGCCA CAGATGTATTCAACTTTCTCTTCTAGGGCCTGCCTACTGCCTCGGA
CDCA5.798	TGCTGTTGACAGTGAGCGAACCCGAGAAACAGAAACGTAATAGTGAAGCCA CAGATGTATTACGTTTCTGTTTCTCGGGTGTGCCTACTGCCTCGGA
CDCA5.1601	TGCTGTTGACAGTGAGCGATGCTGTTGAGATGTTTGGTAATAGTGAAGCCAC AGATGTATTACCAAACATCTCAACAGCAGTGCCTACTGCCTCGGA
CDK9.1879	TGCTGTTGACAGTGAGCGCTGGGACTTGATTGTCAAGTCATAGTGAAGCCA CAGATGTATGACTTGACAATCAAGTCCCAATGCCTACTGCCTCGGA
CDK9.1734	TGCTGTTGACAGTGAGCGATGGGCACAGTTTGGTCCGTTATAGTGAAGCCA CAGATGTATAACGGACCAAACTGTGCCCAGTGCCTACTGCCTCGGA
CDK9.1500	TGCTGTTGACAGTGAGCGACCCACCAGTGACTTTTTCTAATAGTGAAGCCAC AGATGTATTAGAAAAAGTCACTGGTGGGCTGCCTACTGCCTCGGA

shRNA ID	Sequence
CDK9.402	TGCTGTTGACAGTGAGCGCCCGCTGCAAGGGTAGTATATATAGTGAAGCCA CAGATGTATATATACTACCCTTGCAGCGGTTGCCTACTGCCTCGGA
COBRA1.641	TGCTGTTGACAGTGAGCGCGCGGCAGATCTGGCAAGACAATAGTGAAGCCA CAGATGTATTGTCTTGCCAGATCTGCCGCTTGCCTACTGCCTCGGA
COBRA1.293	TGCTGTTGACAGTGAGCGCCGAGCAGTTCCAGACAGAGAATAGTGAAGCCA CAGATGTATTCTCTGTCTGGAACTGCTCGATGCCTACTGCCTCGGA
COBRA1.1345	TGCTGTTGACAGTGAGCGATGGTGGATGACTACACTTTCATAGTGAAGCCA CAGATGTATGAAAGTGTAGTCATCCACCAGTGCCTACTGCCTCGGA
COBRA1.1805	TGCTGTTGACAGTGAGCGAAGAGGCAGTGAAGGAGCTTTATAGTGAAGCCA CAGATGTATAAAGCTCCTTCACTGCCTCTCTGCCTACTGCCTCGGA
CTCF.2059	TGCTGTTGACAGTGAGCGCTGGGAAGGACTTAGAGTTTTATAGTGAAGCCA CAGATGTATAAAACTCTAAGTCCTTCCCAATGCCTACTGCCTCGGA
CTCF.557	TGCTGTTGACAGTGAGCGCCGCCAGTGTAGAAGTCAGCAATAGTGAAGCCA CAGATGTATTGCTGACTTCTACACTGGCGTTGCCTACTGCCTCGGA
CTCF.1905	TGCTGTTGACAGTGAGCGATGAGGACAGTGTTGACAACTATAGTGAAGCCA CAGATGTATAGTTGTCAACACTGTCCTCAGTGCCTACTGCCTCGGA
CTCF.620	TGCTGTTGACAGTGAGCGCTCAGTGCAGTTTGTGCAGTTATAGTGAAGCCAC AGATGTATAACTGCACAAACTGCACTGAATGCCTACTGCCTCGGA
CTR9.2005	TGCTGTTGACAGTGAGCGCCCCGAGATCGAGAAAAGGAAATAGTGAAGCCA CAGATGTATTTCCTTTTCTCGATCTCGGGTTGCCTACTGCCTCGGA
CTR9.4186	TGCTGTTGACAGTGAGCGAAGCGATGTAGATAAAATCACATAGTGAAGCCA CAGATGTATGTGATTTTATCTACATCGCTGTGCCTACTGCCTCGGA
CTR9.2339	TGCTGTTGACAGTGAGCGATGGCAAGTTACAGGAATGCAATAGTGAAGCCA CAGATGTATTGCATTCCTGTAACTTGCCACTGCCTACTGCCTCGGA
CTR9.4202	TGCTGTTGACAGTGAGCGCTCACAAATGTATAATGTGTTATAGTGAAGCCAC AGATGTATAACACATTATACATTTGTGATTGCCTACTGCCTCGGA
CXCR4.1354	TGCTGTTGACAGTGAGCGACCAGCTGTTTATGCATAGATATAGTGAAGCCA CAGATGTATATCTATGCATAAACAGCTGGGTGCCTACTGCCTCGGA
CXCR4.1514	TGCTGTTGACAGTGAGCGCACAGTGTACAGTCTTGTATTATAGTGAAGCCAC AGATGTATAATACAAGACTGTACACTGTATGCCTACTGCCTCGGA
CXCR4.1036	TGCTGTTGACAGTGAGCGATCCAGCTAACACAGATGTAAATAGTGAAGCCA CAGATGTATTTACATCTGTGTTAGCTGGAGTGCCTACTGCCTCGGA
CXCR4.1037	TGCTGTTGACAGTGAGCGCCCAGCTAACACAGATGTAAAATAGTGAAGCCA CAGATGTATTTTACATCTGTGTTAGCTGGATGCCTACTGCCTCGGA
DIS3.171	TGCTGTTGACAGTGAGCGACCGCGAGAAGTGTTACTAATATAGTGAAGCCA CAGATGTATATTAGTAACACTTCTCGCGGGTGCCTACTGCCTCGGA
DIS3.3897	TGCTGTTGACAGTGAGCGAACAGAGATTAGTAATGATTAATAGTGAAGCCA CAGATGTATTAATCATTACTAATCTCTGTGTGCCTACTGCCTCGGA
DIS3.2650	TGCTGTTGACAGTGAGCGACCAGTTATTCTTCAAAAGCAATAGTGAAGCCA CAGATGTATTGCTTTTGAAGAATAACTGGGTGCCTACTGCCTCGGA
DIS3.940	TGCTGTTGACAGTGAGCGCTGGCGACAATGAAGAAAATAATAGTGAAGCCA CAGATGTATTATTTTCTTCATTGTCGCCATTGCCTACTGCCTCGGA
EXOSC1.219	TGCTGTTGACAGTGAGCGACCAGATGTGGGAGCTATTGTATAGTGAAGCCA CAGATGTATACAATAGCTCCCACATCTGGCTGCCTACTGCCTCGGA
EXOSC1.177	TGCTGTTGACAGTGAGCGCCCAGTGGTGTCTGTAGTGAGATAGTGAAGCCA CAGATGTATCTCACTACAGACACCACTGGATGCCTACTGCCTCGGA
EXOSC1.854	TGCTGTTGACAGTGAGCGCTCCCTCCTGACAGTCCTTATATAGTGAAGCCAC AGATGTATATAAGGACTGTCAGGAGGGATTGCCTACTGCCTCGGA
EXOSC1.227	TGCTGTTGACAGTGAGCGCGGGAGCTATTGTAACCTGTAATAGTGAAGCCA CAGATGTATTACAGGTTACAATAGCTCCCATGCCTACTGCCTCGGA
EXOSC10.1509	TGCTGTTGACAGTGAGCGAACGGATGAGTCCTACCTTGAATAGTGAAGCCA CAGATGTATTCAAGGTAGGACTCATCCGTGTGCCTACTGCCTCGGA

shRNA ID	Sequence
EXOSC10.2066	TGCTGTTGACAGTGAGCGATCCATTGACAGTTGCACAGAATAGTGAAGCCA CAGATGTATTCTGTGCAACTGTCAATGGACTGCCTACTGCCTCGGA
EXOSC10.2366	TGCTGTTGACAGTGAGCGCACGACTCAAAATTTCCAAGAATAGTGAAGCCA CAGATGTATTCTTGGAAATTTTGAGTCGTTTGCCTACTGCCTCGGA
EXOSC10.2333	TGCTGTTGACAGTGAGCGCCCCAAGGACCACAGAACAGAATAGTGAAGCCA CAGATGTATTCTGTTCTGTGGTCCTTGGGTTGCCTACTGCCTCGGA
EXOSC2.1488	TGCTGTTGACAGTGAGCGACGAGACTGTGTTTCAAAGAAATAGTGAAGCCA CAGATGTATTTCTTTGAAACACAGTCTCGCTGCCTACTGCCTCGGA
EXOSC2.1672	TGCTGTTGACAGTGAGCGATGGGTCCATGTACTCAGATAATAGTGAAGCCA CAGATGTATTATCTGAGTACATGGACCCAGTGCCTACTGCCTCGGA
EXOSC2.1057	TGCTGTTGACAGTGAGCGCTCCACTCACTGTCATTCACAATAGTGAAGCCAC AGATGTATTGTGAATGACAGTGAGTGGAATGCCTACTGCCTCGGA
EXOSC2.662	TGCTGTTGACAGTGAGCGCTCAGATCAAAGACATCTTAAATAGTGAAGCCA CAGATGTATTTAAGATGTCTTTGATCTGATTGCCTACTGCCTCGGA
EXOSC3.412	TGCTGTTGACAGTGAGCGAACAGCTAAATCTGGAGATATATAGTGAAGCCA CAGATGTATATATCTCCAGATTTAGCTGTCTGCCTACTGCCTCGGA
EXOSC3.423	TGCTGTTGACAGTGAGCGATGGAGATATATTCAAAGTTGATAGTGAAGCCA CAGATGTATCAACTTTGAATATATCTCCAGTGCCTACTGCCTCGGA
EXOSC3.884	TGCTGTTGACAGTGAGCGATGGGAGAGAAAATCATTGCAATAGTGAAGCCA CAGATGTATTGCAATGATTTTCTCTCCCACTGCCTACTGCCTCGGA
EXOSC3.1663	TGCTGTTGACAGTGAGCGCAGAACTCTATTTGTTGATAAATAGTGAAGCCAC AGATGTATTTATCAACAAATAGAGTTCTATGCCTACTGCCTCGGA
EXOSC4.319	TGCTGTTGACAGTGAGCGACCCTAGTGAACTGTCAATATATAGTGAAGCCA CAGATGTATATATTGACAGTTCACTAGGGCTGCCTACTGCCTCGGA
EXOSC4.461	TGCTGTTGACAGTGAGCGAACGCTCCCAGATTGATATCTATAGTGAAGCCA CAGATGTATAGATATCAATCTGGGAGCGTGTGCCTACTGCCTCGGA
EXOSC4.467	TGCTGTTGACAGTGAGCGACCAGATTGATATCTATGTGCATAGTGAAGCCA CAGATGTATGCACATAGATATCAATCTGGGTGCCTACTGCCTCGGA
EXOSC4.328	TGCTGTTGACAGTGAGCGCACTGTCAATATAGTTCAGCGATAGTGAAGCCA CAGATGTATCGCTGAACTATATTGACAGTTTGCCTACTGCCTCGGA
EXOSC5.920	TGCTGTTGACAGTGAGCGACCCAGCAAGGATAACATTCAATAGTGAAGCCA CAGATGTATTGAATGTTATCCTTGCTGGGGTGCCTACTGCCTCGGA
EXOSC5.207	TGCTGTTGACAGTGAGCGAGCCGAGGTGAAGGTCAGCAAATAGTGAAGCCA CAGATGTATTTGCTGACCTTCACCTCGGCCTGCCTACTGCCTCGGA
EXOSC5.527	TGCTGTTGACAGTGAGCGCTCCTACATCCAAGCAAGAAAATAGTGAAGCCA CAGATGTATTTTCTTGCTTGGATGTAGGATTGCCTACTGCCTCGGA
EXOSC5.209	TGCTGTTGACAGTGAGCGACGAGGTGAAGGTCAGCAAAGATAGTGAAGCCA CAGATGTATCTTTGCTGACCTTCACCTCGGTGCCTACTGCCTCGGA
EZH2.1418	TGCTGTTGACAGTGAGCGACCGCTGCAAAGCACAGTGCAATAGTGAAGCCA CAGATGTATTGCACTGTGCTTTGCAGCGGCTGCCTACTGCCTCGGA
EZH2.1826	TGCTGTTGACAGTGAGCGCTGCAAAAGTTATGATGGTTAATAGTGAAGCCA CAGATGTATTAACCATCATAACTTTTGCATTGCCTACTGCCTCGGA
EZH2.1700	TGCTGTTGACAGTGAGCGAAGGGAAAGTGTATGATAAATATAGTGAAGCCA CAGATGTATATTTATCATACACTTTCCCTCTGCCTACTGCCTCGGA
EZH2.572	TGCTGTTGACAGTGAGCGAAGCAGAAGAACTAAAGGAAAATAGTGAAGCC ACAGATGTATTTTCCTTTAGTTCTTCTGCTGTGCCTACTGCCTCGGA
FOXO1.5499	TGCTGTTGACAGTGAGCGCCCCGAGTTTAGTAACAGTGCATAGTGAAGCCA CAGATGTATGCACTGTTACTAAACTCGGGTTGCCTACTGCCTCGGA
FOXO1.4493	TGCTGTTGACAGTGAGCGATGCCAAAGTACTTGTGTACTATAGTGAAGCCA CAGATGTATAGTACACAAGTACTTTGGCACTGCCTACTGCCTCGGA
FOXO1.1151	TGCTGTTGACAGTGAGCGCTCCATGGACAACAACAGTAAATAGTGAAGCCA CAGATGTATTTACTGTTGTTGTCCATGGATTGCCTACTGCCTCGGA

shRNA ID	Sequence
FOXO1.2421	TGCTGTTGACAGTGAGCGATCCAAAGATGTCTTTCACCAATAGTGAAGCCA CAGATGTATTGGTGAAAGACATCTTTGGACTGCCTACTGCCTCGGA
ID3.155	TGCTGTTGACAGTGAGCGACCGGAGGAAGCCTGTTTGCAATAGTGAAGCCA CAGATGTATTGCAAACAGGCTTCCTCCGGCTGCCTACTGCCTCGGA
ID3.683	TGCTGTTGACAGTGAGCGATCCGGAACTTGTCATCTCCAATAGTGAAGCCAC AGATGTATTGGAGATGACAAGTTCCGGAGTGCCTACTGCCTCGGA
ID3.159	TGCTGTTGACAGTGAGCGAAGGAAGCCTGTTTGCAATTTATAGTGAAGCCA CAGATGTATAAATTGCAAACAGGCTTCCTCTGCCTACTGCCTCGGA
ID3.1097	TGCTGTTGACAGTGAGCGCTGCCCTGATTTATGAACTCTATAGTGAAGCCAC AGATGTATAGAGTTCATAAATCAGGGCAATGCCTACTGCCTCGGA
IFI35.401	TGCTGTTGACAGTGAGCGACCGGAAGTGCCTAAGTCTTTATAGTGAAGCCA CAGATGTATAAAGACTTAGGCACTTCCGGGTGCCTACTGCCTCGGA
IFI35.496	TGCTGTTGACAGTGAGCGATGAGCAGGTGCTGCAACAAAATAGTGAAGCCA CAGATGTATTTTGTTGCAGCACCTGCTCAGTGCCTACTGCCTCGGA
IFI35.1164	TGCTGTTGACAGTGAGCGCGGAGGTCTGTATGTTCACCAATAGTGAAGCCA CAGATGTATTGGTGAACATACAGACCTCCTTGCCTACTGCCTCGGA
IFI35.1166	TGCTGTTGACAGTGAGCGAAGGTCTGTATGTTCACCAACATAGTGAAGCCA CAGATGTATGTTGGTGAACATACAGACCTCTGCCTACTGCCTCGGA
IKZF1.4556	TGCTGTTGACAGTGAGCGATGCCTCTATGATGATGTTAAATAGTGAAGCCAC AGATGTATTTAACATCATCATAGAGGCAGTGCCTACTGCCTCGGA
IKZF1.3593	TGCTGTTGACAGTGAGCGCTGCGTAGTTTCTTCTGTTGTATAGTGAAGCCAC AGATGTATACAACAGAAGAAACTACGCAATGCCTACTGCCTCGGA
IKZF1.3726	TGCTGTTGACAGTGAGCGACCTGTTGATTACAGCTAGTAATAGTGAAGCCA CAGATGTATTACTAGCTGTAATCAACAGGCTGCCTACTGCCTCGGA
IKZF1.4203	TGCTGTTGACAGTGAGCGCCCAGCAATAATGAATGGTCAATAGTGAAGCCA CAGATGTATTGACCATTCATTATTGCTGGATGCCTACTGCCTCGGA
IKZF3.1374	TGCTGTTGACAGTGAGCGCCCAGTTGTCTATAGAAGAAAATAGTGAAGCCA CAGATGTATTTTCTTCTATAGACAACTGGTTGCCTACTGCCTCGGA
IKZF3.1004	TGCTGTTGACAGTGAGCGACCGGAGATGTGCTATATGAAATAGTGAAGCCA CAGATGTATTTCATATAGCACATCTCCGGGTGCCTACTGCCTCGGA
IKZF3.1003	TGCTGTTGACAGTGAGCGCCCCGGAGATGTGCTATATGAATAGTGAAGCCA CAGATGTATTCATATAGCACATCTCCGGGATGCCTACTGCCTCGGA
IKZF3.7627	TGCTGTTGACAGTGAGCGCTCCCAAGATCATACTAGATCATAGTGAAGCCA CAGATGTATGATCTAGTATGATCTTGGGATTGCCTACTGCCTCGGA
ILF2.33	TGCTGTTGACAGTGAGCGCCGCCTCTTCAGTTGTCTGCTATAGTGAAGCCAC AGATGTATAGCAGACAACTGAAGAGGCGTTGCCTACTGCCTCGGA
ILF2.1560	TGCTGTTGACAGTGAGCGCTGCTGTTGAAATGTTGTGAAATAGTGAAGCCA CAGATGTATTTCACAACATTTCAACAGCAATGCCTACTGCCTCGGA
ILF2.1827	TGCTGTTGACAGTGAGCGCCCAGAGTAAACTAGAATATCATAGTGAAGCCA CAGATGTATGATATTCTAGTTTACTCTGGTTGCCTACTGCCTCGGA
ILF2.466	TGCTGTTGACAGTGAGCGATGACCTGGTGGTGATACTCAATAGTGAAGCCA CAGATGTATTGAGTATCACCACCAGGTCAGTGCCTACTGCCTCGGA
IRF8.1634	TGCTGTTGACAGTGAGCGCTCGGTTAACTATCATTTCCAATAGTGAAGCCAC AGATGTATTGGAAATGATAGTTAACCGAATGCCTACTGCCTCGGA
IRF8.666	TGCTGTTGACAGTGAGCGACCAGATGGTGATCAGCTTCTATAGTGAAGCCA CAGATGTATAGAAGCTGATCACCATCTGGGTGCCTACTGCCTCGGA
IRF8.2538	TGCTGTTGACAGTGAGCGATCCAATTCAAATGAATGTCAATAGTGAAGCCA CAGATGTATTGACATTCATTTGAATTGGAGTGCCTACTGCCTCGGA
IRF8.2501	TGCTGTTGACAGTGAGCGCTGCTGAAGTGTTCATAAGATATAGTGAAGCCA CAGATGTATATCTTATGAACACTTCAGCAATGCCTACTGCCTCGGA
KHDRBS1.1578	TGCTGTTGACAGTGAGCGCCCCAGGATTCCTGTTGCTTTATAGTGAAGCCAC AGATGTATAAAGCAACAGGAATCCTGGGATGCCTACTGCCTCGGA

shRNA ID	Sequence
KHDRBS1.2092	TGCTGTTGACAGTGAGCGAGCGGTTGAATCTTAAGGTTAATAGTGAAGCCA CAGATGTATTAACCTTAAGATTCAACCGCCTGCCTACTGCCTCGGA
KHDRBS1.2603	TGCTGTTGACAGTGAGCGCTCCACTTTGTACATAAGTTAATAGTGAAGCCAC AGATGTATTAACTTATGTACAAAGTGGATTGCCTACTGCCTCGGA
KHDRBS1.1986	TGCTGTTGACAGTGAGCGAAGGGCTGTTTAAGATTTAAAATAGTGAAGCCA CAGATGTATTTTAAATCTTAAACAGCCCTGTGCCTACTGCCTCGGA
LDB1.705	TGCTGTTGACAGTGAGCGCCCGACTCTGTGTGATACTCGATAGTGAAGCCAC AGATGTATCGAGTATCACACAGAGTCGGATGCCTACTGCCTCGGA
LDB1.308	TGCTGTTGACAGTGAGCGAAGGATGGACCAAAGAGATATATAGTGAAGCCA CAGATGTATATATCTCTTTGGTCCATCCTCTGCCTACTGCCTCGGA
LDB1.386	TGCTGTTGACAGTGAGCGCCGGAGCTGTACTATGTTCTTATAGTGAAGCCAC AGATGTATAAGAACATAGTACAGCTCCGTTGCCTACTGCCTCGGA
LDB1.727	TGCTGTTGACAGTGAGCGACCCATGCAAGAGCTCATGTCATAGTGAAGCCA CAGATGTATGACATGAGCTCTTGCATGGGCTGCCTACTGCCTCGGA
LEO1.1469	TGCTGTTGACAGTGAGCGCACGCACAGAAATGATTAAGAATAGTGAAGCCA CAGATGTATTCTTAATCATTTCTGTGCGTTTGCCTACTGCCTCGGA
LEO1.1143	TGCTGTTGACAGTGAGCGCCCCAAAGTAAACACTGATTTATAGTGAAGCCA CAGATGTATAAATCAGTGTTTACTTTGGGTTGCCTACTGCCTCGGA
LEO1.1122	TGCTGTTGACAGTGAGCGAACCAGAATAGAAGTAGAAATATAGTGAAGCCA CAGATGTATATTTCTACTTCTATTCTGGTCTGCCTACTGCCTCGGA
LEO1.1733	TGCTGTTGACAGTGAGCGATCAAAGATTACTCAAAGCAAATAGTGAAGCCA CAGATGTATTTGCTTTGAGTAATCTTTGAGTGCCTACTGCCTCGGA
LRRFIP1.595	TGCTGTTGACAGTGAGCGATCCTACCAAGATGACAAAAGATAGTGAAGCCA CAGATGTATCTTTTGTCATCTTGGTAGGACTGCCTACTGCCTCGGA
LRRFIP1.259	TGCTGTTGACAGTGAGCGCAGCAGAAGTTGAAGAGAAATATAGTGAAGCCA CAGATGTATATTTCTCTTCAACTTCTGCTATGCCTACTGCCTCGGA
LRRFIP1.445	TGCTGTTGACAGTGAGCGACCACAGTATACTGCAATTTCATAGTGAAGCCA CAGATGTATGAAATTGCAGTATACTGTGGGTGCCTACTGCCTCGGA
LRRFIP1.592	TGCTGTTGACAGTGAGCGCAGGTCCTACCAAGATGACAAATAGTGAAGCCA CAGATGTATTTGTCATCTTGGTAGGACCTTTGCCTACTGCCTCGGA
Luc.1309	TGCTGTTGACAGTGAGCGCCCGCCTGAAGTCTCTGATTAATAGTGAAGCCAC AGATGTATTAATCAGAGACTTCAGGCGGTTGCCTACTGCCTCGGA
MCM3AP.6081	TGCTGTTGACAGTGAGCGCCCCGAAGAGTTTCTGTTTTTATAGTGAAGCCAC AGATGTATAAAAACAGAAACTCTTCGGGATGCCTACTGCCTCGGA
MCM3AP.1129	TGCTGTTGACAGTGAGCGAACGATACAGGATGTTTTCAAATAGTGAAGCCA CAGATGTATTTGAAAACATCCTGTATCGTCTGCCTACTGCCTCGGA
MCM3AP.5343	TGCTGTTGACAGTGAGCGCCGCCTTGTGTATCAACCACAATAGTGAAGCCA CAGATGTATTGTGGTTGATACACAAGGCGATGCCTACTGCCTCGGA
MCM3AP.1134	TGCTGTTGACAGTGAGCGCACAGGATGTTTTCAAAAGCAATAGTGAAGCCA CAGATGTATTGCTTTTGAAAACATCCTGTATGCCTACTGCCTCGGA
MED10.232	TGCTGTTGACAGTGAGCGCACCGTTAGAAGTTTTTGAATATAGTGAAGCCAC AGATGTATATTCAAAAACTTCTAACGGTATGCCTACTGCCTCGGA
MED10.47	TGCTGTTGACAGTGAGCGAGCGGAGAAGTTTGACCACCTATAGTGAAGCCA CAGATGTATAGGTGGTCAAACTTCTCCGCCTGCCTACTGCCTCGGA
MED10.331	TGCTGTTGACAGTGAGCGCAGGCAAGATCGACACCATGAATAGTGAAGCCA CAGATGTATTCATGGTGTCGATCTTGCCTTTGCCTACTGCCTCGGA
MED10.845	TGCTGTTGACAGTGAGCGCCCGAAAGCTAAATAACGACTATAGTGAAGCCA CAGATGTATAGTCGTTATTTAGCTTTCGGATGCCTACTGCCTCGGA
MED13.9423	TGCTGTTGACAGTGAGCGAACCAAGTGTACTTATATGTAATAGTGAAGCCA CAGATGTATTACATATAAGTACACTTGGTGTGCCTACTGCCTCGGA
MED13.6902	TGCTGTTGACAGTGAGCGCCCCAGTCATGTTAAATGAATATAGTGAAGCCA CAGATGTATATTCATTTAACATGACTGGGTTGCCTACTGCCTCGGA

shRNA ID	Sequence
MED13.8801	TGCTGTTGACAGTGAGCGATCAGTTGTTTAAAATCACTAATAGTGAAGCCAC AGATGTATTAGTGATTTTAAACAACTGAGTGCCTACTGCCTCGGA
MED13.2621	TGCTGTTGACAGTGAGCGACCCAATGAATATGAATAATAATAGTGAAGCCA CAGATGTATTATTATTCATATTCATTGGGGTGCCTACTGCCTCGGA
MED20.1821	TGCTGTTGACAGTGAGCGCTGGCTGTTATTTTGCAGTTAATAGTGAAGCCAC AGATGTATTAACTGCAAAATAACAGCCAATGCCTACTGCCTCGGA
MED20.1731	TGCTGTTGACAGTGAGCGATGACTACAAGATGAAAGACAATAGTGAAGCCA CAGATGTATTGTCTTTCATCTTGTAGTCACTGCCTACTGCCTCGGA
MED20.1451	TGCTGTTGACAGTGAGCGCAGGGCAGATTTAACTCTACAATAGTGAAGCCA CAGATGTATTGTAGAGTTAAATCTGCCCTATGCCTACTGCCTCGGA
MED20.694	TGCTGTTGACAGTGAGCGCCCAGTACATGGAACTCTTCAATAGTGAAGCCA CAGATGTATTGAAGAGTTCCATGTACTGGATGCCTACTGCCTCGGA
MED21.2395	TGCTGTTGACAGTGAGCGCCCTCAAGATATTAATACTATATAGTGAAGCCAC AGATGTATATAGTATTAATATCTTGAGGTTGCCTACTGCCTCGGA
MED21.1930	TGCTGTTGACAGTGAGCGCACAGGATTTATCACTAGTAAATAGTGAAGCCA CAGATGTATTTACTAGTGATAAATCCTGTTTGCCTACTGCCTCGGA
MED21.1823	TGCTGTTGACAGTGAGCGAAGAGAAGATTCAGTCAGAAAATAGTGAAGCCA CAGATGTATTTTCTGACTGAATCTTCTCTCTGCCTACTGCCTCGGA
MED21.1832	TGCTGTTGACAGTGAGCGCTCAGTCAGAAAACTTATTCAATAGTGAAGCCA CAGATGTATTGAATAAGTTTTCTGACTGAATGCCTACTGCCTCGGA
MED4.1794	TGCTGTTGACAGTGAGCGCCCCAGAGATAATATCAATTCATAGTGAAGCCA CAGATGTATGAATTGATATTATCTCTGGGTTGCCTACTGCCTCGGA
MED4.791	TGCTGTTGACAGTGAGCGCACAGAATTGAATACTGTAGAATAGTGAAGCCA CAGATGTATTCTACAGTATTCAATTCTGTATGCCTACTGCCTCGGA
MED4.1045	TGCTGTTGACAGTGAGCGCTCCAGTTTTGCTGGATTTGTATAGTGAAGCCAC AGATGTATACAAATCCAGCAAAACTGGAATGCCTACTGCCTCGGA
MED4.100	TGCTGTTGACAGTGAGCGATGGCAATTTCAAGAAACCAAATAGTGAAGCCA CAGATGTATTTGGTTTCTTGAAATTGCCAGTGCCTACTGCCTCGGA
MED6.385	TGCTGTTGACAGTGAGCGCTGGGATCAGTTATAAACTCTATAGTGAAGCCA CAGATGTATAGAGTTTATAACTGATCCCAATGCCTACTGCCTCGGA
MED6.1558	TGCTGTTGACAGTGAGCGCCCGGCAAACAGAAAAGTTGTATAGTGAAGCCA CAGATGTATACAACTTTTCTGTTTGCCGGATGCCTACTGCCTCGGA
MED6.544	TGCTGTTGACAGTGAGCGCCCAGTGGATCAAACAAAGAAATAGTGAAGCCA CAGATGTATTTCTTTGTTTGATCCACTGGATGCCTACTGCCTCGGA
MED6.331	TGCTGTTGACAGTGAGCGCTCCCACTAGCTGATTACTATATAGTGAAGCCAC AGATGTATATAGTAATCAGCTAGTGGGATTGCCTACTGCCTCGGA
MMS19.2411	TGCTGTTGACAGTGAGCGACCAGCGGTTCTTCACAGATAATAGTGAAGCCA CAGATGTATTATCTGTGAAGAACCGCTGGCTGCCTACTGCCTCGGA
MMS19.388	TGCTGTTGACAGTGAGCGAACCGACACACAGTCTACAATATAGTGAAGCCA CAGATGTATATTGTAGACTGTGTGTCGGTCTGCCTACTGCCTCGGA
MMS19.1740	TGCTGTTGACAGTGAGCGATCAGTTCTGAGAAAAGTACTATAGTGAAGCCA CAGATGTATAGTACTTTTCTCAGAACTGAGTGCCTACTGCCTCGGA
MMS19.1235	TGCTGTTGACAGTGAGCGACCAGCCAGATCTCCTATCTTATAGTGAAGCCAC AGATGTATAAGATAGGAGATCTGGCTGGGTGCCTACTGCCTCGGA
MYC.1741	TGCTGTTGACAGTGAGCGACCCAAGGTAGTTATCCTTAAATAGTGAAGCCA CAGATGTATTTAAGGATAACTACCTTGGGGTGCCTACTGCCTCGGA
MYC.1714	TGCTGTTGACAGTGAGCGACCGGAGTTGGAAAACAATGAATAGTGAAGCCA CAGATGTATTCATTGTTTTCCAACTCCGGGTGCCTACTGCCTCGGA
MYC.929	TGCTGTTGACAGTGAGCGAACGACGAGACCTTCATCAAAATAGTGAAGCCA CAGATGTATTTTGATGAAGGTCTCGTCGTCTGCCTACTGCCTCGGA
MYC.1839	TGCTGTTGACAGTGAGCGAACGAGAACAGTTGAAACACAATAGTGAAGCCA CAGATGTATTGTGTTTCAACTGTTCTCGTCTGCCTACTGCCTCGGA

shRNA ID	Sequence
NELF.1476	TGCTGTTGACAGTGAGCGCTGGAGAGAAGCTGTTCCAGAATAGTGAAGCCA CAGATGTATTCTGGAACAGCTTCTCTCCATTGCCTACTGCCTCGGA
NELF.2779	TGCTGTTGACAGTGAGCGCACGATGTACGATACCCTCATATAGTGAAGCCA CAGATGTATATGAGGGTATCGTACATCGTTTGCCTACTGCCTCGGA
NELF.1401	TGCTGTTGACAGTGAGCGACCGCCTGATGAGCAAAGTGAATAGTGAAGCCA CAGATGTATTCACTTTGCTCATCAGGCGGCTGCCTACTGCCTCGGA
NELF.1992	TGCTGTTGACAGTGAGCGACCCGAAACAATTCAAAGGGAATAGTGAAGCCA CAGATGTATTCCCTTTGAATTGTTTCGGGGTGCCTACTGCCTCGGA
NIPBL.1269	TGCTGTTGACAGTGAGCGAACGATGGAGATTCTTCAACAATAGTGAAGCCA CAGATGTATTGTTGAAGAATCTCCATCGTCTGCCTACTGCCTCGGA
NIPBL.5065	TGCTGTTGACAGTGAGCGCCCAGGATGTTGTCATTACTAATAGTGAAGCCAC AGATGTATTAGTAATGACAACATCCTGGTTGCCTACTGCCTCGGA
NIPBL.5532	TGCTGTTGACAGTGAGCGACCCAGTGGTTTCGAGACACAATAGTGAAGCCA CAGATGTATTGTGTCTCGAAACCACTGGGCTGCCTACTGCCTCGGA
NIPBL.1068	TGCTGTTGACAGTGAGCGCCCCATCCTTCAAGTTACACAATAGTGAAGCCAC AGATGTATTGTGTAACTTGAAGGATGGGATGCCTACTGCCTCGGA
NLRC3.4143	TGCTGTTGACAGTGAGCGCTGGGAAAATTGTGAAGATAAATAGTGAAGCCA CAGATGTATTTATCTTCACAATTTTCCCATTGCCTACTGCCTCGGA
NLRC3.6299	TGCTGTTGACAGTGAGCGAAGGGATCTATTTCTTCTTGTATAGTGAAGCCAC AGATGTATACAAGAAGAAATAGATCCCTCTGCCTACTGCCTCGGA
NLRC3.1642	TGCTGTTGACAGTGAGCGCTGGGCTGCTCAAGAAGAAATATAGTGAAGCCA CAGATGTATATTTCTTCTTGAGCAGCCCATTGCCTACTGCCTCGGA
NLRC3.2767	TGCTGTTGACAGTGAGCGCAGAGCTCATGTTCTCCAGTAATAGTGAAGCCA CAGATGTATTACTGGAGAACATGAGCTCTTTGCCTACTGCCTCGGA
PAF1.791	TGCTGTTGACAGTGAGCGCACCGAGGAAGAAATATACAAATAGTGAAGCCA CAGATGTATTTGTATATTTCTTCCTCGGTATGCCTACTGCCTCGGA
PAF1.587	TGCTGTTGACAGTGAGCGACCCAATGTTCTTCTAGATCCATAGTGAAGCCAC AGATGTATGGATCTAGAAGAACATTGGGGTGCCTACTGCCTCGGA
PAF1.1147	TGCTGTTGACAGTGAGCGAACCAGATGATGTGTATGACTATAGTGAAGCCA CAGATGTATAGTCATACACATCATCTGGTGTGCCTACTGCCTCGGA
PAF1.1009	TGCTGTTGACAGTGAGCGATGCGTTGGAGATGATGTCTCATAGTGAAGCCA CAGATGTATGAGACATCATCTCCAACGCAGTGCCTACTGCCTCGGA
PDCD2.600	TGCTGTTGACAGTGAGCGCAGGAAGATTACTCAGAGATTATAGTGAAGCCA CAGATGTATAATCTCTGAGTAATCTTCCTTTGCCTACTGCCTCGGA
PDCD2.497	TGCTGTTGACAGTGAGCGCCCAGATCATCTGGACCATATATAGTGAAGCCA CAGATGTATATATGGTCCAGATGATCTGGTTGCCTACTGCCTCGGA
PDCD2.377	TGCTGTTGACAGTGAGCGATGCAGGGTTTGTGGCTGTTTATAGTGAAGCCAC AGATGTATAAACAGCCACAAACCCTGCAGTGCCTACTGCCTCGGA
PDCD2.542	TGCTGTTGACAGTGAGCGCCCAGAATTTGAAATTGTAATATAGTGAAGCCA CAGATGTATATTACAATTTCAAATTCTGGATGCCTACTGCCTCGGA
PIK3CA.3060	TGCTGTTGACAGTGAGCGACCCAAGAATGCACAAAGACAATAGTGAAGCCA CAGATGTATTGTCTTTGTGCATTCTTGGGCTGCCTACTGCCTCGGA
PIK3CA.1900	TGCTGTTGACAGTGAGCGACCAGATGTATTGCTTGGTAAATAGTGAAGCCA CAGATGTATTTACCAAGCAATACATCTGGGTGCCTACTGCCTCGGA
PIK3CA.1899	TGCTGTTGACAGTGAGCGACCCAGATGTATTGCTTGGTAATAGTGAAGCCA CAGATGTATTACCAAGCAATACATCTGGGCTGCCTACTGCCTCGGA
PIK3CA.2513	TGCTGTTGACAGTGAGCGACCAGACATCATGTCAGAGTTATAGTGAAGCCA CAGATGTATAACTCTGACATGATGTCTGGGTGCCTACTGCCTCGGA
POLE.4376	TGCTGTTGACAGTGAGCGATCGCTCCAACATGGTCTACAATAGTGAAGCCA CAGATGTATTGTAGACCATGTTGGAGCGAGTGCCTACTGCCTCGGA
POLE.1082	TGCTGTTGACAGTGAGCGCCCAGATTATGATGATTTCCTATAGTGAAGCCAC AGATGTATAGGAAATCATCATAATCTGGTTGCCTACTGCCTCGGA

shRNA ID	Sequence
POLE.3884	TGCTGTTGACAGTGAGCGACCCTGTCACTGTGAAGAGGAATAGTGAAGCCA CAGATGTATTCCTCTTCACAGTGACAGGGGTGCCTACTGCCTCGGA
POLE.3042	TGCTGTTGACAGTGAGCGATCCAAGGAAGAAGGCAAGAAATAGTGAAGCC ACAGATGTATTTCTTGCCTTCTTCCTTGGAGTGCCTACTGCCTCGGA
PTEN.2130	TGCTGTTGACAGTGAGCGCCCAGATGTTAGTGACAATGAATAGTGAAGCCA CAGATGTATTCATTGTCACTAACATCTGGTTGCCTACTGCCTCGGA
PTEN.1685	TGCTGTTGACAGTGAGCGACCAGCTAAAGGTGAAGATATATAGTGAAGCCA CAGATGTATATATCTTCACCTTTAGCTGGCTGCCTACTGCCTCGGA
PTEN.1093	TGCTGTTGACAGTGAGCGCTCGACTTAGACTTGACCTATATAGTGAAGCCAC AGATGTATATAGGTCAAGTCTAAGTCGAATGCCTACTGCCTCGGA
PTEN.3825	TGCTGTTGACAGTGAGCGACGGAGTACAACTACTATTGTATAGTGAAGCCA CAGATGTATACAATAGTAGTTGTACTCCGCTGCCTACTGCCTCGGA
RAD21.827	TGCTGTTGACAGTGAGCGAACGACATGTTAGTAAGCACTATAGTGAAGCCA CAGATGTATAGTGCTTACTAACATGTCGTCTGCCTACTGCCTCGGA
RAD21.2373	TGCTGTTGACAGTGAGCGCTGGGTTGTGTTTGTGTTCTGATAGTGAAGCCAC AGATGTATCAGAACACAAACACAACCCATTGCCTACTGCCTCGGA
RAD21.2439	TGCTGTTGACAGTGAGCGCTCAGAACAGATGTGTGCAATATAGTGAAGCCA CAGATGTATATTGCACACATCTGTTCTGAATGCCTACTGCCTCGGA
RAD21.375	TGCTGTTGACAGTGAGCGCAGCCCATGTGTTCGAGTGTAATAGTGAAGCCA CAGATGTATTACACTCGAACACATGGGCTTTGCCTACTGCCTCGGA
RPTOR.618	TGCTGTTGACAGTGAGCGCCCCGCTCAGAGTTAGAGATAATAGTGAAGCCA CAGATGTATTATCTCTAACTCTGAGCGGGATGCCTACTGCCTCGGA
RPTOR.1177	TGCTGTTGACAGTGAGCGCCCCAACTGTGGATGAAGTCAATAGTGAAGCCA CAGATGTATTGACTTCATCCACAGTTGGGTTGCCTACTGCCTCGGA
RPTOR.3156	TGCTGTTGACAGTGAGCGAACGATGCTGCTGGACACAAAATAGTGAAGCCA CAGATGTATTTTGTGTCCAGCAGCATCGTCTGCCTACTGCCTCGGA
RPTOR.1369	TGCTGTTGACAGTGAGCGACCCGTCGATCTTCGTCTACGATAGTGAAGCCAC AGATGTATCGTAGACGAAGATCGACGGGCTGCCTACTGCCTCGGA
RBBP8.1545	TGCTGTTGACAGTGAGCGCTGGAGACTGTGTGATGGATAATAGTGAAGCCA CAGATGTATTATCCATCACACAGTCTCCATTGCCTACTGCCTCGGA
RBBP8.877	TGCTGTTGACAGTGAGCGCACACTTGGACTTGGTGTTCAATAGTGAAGCCAC AGATGTATTGAACACCAAGTCCAAGTGTTTGCCTACTGCCTCGGA
RBBP8.1974	TGCTGTTGACAGTGAGCGCTGCATCAGTTCTTCAGTTAAATAGTGAAGCCAC AGATGTATTTAACTGAAGAACTGATGCAATGCCTACTGCCTCGGA
RBBP8.2574	TGCTGTTGACAGTGAGCGCTCCACTCAGACTTGTATGGAATAGTGAAGCCA CAGATGTATTCCATACAAGTCTGAGTGGAATGCCTACTGCCTCGGA
RDBP.1381	TGCTGTTGACAGTGAGCGACCCAATGCTGGTCTCAGTAAATAGTGAAGCCA CAGATGTATTTACTGAGACCAGCATTGGGGTGCCTACTGCCTCGGA
RDBP.984	TGCTGTTGACAGTGAGCGCAGGGAATACTCTCTATGTATATAGTGAAGCCA CAGATGTATATACATAGAGAGTATTCCCTTTGCCTACTGCCTCGGA
RDBP.325	TGCTGTTGACAGTGAGCGCACCAGCCAAGGTGGTGTCAAATAGTGAAGCCA CAGATGTATTTGACACCACCTTGGCTGGTTTGCCTACTGCCTCGGA
RDBP.1486	TGCTGTTGACAGTGAGCGATCAGGTTTGATCTCAGTGTAATAGTGAAGCCAC AGATGTATTACACTGAGATCAAACCTGACTGCCTACTGCCTCGGA
RELA.1019	TGCTGTTGACAGTGAGCGCCCGGATTGAGGAGAAACGTAATAGTGAAGCCA CAGATGTATTACGTTTCTCCTCAATCCGGTTGCCTACTGCCTCGGA
RELA.761	TGCTGTTGACAGTGAGCGCTGAGATCTTCCTACTGTGTGATAGTGAAGCCAC AGATGTATCACACAGTAGGAAGATCTCATTGCCTACTGCCTCGGA
RELA.1018	TGCTGTTGACAGTGAGCGAACCGGATTGAGGAGAAACGTATAGTGAAGCCA CAGATGTATACGTTTCTCCTCAATCCGGTGTGCCTACTGCCTCGGA
RELA.533	TGCTGTTGACAGTGAGCGATCAGCGCATCCAGACCAACAATAGTGAAGCCA CAGATGTATTGTTGGTCTGGATGCGCTGACTGCCTACTGCCTCGGA

shRNA ID	Sequence
RELB.2089	TGCTGTTGACAGTGAGCGCCCCGTGCACTAGCTTGTTACATAGTGAAGCCAC AGATGTATGTAACAAGCTAGTGCACGGGATGCCTACTGCCTCGGA
RELB.302	TGCTGTTGACAGTGAGCGCCGACGAGTACATCAAGGAGAATAGTGAAGCCA CAGATGTATTCTCCTTGATGTACTCGTCGATGCCTACTGCCTCGGA
RELB.301	TGCTGTTGACAGTGAGCGCTCGACGAGTACATCAAGGAGATAGTGAAGCCA CAGATGTATCTCCTTGATGTACTCGTCGATTGCCTACTGCCTCGGA
RELB.1003	TGCTGTTGACAGTGAGCGACCGTCTATGACAAGAAATCCATAGTGAAGCCA CAGATGTATGGATTTCTTGTCATAGACGGGTGCCTACTGCCTCGGA
Ren.660	TGCTGTTGACAGTGAGCGACTCGTGAAATCCCGTTAGTAATAGTGAAGCCA CAGATGTATTACTAACGGGATTTCACGAGGTGCCTACTGCCTCGGA
Ren.713	TGCTGTTGACAGTGAGCGCAGGAATTATAATGCTTATCTATAGTGAAGCCAC AGATGTATAGATAAGCATTATAATTCCTATGCCTACTGCCTCGGA
Ren.826	TGCTGTTGACAGTGAGCGCTACTGAATTTGTCAAAGTAAATAGTGAAGCCA CAGATGTATTTACTTTGACAAATTCAGTATTGCCTACTGCCTCGGA
RICTOR.4317	TGCTGTTGACAGTGAGCGACCGGTGGATATAAATGATATATAGTGAAGCCA CAGATGTATATATCATTTATATCCACCGGGTGCCTACTGCCTCGGA
RICTOR.4752	TGCTGTTGACAGTGAGCGCCGCAAAGAAGTTCTAAGATTATAGTGAAGCCA CAGATGTATAATCTTAGAACTTCTTTGCGATGCCTACTGCCTCGGA
RICTOR.3626	TGCTGTTGACAGTGAGCGACCGAGAGAGGTTAGTAGTAGATAGTGAAGCCA CAGATGTATCTACTACTAACCTCTCTCGGCTGCCTACTGCCTCGGA
RICTOR.5890	TGCTGTTGACAGTGAGCGCAGCACTGTTAGTTTAGACCTATAGTGAAGCCAC AGATGTATAGGTCTAAACTAACAGTGCTTTGCCTACTGCCTCGGA
RNF20.1058	TGCTGTTGACAGTGAGCGATCGGAAGTTTGAGGAAATGAATAGTGAAGCCA CAGATGTATTCATTTCCTCAAACTTCCGAGTGCCTACTGCCTCGGA
RNF20.2083	TGCTGTTGACAGTGAGCGCTGGCAGCTGAGAAGAAGTCTATAGTGAAGCCA CAGATGTATAGACTTCTTCTCAGCTGCCATTGCCTACTGCCTCGGA
RNF20.219	TGCTGTTGACAGTGAGCGCACGGAGGAACTAGACATTAGATAGTGAAGCCA CAGATGTATCTAATGTCTAGTTCCTCCGTTTGCCTACTGCCTCGGA
RNF20.1034	TGCTGTTGACAGTGAGCGCTGGCGGCACAATCACTATCAATAGTGAAGCCA CAGATGTATTGATAGTGATTGTGCCGCCATTGCCTACTGCCTCGGA
RNF40.287	TGCTGTTGACAGTGAGCGCCCTGAAGGTACTACAGTTCAATAGTGAAGCCA CAGATGTATTGAACTGTAGTACCTTCAGGTTGCCTACTGCCTCGGA
RNF40.339	TGCTGTTGACAGTGAGCGCCGGCAGGCTTGTGAAGATGAATAGTGAAGCCA CAGATGTATTCATCTTCACAAGCCTGCCGTTGCCTACTGCCTCGGA
RNF40.1229	TGCTGTTGACAGTGAGCGCGCGCATCGAGTTTGAGCAGAATAGTGAAGCCA CAGATGTATTCTGCTCAAACTCGATGCGCATGCCTACTGCCTCGGA
RNF40.2534	TGCTGTTGACAGTGAGCGATCGTGAGAAAGAGAGCTTCAATAGTGAAGCCA CAGATGTATTGAAGCTCTCTTTCTCACGAGTGCCTACTGCCTCGGA
RPA3.126	TGCTGTTGACAGTGAGCGAAGGGAATACAATAGTTTCAGATAGTGAAGCCA CAGATGTATCTGAAACTATTGTATTCCCTGTGCCTACTGCCTCGGA
RPA3.1545	TGCTGTTGACAGTGAGCGCTGGATTTTCATACGATTGTAATAGTGAAGCCAC AGATGTATTACAATCGTATGAAAATCCATTGCCTACTGCCTCGGA
RPA3.1150	TGCTGTTGACAGTGAGCGAAGCCGCAGTCTTGGACCATAATAGTGAAGCCA CAGATGTATTATGGTCCAAGACTGCGGCTGTGCCTACTGCCTCGGA
RPA3.851	TGCTGTTGACAGTGAGCGCTCCTGTGATCGCAGAAAGGTATAGTGAAGCCA CAGATGTATACCTTTCTGCGATCACAGGATTGCCTACTGCCTCGGA
RRM1.2923	TGCTGTTGACAGTGAGCGCACGGATATATTGAGAATCAAATAGTGAAGCCA CAGATGTATTTGATTCTCAATATATCCGTTTGCCTACTGCCTCGGA
RRM1.1138	TGCTGTTGACAGTGAGCGCACCGATGCTGAGAGTATATAATAGTGAAGCCA CAGATGTATTATATACTCTCAGCATCGGTATGCCTACTGCCTCGGA
RRM1.1459	TGCTGTTGACAGTGAGCGCCCGCAAAGTTGTAAAAGCTCATAGTGAAGCCA CAGATGTATGAGCTTTTACAACTTTGCGGATGCCTACTGCCTCGGA

shRNA ID	Sequence
RRM1.1163	TGCTGTTGACAGTGAGCGAACAGCTCGATATGTGGATCAATAGTGAAGCCA CAGATGTATTGATCCACATATCGAGCTGTGTGCCTACTGCCTCGGA
SMARCC1.4331	TGCTGTTGACAGTGAGCGACCCAATGAGTCTAGCCTACTATAGTGAAGCCA CAGATGTATAGTAGGCTAGACTCATTGGGGTGCCTACTGCCTCGGA
SMARCC1.1250	TGCTGTTGACAGTGAGCGCCCCAATATAGAAGAAGTAGTATAGTGAAGCCA CAGATGTATACTACTTCTTCTATATTGGGTTGCCTACTGCCTCGGA
SMARCC1.1048	TGCTGTTGACAGTGAGCGATCCAGAAAGAAGAGATAGAAATAGTGAAGCC ACAGATGTATTTCTATCTCTTCTTTCTGGACTGCCTACTGCCTCGGA
SMARCC1.749	TGCTGTTGACAGTGAGCGCTCCTCACAAGACGATGAAGAATAGTGAAGCCA CAGATGTATTCTTCATCGTCTTGTGAGGAATGCCTACTGCCTCGGA
SMARCC2.4912	TGCTGTTGACAGTGAGCGATCGAGACTATGTGCTGGTATATAGTGAAGCCA CAGATGTATATACCAGCACATAGTCTCGACTGCCTACTGCCTCGGA
SMARCC2.910	TGCTGTTGACAGTGAGCGACCGAAAGAAGATTTCAGCCAATAGTGAAGCCA CAGATGTATTGGCTGAAATCTTCTTTCGGCTGCCTACTGCCTCGGA
SMARCC2.1238	TGCTGTTGACAGTGAGCGAACCGACCTGGATGAACAGGAATAGTGAAGCCA CAGATGTATTCCTGTTCATCCAGGTCGGTCTGCCTACTGCCTCGGA
SMARCC2.1837	TGCTGTTGACAGTGAGCGCGCGCACAGACATGTACACAAATAGTGAAGCCA CAGATGTATTTGTGTACATGTCTGTGCGCATGCCTACTGCCTCGGA
SMC1A.1361	TGCTGTTGACAGTGAGCGCACGGAAGAAAGTAGAGACAGATAGTGAAGCC ACAGATGTATCTGTCTCTACTTTCTTCCGTTTGCCTACTGCCTCGGA
SMC1A.6060	TGCTGTTGACAGTGAGCGATCAGACCTGGTTCAAATTCTATAGTGAAGCCAC AGATGTATAGAATTTGAACCAGGTCTGACTGCCTACTGCCTCGGA
SMC1A.7672	TGCTGTTGACAGTGAGCGCCCAGATCAGATGTGAACTTAATAGTGAAGCCA CAGATGTATTAAGTTCACATCTGATCTGGATGCCTACTGCCTCGGA
SMC1A.8484	TGCTGTTGACAGTGAGCGCACCCAGAATGTCAGACTTGTATAGTGAAGCCA CAGATGTATACAAGTCTGACATTCTGGGTATGCCTACTGCCTCGGA
SMC2.3567	TGCTGTTGACAGTGAGCGACCAGATTTACTCAATGTCAAATAGTGAAGCCA CAGATGTATTTGACATTGAGTAAATCTGGCTGCCTACTGCCTCGGA
SMC2.3098	TGCTGTTGACAGTGAGCGCTCCAAAATTCTTACAACTATATAGTGAAGCCAC AGATGTATATAGTTGTAAGAATTTTGGATTGCCTACTGCCTCGGA
SMC2.2087	TGCTGTTGACAGTGAGCGCACCAAGTTTCAAGAACTCAAATAGTGAAGCCA CAGATGTATTTGAGTTCTTGAAACTTGGTTTGCCTACTGCCTCGGA
SMC2.983	TGCTGTTGACAGTGAGCGACGAGTTAATACTAAATCTCAATAGTGAAGCCA CAGATGTATTGAGATTTAGTATTAACTCGCTGCCTACTGCCTCGGA
SMC3.3804	TGCTGTTGACAGTGAGCGCTGGGAGATGTATATAGTAATATAGTGAAGCCA CAGATGTATATTACTATATACATCTCCCAATGCCTACTGCCTCGGA
SMC3.939	TGCTGTTGACAGTGAGCGCACGCCAAGTTAGAGAATTGAATAGTGAAGCCA CAGATGTATTCAATTCTCTAACTTGGCGTTTGCCTACTGCCTCGGA
SMC3.2558	TGCTGTTGACAGTGAGCGCACAGACAGTTGCTAAATGAAATAGTGAAGCCA CAGATGTATTTCATTTAGCAACTGTCTGTTTGCCTACTGCCTCGGA
SMC3.3386	TGCTGTTGACAGTGAGCGATCCCATCAGTTGACCAGTTTATAGTGAAGCCAC AGATGTATAAACTGGTCAACTGATGGGACTGCCTACTGCCTCGGA
SMC4.4502	TGCTGTTGACAGTGAGCGACGGGTACTAAATGCATTTCAATAGTGAAGCCA CAGATGTATTGAAATGCATTTAGTACCCGGTGCCTACTGCCTCGGA
SMC4.408	TGCTGTTGACAGTGAGCGCACAGTAGAAGTTCATTTTCAATAGTGAAGCCA CAGATGTATTGAAAATGAACTTCTACTGTATGCCTACTGCCTCGGA
SMC4.2904	TGCTGTTGACAGTGAGCGCCGCAATCTGCTTCAAGAATTATAGTGAAGCCA CAGATGTATAATTCTTGAAGCAGATTGCGATGCCTACTGCCTCGGA
SMC4.3862	TGCTGTTGACAGTGAGCGCAGGATTATGAGTTGTATAAAATAGTGAAGCCA CAGATGTATTTTATACAACTCATAATCCTTTGCCTACTGCCTCGGA
SMYD3.37	TGCTGTTGACAGTGAGCGATCGCCAAATACTGTAGTGCTATAGTGAAGCCA CAGATGTATAGCACTACAGTATTTGGCGACTGCCTACTGCCTCGGA

shRNA ID	Sequence
SMYD3.1147	TGCTGTTGACAGTGAGCGCACGGCGTGTGTCTTTGTTGAATAGTGAAGCCAC AGATGTATTCAACAAAGACACACGCCGTATGCCTACTGCCTCGGA
SMYD3.431	TGCTGTTGACAGTGAGCGCTCCCAGTATCTCTTTGCTCAATAGTGAAGCCAC AGATGTATTGAGCAAAGAGATACTGGGATTGCCTACTGCCTCGGA
SMYD3.1278	TGCTGTTGACAGTGAGCGCCCACAAGAATCATTAGTTGTATAGTGAAGCCA CAGATGTATACAACTAATGATTCTTGTGGTTGCCTACTGCCTCGGA
SND1.2839	TGCTGTTGACAGTGAGCGCCCAGAAAGTGATCACAGAATATAGTGAAGCCA CAGATGTATATTCTGTGATCACTTTCTGGATGCCTACTGCCTCGGA
SND1.3252	TGCTGTTGACAGTGAGCGAAGGAAGAAACATCAAAGACTATAGTGAAGCCA CAGATGTATAGTCTTTGATGTTTCTTCCTGTGCCTACTGCCTCGGA
SND1.1373	TGCTGTTGACAGTGAGCGAACCCAGGATAAGAACAAGAAATAGTGAAGCC ACAGATGTATTTCTTGTTCTTATCCTGGGTGTGCCTACTGCCTCGGA
SND1.2921	TGCTGTTGACAGTGAGCGCCGAGCTGATGATGCAGACGAATAGTGAAGCCA CAGATGTATTCGTCTGCATCATCAGCTCGATGCCTACTGCCTCGGA
SPI1.880	TGCTGTTGACAGTGAGCGCCCGCAAGAAGATGACCTACCATAGTGAAGCCA CAGATGTATGGTAGGTCATCTTCTTGCGGTTGCCTACTGCCTCGGA
SPI1.296	TGCTGTTGACAGTGAGCGAACGGATCTATACCAACGCCAATAGTGAAGCCA CAGATGTATTGGCGTTGGTATAGATCCGTGTGCCTACTGCCTCGGA
SPI1.940	TGCTGTTGACAGTGAGCGACGAGGTCAAGAAGGTGAAGAATAGTGAAGCC ACAGATGTATTCTTCACCTTCTTGACCTCGCTGCCTACTGCCTCGGA
SPI1.937	TGCTGTTGACAGTGAGCGAGGGCGAGGTCAAGAAGGTGAATAGTGAAGCC ACAGATGTATTCACCTTCTTGACCTCGCCCGTGCCTACTGCCTCGGA
SPIB.858	TGCTGTTGACAGTGAGCGAAGGGTAGCAGTTCTTCCAGAATAGTGAAGCCA CAGATGTATTCTGGAAGAACTGCTACCCTGTGCCTACTGCCTCGGA
SPIB.3229	TGCTGTTGACAGTGAGCGCAGCTATCATTGTCACCCTAAATAGTGAAGCCAC AGATGTATTTAGGGTGACAATGATAGCTTTGCCTACTGCCTCGGA
SPIB.1575	TGCTGTTGACAGTGAGCGATCGCTGGTGGTGGTCTCTTTATAGTGAAGCCAC AGATGTATAAAGAGACCACCACCAGCGACTGCCTACTGCCTCGGA
SPIB.2753	TGCTGTTGACAGTGAGCGCTCACCTGAGGTAGGAGTTCAATAGTGAAGCCA CAGATGTATTGAACTCCTACCTCAGGTGATTGCCTACTGCCTCGGA
SSRP1.1950	TGCTGTTGACAGTGAGCGACCGAGAGAAGATCAAGTCAGATAGTGAAGCCA CAGATGTATCTGACTTGATCTTCTCTCGGCTGCCTACTGCCTCGGA
SSRP1.810	TGCTGTTGACAGTGAGCGCTGCCCAGAATGTGTTGTCAAATAGTGAAGCCA CAGATGTATTTGACAACACATTCTGGGCAATGCCTACTGCCTCGGA
SSRP1.2285	TGCTGTTGACAGTGAGCGCCGGGAGAGAACAAGAGCAAAATAGTGAAGCC ACAGATGTATTTTGCTCTTGTTCTCTCCCGATGCCTACTGCCTCGGA
SSRP1.945	TGCTGTTGACAGTGAGCGCTGGCAAGACCTTTGACTACAATAGTGAAGCCA CAGATGTATTGTAGTCAAAGGTCTTGCCATTGCCTACTGCCTCGGA
STAG1.2017	TGCTGTTGACAGTGAGCGCTCCTATGTTACTGTCAAAGTATAGTGAAGCCAC AGATGTATACTTTGACAGTAACATAGGAATGCCTACTGCCTCGGA
STAG1.1111	TGCTGTTGACAGTGAGCGCACGCAAAGAGCTGCAAGAAAATAGTGAAGCCA CAGATGTATTTTCTTGCAGCTCTTTGCGTTTGCCTACTGCCTCGGA
STAG1.4919	TGCTGTTGACAGTGAGCGCACGAAAATGATTCTCAAGTTATAGTGAAGCCA CAGATGTATAACTTGAGAATCATTTTCGTTTGCCTACTGCCTCGGA
STAG1.1237	TGCTGTTGACAGTGAGCGCTGGAGTATGGATGAAAATGTATAGTGAAGCCA CAGATGTATACATTTTCATCCATACTCCAATGCCTACTGCCTCGGA
STAG2.387	TGCTGTTGACAGTGAGCGATCGGTGGTAGATGATTGGATATAGTGAAGCCA CAGATGTATATCCAATCATCTACCACCGACTGCCTACTGCCTCGGA
STAG2.968	TGCTGTTGACAGTGAGCGCAGGAGTGTTTGTACATAGATATAGTGAAGCCA CAGATGTATATCTATGTACAAACACTCCTTTGCCTACTGCCTCGGA
STAG2.4656	TGCTGTTGACAGTGAGCGCAGGGATCTAAATTTTAAATAATAGTGAAGCCA CAGATGTATTATTTAAAATTTAGATCCCTATGCCTACTGCCTCGGA

shRNA ID	Sequence
STAG2.1098	TGCTGTTGACAGTGAGCGCTGGACTATGCATGATAAGCAATAGTGAAGCCA CAGATGTATTGCTTATCATGCATAGTCCAATGCCTACTGCCTCGGA
STAT6.1182	TGCTGTTGACAGTGAGCGACCAGAAGATCTTCAATGACAATAGTGAAGCCA CAGATGTATTGTCATTGAAGATCTTCTGGGTGCCTACTGCCTCGGA
STAT6.1567	TGCTGTTGACAGTGAGCGACGGGATCTTGCTCAGCTCAAATAGTGAAGCCA CAGATGTATTTGAGCTGAGCAAGATCCCGGTGCCTACTGCCTCGGA
STAT6.2311	TGCTGTTGACAGTGAGCGCCCCAAAGAGACAGCTCTTCTATAGTGAAGCCA CAGATGTATAGAAGAGCTGTCTCTTTGGGTTGCCTACTGCCTCGGA
STAT6.2981	TGCTGTTGACAGTGAGCGCACCTAGGATATGTTAAATGTATAGTGAAGCCA CAGATGTATACATTTAACATATCCTAGGTATGCCTACTGCCTCGGA
SUPT16H.1417	TGCTGTTGACAGTGAGCGATCCCTAGTAATCAATAGCAAATAGTGAAGCCA CAGATGTATTTGCTATTGATTACTAGGGAGTGCCTACTGCCTCGGA
SUPT16H.1854	TGCTGTTGACAGTGAGCGATCGCAAGTCTAATGTGTCCTATAGTGAAGCCAC AGATGTATAGGACACATTAGACTTGCGAGTGCCTACTGCCTCGGA
SUPT16H.4123	TGCTGTTGACAGTGAGCGATCGGTTATTTTCTCCATACAATAGTGAAGCCAC AGATGTATTGTATGGAGAAAATAACCGAGTGCCTACTGCCTCGGA
SUPT16H.2307	TGCTGTTGACAGTGAGCGCCCGGAGTAATCCGAAACTGAATAGTGAAGCCA CAGATGTATTCAGTTTCGGATTACTCCGGTTGCCTACTGCCTCGGA
SUPT4H1.735	TGCTGTTGACAGTGAGCGCAGGACTCTAGGTCAAATGTCATAGTGAAGCCA CAGATGTATGACATTTGACCTAGAGTCCTTTGCCTACTGCCTCGGA
SUPT4H1.1304	TGCTGTTGACAGTGAGCGCAGGTGAGTTGTTTTAGGATAATAGTGAAGCCA CAGATGTATTATCCTAAAACAACTCACCTTTGCCTACTGCCTCGGA
SUPT4H1.440	TGCTGTTGACAGTGAGCGAACAGCTATAAAGACCTAGCAATAGTGAAGCCA CAGATGTATTGCTAGGTCTTTATAGCTGTGTGCCTACTGCCTCGGA
SUPT4H1.995	TGCTGTTGACAGTGAGCGACCCTGGTGTCTCACACTTGTATAGTGAAGCCAC AGATGTATACAAGTGTGAGACACCAGGGCTGCCTACTGCCTCGGA
SUPT5H.534	TGCTGTTGACAGTGAGCGCTCTGTGGACTGTCAAATGTAATAGTGAAGCCA CAGATGTATTACATTTGACAGTCCACAGATTGCCTACTGCCTCGGA
SUPT5H.3479	TGCTGTTGACAGTGAGCGACCACAGCTTGCTTTTGTTGTATAGTGAAGCCAC AGATGTATACAACAAAAGCAAGCTGTGGGTGCCTACTGCCTCGGA
SUPT5H.2604	TGCTGTTGACAGTGAGCGAACAGGTCAACCCACAATACAATAGTGAAGCCA CAGATGTATTGTATTGTGGGTTGACCTGTGTGCCTACTGCCTCGGA
SUPT5H.297	TGCTGTTGACAGTGAGCGAAGCAGAGGACATTCTAGAGAATAGTGAAGCCA CAGATGTATTCTCTAGAATGTCCTCTGCTCTGCCTACTGCCTCGGA
TCF3.3359	TGCTGTTGACAGTGAGCGACCCATGGTAGATGCAAGGGAATAGTGAAGCCA CAGATGTATTCCCTTGCATCTACCATGGGGTGCCTACTGCCTCGGA
TCF3.1062	TGCTGTTGACAGTGAGCGACCCGGATCACTCAAGCAATAATAGTGAAGCCA CAGATGTATTATTGCTTGAGTGATCCGGGGTGCCTACTGCCTCGGA
TCF3.3482	TGCTGTTGACAGTGAGCGACCCTGGGAGTTTGATCTCTTATAGTGAAGCCAC AGATGTATAAGAGATCAAACTCCCAGGGCTGCCTACTGCCTCGGA
TCF3.2135	TGCTGTTGACAGTGAGCGAAGCATTGCATTTCTTGATCAATAGTGAAGCCAC AGATGTATTGATCAAGAAATGCAATGCTCTGCCTACTGCCTCGGA
TCF4.6910	TGCTGTTGACAGTGAGCGCTGGGTTTAGTGACTAGTTGAATAGTGAAGCCA CAGATGTATTCAACTAGTCACTAAACCCAATGCCTACTGCCTCGGA
TCF4.5609	TGCTGTTGACAGTGAGCGCACCATGTGTTACCAAGATGAATAGTGAAGCCA CAGATGTATTCATCTTGGTAACACATGGTATGCCTACTGCCTCGGA
TCF4.6841	TGCTGTTGACAGTGAGCGCACAGAATGTTCTAATCAAGTATAGTGAAGCCA CAGATGTATACTTGATTAGAACATTCTGTTTGCCTACTGCCTCGGA
TCF4.6516	TGCTGTTGACAGTGAGCGCACAGTTAATATTAACACTGTATAGTGAAGCCA CAGATGTATACAGTGTTAATATTAACTGTATGCCTACTGCCTCGGA
TH1L.694	TGCTGTTGACAGTGAGCGAACCTCTCTAGCTACAATTTTATAGTGAAGCCAC AGATGTATAAAATTGTAGCTAGAGAGGTCTGCCTACTGCCTCGGA

shRNA ID	Sequence
TH1L.1904	TGCTGTTGACAGTGAGCGCCCCTTTCAAGAAGCTGTTTTATAGTGAAGCCAC AGATGTATAAAACAGCTTCTTGAAAGGGTTGCCTACTGCCTCGGA
TH1L.1007	TGCTGTTGACAGTGAGCGCCCGTCCTGTTCAAGATGTTCATAGTGAAGCCAC AGATGTATGAACATCTTGAACAGGACGGTTGCCTACTGCCTCGGA
TH1L.1010	TGCTGTTGACAGTGAGCGATCCTGTTCAAGATGTTCACAATAGTGAAGCCAC AGATGTATTGTGAACATCTTGAACAGGACTGCCTACTGCCTCGGA
TRIP4.1099	TGCTGTTGACAGTGAGCGCAGCAGACTAGATGAGACAATATAGTGAAGCCA CAGATGTATATTGTCTCATCTAGTCTGCTATGCCTACTGCCTCGGA
TRIP4.625	TGCTGTTGACAGTGAGCGACGCATTGTCTGTGAACAAGAATAGTGAAGCCA CAGATGTATTCTTGTTCACAGACAATGCGCTGCCTACTGCCTCGGA
TRIP4.1807	TGCTGTTGACAGTGAGCGCCCCAGGAGAAAAGGAACTATATAGTGAAGCCA CAGATGTATATAGTTCCTTTTCTCCTGGGTTGCCTACTGCCTCGGA
TRIP4.585	TGCTGTTGACAGTGAGCGACCAGAAGCACAAGCTCATCAATAGTGAAGCCA CAGATGTATTGATGAGCTTGTGCTTCTGGCTGCCTACTGCCTCGGA
TSC1.3801	TGCTGTTGACAGTGAGCGATCGCAGTGTGTTCTAATCCAATAGTGAAGCCAC AGATGTATTGGATTAGAACACACTGCGAGTGCCTACTGCCTCGGA
TSC1.2292	TGCTGTTGACAGTGAGCGCTCAGTTGAAGTTACAAGAGAATAGTGAAGCCA CAGATGTATTCTCTTGTAACTTCAACTGATTGCCTACTGCCTCGGA
TSC1.4826	TGCTGTTGACAGTGAGCGAAGAGATGACACTAGTCAGAAATAGTGAAGCCA CAGATGTATTTCTGACTAGTGTCATCTCTCTGCCTACTGCCTCGGA
TSC1.1426	TGCTGTTGACAGTGAGCGAAGCAAAGGTTCTGTCACTCTATAGTGAAGCCA CAGATGTATAGAGTGACAGAACCTTTGCTGTGCCTACTGCCTCGGA
TSC2.2911	TGCTGTTGACAGTGAGCGACCGGAGTACTAGTCTCAACGATAGTGAAGCCA CAGATGTATCGTTGAGACTAGTACTCCGGGTGCCTACTGCCTCGGA
TSC2.1027	TGCTGTTGACAGTGAGCGCCCGGCTCTATTCTCTCAGGAATAGTGAAGCCAC AGATGTATTCCTGAGAGAATAGAGCCGGTTGCCTACTGCCTCGGA
TSC2.690	TGCTGTTGACAGTGAGCGCTCGCAAGGATGGTTCAGATGATAGTGAAGCCA CAGATGTATCATCTGAACCATCCTTGCGATTGCCTACTGCCTCGGA
TSC2.733	TGCTGTTGACAGTGAGCGACGCGTCCTCTGTGGACATAGATAGTGAAGCCA CAGATGTATCTATGTCCACAGAGGACGCGGTGCCTACTGCCTCGGA
WAPAL.1541	TGCTGTTGACAGTGAGCGATCAGTTTGTAATGTTACCATATAGTGAAGCCAC AGATGTATATGGTAACATTACAAACTGACTGCCTACTGCCTCGGA
WAPAL.4862	TGCTGTTGACAGTGAGCGAAGGAAAGACTCTCTAGATATATAGTGAAGCCA CAGATGTATATATCTAGAGAGTCTTTCCTCTGCCTACTGCCTCGGA
WAPAL.4806	TGCTGTTGACAGTGAGCGCCCCAGGGTAGTTTACACTTAATAGTGAAGCCA CAGATGTATTAAGTGTAAACTACCCTGGGTTGCCTACTGCCTCGGA
WAPAL.2089	TGCTGTTGACAGTGAGCGAACCCACAAAAGCTGTATATAATAGTGAAGCCA CAGATGTATTATATACAGCTTTTGTGGGTGTGCCTACTGCCTCGGA
YWHAB.2958	TGCTGTTGACAGTGAGCGCCCCGTGGTTGTGAAAATAGTATAGTGAAGCCA CAGATGTATACTATTTTCACAACCACGGGTTGCCTACTGCCTCGGA
YWHAB.2997	TGCTGTTGACAGTGAGCGACCCGGTTATTGATGTACTAGATAGTGAAGCCA CAGATGTATCTAGTACATCAATAACCGGGGTGCCTACTGCCTCGGA
YWHAB.1187	TGCTGTTGACAGTGAGCGCACAGCTGGTATTTGTATCTAATAGTGAAGCCAC AGATGTATTAGATACAAATACCAGCTGTTTGCCTACTGCCTCGGA
YWHAB.630	TGCTGTTGACAGTGAGCGCACCAGAAAGTAAGGTGTTCTATAGTGAAGCCA CAGATGTATAGAACACCTTACTTTCTGGTTTGCCTACTGCCTCGGA
YWHAE.1290	TGCTGTTGACAGTGAGCGAAGAGAGGTTAATCACACTATATAGTGAAGCCA CAGATGTATATAGTGTGATTAACCTCTCTCTGCCTACTGCCTCGGA
YWHAE.1466	TGCTGTTGACAGTGAGCGCCCGCTGAAATGTTGCTGAAAATAGTGAAGCCA CAGATGTATTTTCAGCAACATTTCAGCGGATGCCTACTGCCTCGGA
YWHAE.501	TGCTGTTGACAGTGAGCGATCCAAGGTTTTCTATTATAAATAGTGAAGCCAC AGATGTATTTATAATAGAAAACCTTGGACTGCCTACTGCCTCGGA

shRNA ID	Sequence
YWHAE.1562	TGCTGTTGACAGTGAGCGCAGCGTTGAAGGTGGTATGGAATAGTGAAGCCA CAGATGTATTCCATACCACCTTCAACGCTATGCCTACTGCCTCGGA
YWHAG.3449	TGCTGTTGACAGTGAGCGAACCAGTGTTAGCTTAATCTTATAGTGAAGCCAC AGATGTATAAGATTAAGCTAACACTGGTGTGCCTACTGCCTCGGA
YWHAG.2765	TGCTGTTGACAGTGAGCGCCCAGAAGAACTAGTAGAAGAATAGTGAAGCCA CAGATGTATTCTTCTACTAGTTCTTCTGGTTGCCTACTGCCTCGGA
YWHAG.1264	TGCTGTTGACAGTGAGCGCTCAGATGAAAAGAAAACTTAATAGTGAAGCCA CAGATGTATTAAGTTTTCTTTTCATCTGAATGCCTACTGCCTCGGA
YWHAG.3450	TGCTGTTGACAGTGAGCGCCCAGTGTTAGCTTAATCTTAATAGTGAAGCCAC AGATGTATTAAGATTAAGCTAACACTGGTTGCCTACTGCCTCGGA
YWHAH.1123	TGCTGTTGACAGTGAGCGCTGGGAAGCAGTTTCAGATAAATAGTGAAGCCA CAGATGTATTTATCTGAAACTGCTTCCCAATGCCTACTGCCTCGGA
YWHAH.1302	TGCTGTTGACAGTGAGCGCCCAGGCTACAGTTGATATTTATAGTGAAGCCAC AGATGTATAAATATCAACTGTAGCCTGGTTGCCTACTGCCTCGGA
YWHAH.553	TGCTGTTGACAGTGAGCGCTGACAAGTTCCTGATCAAGAATAGTGAAGCCA CAGATGTATTCTTGATCAGGAACTTGTCAATGCCTACTGCCTCGGA
YWHAH.1398	TGCTGTTGACAGTGAGCGATCCCACCTCTTTCTTCAATTATAGTGAAGCCAC AGATGTATAATTGAAGAAAGAGGTGGGAGTGCCTACTGCCTCGGA
YWHAQ.1842	TGCTGTTGACAGTGAGCGAACCACTATTGTGTGTTGCTAATAGTGAAGCCAC AGATGTATTAGCAACACACAATAGTGGTCTGCCTACTGCCTCGGA
YWHAQ.608	TGCTGTTGACAGTGAGCGCTCGAAAACAAACGATAGATAATAGTGAAGCCA CAGATGTATTATCTATCGTTTGTTTTCGATTGCCTACTGCCTCGGA
YWHAQ.1905	TGCTGTTGACAGTGAGCGACCCTATGTAACAGCAGAGTAATAGTGAAGCCA CAGATGTATTACTCTGCTGTTACATAGGGCTGCCTACTGCCTCGGA
YWHAQ.594	TGCTGTTGACAGTGAGCGCGCGTGTGGTGATGATCGAAAATAGTGAAGCCA CAGATGTATTTTCGATCATCACCACACGCATGCCTACTGCCTCGGA
YWHAZ.1676	TGCTGTTGACAGTGAGCGCTCCAAGCATAATTGTTAAGAATAGTGAAGCCA CAGATGTATTCTTAACAATTATGCTTGGATTGCCTACTGCCTCGGA
YWHAZ.409	TGCTGTTGACAGTGAGCGCGCCGCTGGTGATGACAAGAAATAGTGAAGCCA CAGATGTATTTCTTGTCATCACCAGCGGCATGCCTACTGCCTCGGA
YWHAZ.2832	TGCTGTTGACAGTGAGCGCCCCAAATATTACATTCAAATATAGTGAAGCCA CAGATGTATATTTGAATGTAATATTTGGGATGCCTACTGCCTCGGA
YWHAZ.190	TGCTGTTGACAGTGAGCGAAGGGTCGTCTCAAGTATTGAATAGTGAAGCCA CAGATGTATTCAATACTTGAGACGACCCTCTGCCTACTGCCTCGGA
SFN.1233	TGCTGTTGACAGTGAGCGACCAGTGCAAGACCGAGATTGATAGTGAAGCCA CAGATGTATCAATCTCGGTCTTGCACTGGCTGCCTACTGCCTCGGA
SFN.911	TGCTGTTGACAGTGAGCGCAGGCGCTGTTCTTGCTCCAAATAGTGAAGCCAC AGATGTATTTGGAGCAAGAACAGCGCCTATGCCTACTGCCTCGGA
SFN.197	TGCTGTTGACAGTGAGCGCCCTGCTCTCAGTAGCCTATAATAGTGAAGCCAC AGATGTATTATAGGCTACTGAGAGCAGGTTGCCTACTGCCTCGGA
SFN.1274	TGCTGTTGACAGTGAGCGATGTGACCATGTTTCCTCTCAATAGTGAAGCCAC AGATGTATTGAGAGGAAACATGGTCACACTGCCTACTGCCTCGGA

Table S4: Detailed results of the RNAi screen. For all the shRNAs used, the ratio of IgM positive (IgM^+) on IgM negative (IgM^-) cells is shown, organized by rounds of infection. Cells highlighted in red indicate hairpins scoring in the screen. The cut-off is set calculating the average of each round's Renilla ratios of IgM^+ on IgM^- cells, multiplied by two (2.564). ShRNAs in grey are inconsistent or not verified.

Round 1					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
Ren.713	1,647761194	C1D.179	1,102702703	CCNT2.4496	0,813084112
Ren.713	1,84244373	CCNC.677	1,264305177	CCNT2.2304	0,88988764
EXOSC3.423	2,987068966	CCNC.677	1,238888889	CCNT2.2304	0,899280576
EXOSC3.423	2,756198347	CCNC.677	1,148541114	CCNT2.2304	0,929906542
BCL6.2465	0,888636364	CCNC.1852	1,904290429	CDC73.597	0,858769932
BCL6.2465	0,61663286	CCNC.1852	1,672131148	CDC73.597	0,722338205
BCL6.2465	0,785714286	CCNC.1852	1,684848485	CDC73.597	0,732776618
BCL6.237	1,082687339	CCNC.456	1,290575916	CDC73.5067	1,17539267
BCL6.237	1,25	CCNC.456	1,226463104	CDC73.5067	0,984962406
BCL6.237	1,108040201	CCNC.456	1,041463415	CDC73.5067	0,858173077
BTBD1.2904	1,079155673	CCNC.1143	1,432276657	CDC73.812	1,780821918
BTBD1.2904	1,168865435	CCNC.1143	1,332386364	CDC73.812	1,328402367
BTBD1.2904	1,174934726	CCNC.1143	1,285326087	CDC73.812	1,521341463
BTBD1.1989	0,698030635	CCNH.782	1,049140049	CDC73.1639	1,407294833
BTBD1.1989	0,668041237	CCNH.782	1,402332362	CDC73.1639	1,26686217
BTBD1.1989	0,657327586	CCNH.782	1,118090452	CDC73.1639	1,794701987
BTBD1.1597	1,334285714	CCNH.354	1,426470588	CDCA5.1551	1,286472149
BTBD1.1597	0,458167331	CCNH.354	1,847750865	CDCA5.1551	1,083743842
BTBD1.1597	1,574534161	CCNH.354	1,389057751	CDCA5.1551	1,270341207
C1D.109	1,080645161	CCNH.841	1,608832808	CDCA5.1838	2,768292683
C1D.109	1,271186441	CCNH.841	1,385294118	CDCA5.1838	1,993506494
Ren.713	1,647761194	CCNH.841	1,483870968	CDCA5.1838	2,166101695
Ren.713	1,84244373	CCNL1.1015	1,584337349	CDCA5.798	2,65234375
EXOSC3.423	2,987068966	CCNL1.1015	1,107142857	CDCA5.798	1,54494382
EXOSC3.423	2,756198347	CCNL1.1015	1,37535014	CDCA5.798	1,719879518
BCL6.2465	0,888636364	CCNL1.1744	1,32122905	CDCA5.1601	2,14334471
BCL6.2465	0,61663286	CCNL1.1744	1,23495702	CDCA5.1601	1,966442953
BCL6.2465	0,785714286	CCNL1.1744	0,987309645	CDCA5.1601	2,250883392
BCL6.237	1,082687339	CCNL2.567	1,531561462	CDK9.1500	2,64940239
BCL6.237	1,25	CCNL2.567	1,17877095	CDK9.1500	2,48
BCL6.237	1,108040201	CCNL2.567	0,822072072	CDK9.1500	2,08650519
BTBD1.2904	1,079155673	CCNL2.350	1,354466859	CDK9.402	2,104377104
BTBD1.2904	1,168865435	CCNL2.350	1,425981873	CDK9.402	1,370165746
BTBD1.2904	1,174934726	CCNL2.350	1,578461538	CDK9.402	1,613981763
BTBD1.1989	0,698030635	CCNL2.604	1,655737705	COBRA1.293	3,590425532
BTBD1.1989	0,668041237	CCNL2.604	1,461309524	COBRA1.293	3,37628866
BTBD1.1989	0,657327586	CCNL2.604	1,522012579	COBRA1.293	3,158163265
BTBD1.1597	1,334285714	CCNL2.349	1,39039039	CTCF.2059	3,649746193
BTBD1.1597	0,458167331	CCNL2.349	1,308823529	CTCF.2059	4,744680851
BTBD1.1597	1,574534161	CCNL2.349	1,714285714	CTCF.2059	5,185714286
C1D.109	1,080645161	CCNT1.581	1,296735905	CTCF.557	6,008474576
C1D.109	1,271186441	CCNT1.581	1,539634146	CTCF.557	6,627118644

Round 1					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
C1D.109	1,232620321	CCNT1.581	1,537091988	CTCF.557	6,938053097
C1D.625	0,677824268	CCNT1.2675	3,147783251	CTR9.4186	3,084158416
C1D.625	0,748478702	CCNT1.2675	2,369047619	CTR9.4186	2,803827751
C1D.625	0,694736842	CCNT1.2675	2,128404669	CTR9.4186	2,37944664
C1D.905	1,506535948	CCNT1.3065	1,222527473	CTR9.2339	3,093264249
C1D.905	1,757475083	CCNT1.3065	1,207282913	CTR9.2339	3,389473684
C1D.905	1,118110236	CCNT1.3065	1,416666667	CTR9.2339	2,48
C1D.179	1,559766764	CCNT2.4496	1,078125		
C1D.179	1,505952381	CCNT2.4496	0,994936709		

Round 2					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
Ren.713	1,072727273	EXOSC10.2366	0,238493724	EXOSC5.527	1,2425
Ren.713	1,37458194	EXOSC10.2333	0,844262295	EZH2.1826	0,789366053
EXOSC3.423	2,397196262	EXOSC10.2333	1,121140143	EZH2.1826	0,776422764
EXOSC3.423	1,24120603	EXOSC10.2333	0,834016393	EZH2.1826	0,836864407
BACH1.4982	1,056179775	EXOSC2.1488	1,022727273	EZH2.1700	1,036281179
BACH1.4982	0,870967742	EXOSC2.1488	0,923240938	EZH2.1700	0,89957265
BACH1.4982	0,946547884	EXOSC2.1488	0,982378855	EZH2.1700	0,927966102
BACH1.2716	1,435393258	EXOSC2.1672	0,923076923	EZH2.572	0,920879121
BACH1.2716	1,438547486	EXOSC2.1672	0,764356436	EZH2.572	0,732251521
BACH1.2716	1,260526316	EXOSC2.1672	0,826175869	EZH2.572	0,741106719
CTR9.4202	2,515267176	EXOSC2.1057	0,502645503	FOXO1.5499	1,004618938
CTR9.4202	2,941176471	EXOSC2.1057	0,526978417	FOXO1.5499	0,898089172
CTR9.4202	2,917355372	EXOSC2.1057	0,447457627	FOXO1.5499	0,831556503
CXCR4.1354	2,936170213	EXOSC2.662	3,026666667	FOXO1.4493	1,614035088
CXCR4.1354	2,444029851	EXOSC2.662	2,969565217	FOXO1.4493	1,537572254
CXCR4.1354	2,298245614	EXOSC2.662	3,100456621	FOXO1.4493	1,311053985
CXCR4.1514	1,231343284	EXOSC3.412	2,666666667	FOXO1.1151	1,409937888
CXCR4.1514	1,307106599	EXOSC3.412	2,325670498	FOXO1.1151	1,537931034
CXCR4.1514	1,423180593	EXOSC3.412	2,95154185	FOXO1.1151	1,666666667
CXCR4.1036	0,949115044	EXOSC3.423	0,678846154	FOXO1.2421	2,740196078
CXCR4.1036	1,013888889	EXOSC3.423	0,683908046	FOXO1.2421	3,543209877
CXCR4.1036	0,728542914	EXOSC3.423	0,753451677	FOXO1.2421	3,162436548
CXCR4.1037	1,433802817	EXOSC3.884	1,568181818	ID3.155	1,382142857
CXCR4.1037	1,602985075	EXOSC3.884	1,795731707	ID3.155	2,5
CXCR4.1037	1,697247706	EXOSC3.884	1,765243902	ID3.155	1,158959538
DIS3.3897	1,325459318	EXOSC3.1663	1,130434783	ID3.683	1,221830986
DIS3.3897	1,285347044	EXOSC3.1663	0,88034188	ID3.683	1,149253731
DIS3.3897	1,185990338	EXOSC3.1663	0,947019868	ID3.683	0,729292929
DIS3.2650	0,803382664	EXOSC4.319	1,976430976	ID3.159	2,341365462
DIS3.2650	0,811088296	EXOSC4.319	1,779874214	ID3.159	2
DIS3.2650	0,794297352	EXOSC4.319	1,872131148	ID3.159	2,64806867
DIS3.940	2,006711409	EXOSC4.461	1,120481928	IKZF1.3593	0,613226453
DIS3.940	1,828478964	EXOSC4.461	1,042352941	IKZF1.3593	0,564377682
DIS3.940	1,575144509	EXOSC4.461	1,011337868	IKZF1.3593	0,545634921
EXOSC10.1509	1,23989899	EXOSC4.328	0,832985386	IKZF1.3726	2,187250996

Round 2					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
EXOSC10.1509	1,255	EXOSC4.328	0,81390593	IKZF1.3726	1,916955017
EXOSC10.1509	1,393617021	EXOSC4.328	0,915217391	IKZF1.3726	1,868512111
EXOSC10.2066	3,190909091	EXOSC5.207	2,204225352	IKZF1.4203	1,627659574
EXOSC10.2066	3,43902439	EXOSC5.207	2,321299639	IKZF1.4203	2,13877551
EXOSC10.2066	3,230046948	EXOSC5.207	2,564202335	IKZF1.4203	1,94214876
EXOSC10.2366	0,138650307	EXOSC5.527	1,049065421		
EXOSC10.2366	0,193166886	EXOSC5.527	1,180693069		

Round 3					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
Ren.713	1,785932722	NIPBL.1068	1,724035608	RDBP.1381	3,287234043
EXOSC3.423	2,383211679	NIPBL.1068	1,42972973	RDBP.984	3,976331361
EXOSC3.423	3,317307692	NIPBL.1068	1,595307918	RDBP.984	3,767195767
CDCA5.1838	1	PAF1.791	1,136612022	RDBP.984	2,751054852
CDCA5.1838	0	PAF1.791	1,543352601	RDBP.325	1,474777448
CDCA5.1838	1,88961039	PAF1.791	1,542056075	RDBP.325	1,102941176
CDCA5.798	2,154362416	PAF1.1147	2,242857143	RDBP.325	1,401215805
CDCA5.798	2,071895425	PAF1.1147	2,254901961	RDBP.1486	1,765625
CDCA5.798	1,980456026	PAF1.1147	2,217391304	RDBP.1486	2,099656357
CDCA5.1601	1,928813559	RAD21.827	1,090692124	RDBP.1486	1,839590444
CDCA5.1601	1,959183673	RAD21.827	1,027459954	SMC1A.1361	1,909090909
CDCA5.1601	2,093283582	RAD21.827	0,836244541	SMC1A.1361	2,009836066
CDK9.1500	2,769565217	RAD21.2439	2,633466135	SMC1A.1361	2,2
CDK9.1500	2,951923077	RAD21.2439	2,707317073	SMC1A.8484	0,556521739
CDK9.1500	2,52016129	RAD21.2439	2,376383764	SMC1A.8484	0,484400657
CDK9.402	2,042145594	RAD21.375	2,350943396	SMC1A.8484	0,486312399
CDK9.402	1,541916168	RAD21.375	2,854077253	SMC2.3567	2,516483516
CDK9.402	1,331412104	RAD21.375	2,330827068	SMC2.3567	2,156666667
FOXO1.4493	2,032467532	RDBP.1381	2,704347826	SMC2.3567	2,733333333
FOXO1.4493	1,657894737	RDBP.1381	2,462151394	SMC2.3098	2,367647059
FOXO1.4493	1,918032787	RDBP.1381	3,287234043	SMC2.3098	2,174496644
FOXO1.1151	1,567010309	RDBP.984	3,976331361	SMC2.3098	2,789473684
FOXO1.1151	2,086466165	RDBP.984	3,767195767	SMC2.2087	2,540540541
FOXO1.1151	2,02962963	RDBP.984	2,751054852	SMC2.2087	1,659883721
FOXO1.2421	2,237918216	RDBP.325	1,474777448	SMC2.2087	1,894230769
FOXO1.2421	2,153284672	RDBP.325	1,102941176	SMC3.939	2,347826087
FOXO1.2421	2,463035019	RDBP.325	1,401215805	SMC3.939	2,183391003
LEO1.1469	2,450381679	RDBP.1486	1,765625	SMC3.939	1,900958466
LEO1.1469	1,590116279	RDBP.1486	2,099656357	SMC3.2558	1,976744186
LEO1.1469	1,996655518	RDBP.1486	1,839590444	SMC3.2558	2,09122807
LEO1.1143	3,29245283	NIPBL.5532	1,980392157	SMC3.2558	2,356877323
LEO1.1143	3,317757009	NIPBL.1068	1,724035608	SMC3.3386	1,810725552
LEO1.1143	3,119469027	NIPBL.1068	1,42972973	SMC3.3386	2,013114754
LEO1.1122	3,022522523	NIPBL.1068	1,595307918	SMC3.3386	3,311627907
LEO1.1122	2,931914894	PAF1.791	1,136612022	SMC4.408	1,660869565
LEO1.1122	2,902542373	PAF1.791	1,543352601	SMC4.408	1,621776504

Round 3					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
LEO1.1733	2,375478927	PAF1.791	1,542056075	SMC4.2904	2,418772563
LEO1.1733	2,505882353	PAF1.1147	2,242857143	SMC4.2904	2,311827957
LEO1.1733	2,223443223	PAF1.1147	2,254901961	SMC4.2904	2,255319149
NELF.2779	2,070707071	PAF1.1147	2,217391304	SMC4.3862	1,337662338
NELF.2779	2,766393443	RAD21.827	1,090692124	SMC4.3862	1,349869452
NELF.2779	3,021645022	RAD21.827	1,027459954	SMC4.3862	1,085846868
NELF.1401	3,255102041	RAD21.827	0,836244541	STAG1.2017	1,372340426
NELF.1401	2,722222222	RAD21.2439	2,633466135	STAG1.2017	1,655882353
NELF.1401	2,4743083	RAD21.2439	2,707317073	STAG1.2017	1,782874618
NIPBL.1269	2,451737452	RAD21.2439	2,376383764	STAG1.1111	1,041763341
NIPBL.1269	1,80733945	RAD21.375	2,350943396	STAG1.1111	1,370666667
NIPBL.5065	2,301470588	RAD21.375	2,854077253	STAG1.1111	1,147268409
NIPBL.5065	2,01986755	RAD21.375	2,330827068	STAG1.4919	0,930131004
NIPBL.5532	1,789473684	RDBP.1381	2,704347826	STAG1.4919	0,898268398
STAG1.4919	0,753424658	TH1L.1904	2,263157895	WAPAL.1541	1,634877384
STAG2.968	1,130541872	TH1L.1904	3,871134021	WAPAL.1541	5,281879195
STAG2.968	1,628742515	TH1L.1904	2,873469388	WAPAL.1541	2,336805556
STAG2.968	1,371900826	TH1L.1007	2,00952381	WAPAL.4862	2,003003003
STAG2.1098	1,16464891	TH1L.1007	2,193103448	WAPAL.4862	2,003003003
STAG2.1098	1,041666667	TH1L.1007	1,505376344	WAPAL.4862	0
STAG2.1098	1,068075117	TH1L.1010	2,663003663	WAPAL.4806	1,413881748
TH1L.694	2,391791045	TH1L.1010	1,602040816	WAPAL.4806	1,770833333
TH1L.694	2,081911263	TH1L.1010	2,197841727	WAPAL.4806	2,297202797
TH1L.694	2,14532872	WAPAL.1541	1,634877384		

Round 4					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
Ren.713	0,817269076	LDB1.386	3,023148148	MED20.694	1,423529412
Ren.713	0,84040404	LDB1.727	1,24205379	MED21.2395	1,83171521
EXOSC3.423	2,268551237	LDB1.727	1,244215938	MED21.2395	1,37434555
EXOSC3.423	2,197231834	LDB1.727	1,193627451	MED21.2395	1,4140625
IKZF3.1374	1,471471471	LRRFIP1.595	1,622291022	MED21.1930	1,151193634
IKZF3.1374	1,152230971	LRRFIP1.595	1,463126844	MED21.1930	1,179028133
IKZF3.1374	1,245179063	LRRFIP1.595	1,242647059	MED21.1930	1,031100478
IKZF3.1003	1,75483871	LRRFIP1.259	1,475543478	MED21.1823	1,932659933
IKZF3.1003	1,854092527	LRRFIP1.259	1,68597561	MED21.1823	1,521621622
IKZF3.1003	1,834482759	LRRFIP1.259	1,974440895	MED21.1823	1,50273224
IKZF3.7627	1,100775194	LRRFIP1.445	1,670588235	MED21.1832	1,206235012
IKZF3.7627	1,112171838	LRRFIP1.445	1,476744186	MED21.1832	1,301507538
IKZF3.7627	1,085918854	LRRFIP1.445	1,197619048	MED21.1832	1,524590164
ILF2.1560	2,752066116	LRRFIP1.592	1,467605634	MED4.1794	1,657060519
ILF2.1560	2,533864542	LRRFIP1.592	1,740181269	MED4.1794	1,387596899
ILF2.1560	2,823275862	LRRFIP1.592	1,664670659	MED4.1794	1,445623342
ILF2.1827	1,242268041	MCM3AP.1129	1,638483965	MED4.791	1,23573201
ILF2.1827	1,025056948	MCM3AP.1129	1,674486804	MED4.791	1,706395349

Round 4					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
ILF2.1827	1,018058691	MCM3AP.1129	1,571847507	MED4.791	1,44973545
IRF8.1634	1,846153846	MCM3AP.1134	2,325	MED4.100	1,545454545
IRF8.1634	1,69470405	MCM3AP.1134	2,180887372	MED4.100	1,498659517
IRF8.1634	1,5	MCM3AP.1134	1,911949686	MED4.100	1,331670823
IRF8.666	1,493670886	MED10.232	2,147727273	MED6.385	1,013129103
IRF8.666	1,132678133	MED10.232	1,721003135	MED6.385	1,385416667
IRF8.666	1,191646192	MED10.232	1,913043478	MED6.385	1,25
IRF8.2538	1,293150685	MED10.845	1,912903226	MED6.1558	0,932346723
IRF8.2538	1,1575	MED10.845	2,33203125	MED6.1558	0,944325482
IRF8.2538	1,249275362	MED10.845	1,914965986	MED6.1558	1,186761229
KHDRBS1.1578	3,976608187	MED13.9423	1,534818942	MED6.544	1,106818182
KHDRBS1.1578	4,075144509	MED13.9423	1,181818182	MED6.544	0,858
KHDRBS1.1578	3,578125	MED13.9423	1,072093023	MED6.544	0,894409938
KHDRBS1.2603	1,90070922	MED13.2621	1,838815789	MMS19.2411	1,014989293
KHDRBS1.2603	1,834532374	MED13.2621	1,438642298	MMS19.2411	1,372093023
KHDRBS1.2603	1,791808874	MED13.2621	1,560830861	MMS19.2411	0,939958592
KHDRBS1.1986	2,087912088	MED20.1821	0,622176591	MMS19.388	1
KHDRBS1.1986	1,808724832	MED20.1821	0,659528908	MMS19.388	1,4140625
KHDRBS1.1986	1,85483871	MED20.1821	0,663346614	MMS19.388	1,493297587
LDB1.705	2,088652482	MED20.1731	2,035211268	MMS19.1235	1,290640394
LDB1.705	1,769230769	MED20.1731	2,26953125	MMS19.1235	1,024070022
LDB1.705	1,735973597	MED20.1731	2,155234657	MMS19.1235	0,757170172
LDB1.308	1,957236842	MED20.1451	2,323636364	MYC.1741	1,037280702
LDB1.308	1,993265993	MED20.1451	2,08	MYC.1741	1,028571429
LDB1.308	2,256809339	MED20.1451	1,896024465	MYC.1741	1,058295964
LDB1.386	4,06741573	MED20.694	1,351744186		
LDB1.386	3,684210526	MED20.694	1,324022346		

Round 5					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
Ren.713	0,962352941	POLE.4376	1,06712963	RBBP8.2574	1,766025641
Ren.713	0,904109589	POLE.1082	1,023148148	RELA.1019	0,718562874
EXOSC3.423	1,360335196	POLE.1082	1,006711409	RELA.1019	0,781779661
EXOSC3.423	1,923076923	POLE.1082	1,480555556	RELA.1019	0,944071588
NLRC3.4143	1,601173021	POLE.3884	1,510510511	RELA.761	1,92358804
NLRC3.4143	1,348441926	POLE.3884	1,077464789	RELA.761	1,470254958
NLRC3.4143	1,692307692	POLE.3884	1,51810585	RELA.761	1,594752187
NLRC3.6299	1,126520681	POLE.3042	2,25631769	RELA.533	1,840764331
NLRC3.6299	1,187341772	POLE.3042	2,177536232	RELA.533	1,996515679
NLRC3.6299	1,732307692	POLE.3042	1,951140065	RELA.533	2,093425606
NLRC3.2767	0,649727768	PTEN.2130	2,117870722	RELB.2089	0,948235294
NLRC3.2767	0,838174274	PTEN.2130	2,159574468	RELB.2089	1,004807692
NLRC3.2767	0,930585683	PTEN.2130	1,99	RELB.2089	1,027777778
PDCD2.600	2,305970149	PTEN.1685	1,618075802	RELB.302	1,176616915
PDCD2.600	2,486381323	PTEN.1685	1,882352941	RELB.302	1,510204082

Round 5					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
PDCD2.600	2,916666667	PTEN.1685	1,706060606	RELB.302	1,483146067
PDCD2.497	1,927868852	PTEN.1093	1,721649485	RELB.301	1,232375979
PDCD2.497	1,855769231	PTEN.1093	1,363395225	RELB.301	1,293010753
PDCD2.497	1,837209302	PTEN.1093	1,479224377	RELB.301	1,316939891
PDCD2.542	1,539156627	PTEN.3825	2,982142857	RELB.1003	1,00462963
PDCD2.542	0,973333333	PTEN.3825	2,804444444	RELB.1003	1,019002375
PDCD2.542	1,667692308	PTEN.3825	2,560311284	RELB.1003	0,956818182
PIK3CA.3060	1,494413408	RPTOR.618	1,104477612	RICTOR.4317	0,997747748
PIK3CA.3060	2,058020478	RPTOR.618	1,076738609	RICTOR.4317	0,851694915
PIK3CA.3060	1,71686747	RPTOR.618	1,091566265	RICTOR.4317	0,85193133
PIK3CA.1900	1,713846154	RPTOR.3156	1,709480122	RICTOR.3626	1,028169014
PIK3CA.1900	1,56626506	RPTOR.3156	1,835443038	RICTOR.3626	1,023041475
PIK3CA.1900	1,527932961	RPTOR.3156	1,88961039	RICTOR.3626	0,96
PIK3CA.1899	1,84789644	RBBP8.1545	1,087939698	RICTOR.5890	0,884782609
PIK3CA.1899	2,017123288	RBBP8.1545	1,169444444	RICTOR.5890	1,106493506
PIK3CA.1899	2,904977376	RBBP8.1545	1,119221411	RICTOR.5890	0,977116705
PIK3CA.2513	1,528901734	RBBP8.877	1,260638298	RNF20.219	0,846315789
PIK3CA.2513	1,488372093	RBBP8.877	1,204488778	RNF20.219	0,895652174
PIK3CA.2513	1,88961039	RBBP8.877	1,2	RNF20.219	0,926174497
POLE.4376	1,418539326	RBBP8.2574	1,917763158	RNF20.1034	0,888888889
POLE.4376	1,190231362	RBBP8.2574	2,073170732	RNF20.1034	0,831081081
RNF20.1034	0,9030837	SMARCC1.1048	1,084367246	SMYD3.37	1,217391304
RNF40.287	1,635542169	SMARCC1.1048	1,079404467	SMYD3.37	1,024937656
RNF40.287	1,403183024	SMARCC1.1048	1,099255583	SMYD3.1147	1,617647059
RNF40.287	1,659763314	SMARCC1.749	1,326145553	SMYD3.1147	1,316753927
RNF40.2534	3,626943005	SMARCC1.749	1,199488491	SMYD3.1147	1,478873239
RNF40.2534	3,014018692	SMARCC1.749	1,322916667	SMYD3.431	1,179948586
RNF40.2534	2,871244635	SMARCC2.4912	1,891891892	SMYD3.431	1,142144638
SMARCC1.1250	1,2899729	SMARCC2.4912	1,918918919	SMYD3.431	0,937078652
SMARCC1.1250	1,210918114	SMARCC2.4912	4,227272727		
SMARCC1.1250	1,132387707	SMYD3.37	1,175531915		

Round 6					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
Ren.713	1,093506494	SUPT16H.2307	2,861344538	YWHAB.630	6,739495798
Ren.713	1,152777778	SUPT4H1.1304	0,518304432	YWHAE.1290	4,779874214
EXOSC3.423	2,261029412	SUPT4H1.1304	0,536630037	YWHAE.1290	3,346938776
EXOSC3.423	2,247311828	SUPT4H1.1304	0,965417867	YWHAE.1290	3,174528302
SMYD3.1278	1,272988506	SUPT4H1.440	1,408045977	YWHAE.501	1,871527778
SMYD3.1278	1,617363344	SUPT4H1.440	1,094292804	YWHAE.501	1,967153285
SMYD3.1278	0,970720721	SUPT4H1.440	2,019607843	YWHAE.501	2,091911765
SND1.2839	0,730337079	SUPT5H.534	0,804147465	YWHAE.1562	1,599406528
SND1.2839	0,643006263	SUPT5H.534	0,671936759	YWHAE.1562	2,135036496
SND1.2839	0,992537313	SUPT5H.534	1,085959885	YWHAE.1562	1,620795107
SND1.1373	1,465875371	SUPT5H.3479	1,121287129	YWHAG.3449	2,737556561

Round 6					
shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells	shRNA ID	Ratio IgM+/IgM- cells
SND1.1373	1,551724138	SUPT5H.3479	1,075471698	YWHAG.3449	2,031358885
SND1.1373	1,37797619	SUPT5H.3479	1,25	YWHAG.3449	2,10989011
SPI1.296	0,706638116	SUPT5H.2604	2,018450185	YWHAG.2765	2,719298246
SPI1.296	0,729927007	SUPT5H.2604	1,685714286	YWHAG.2765	2,295275591
SPI1.296	1,265175719	SUPT5H.2604	1,804216867	YWHAG.2765	3,404494382
SPI1.940	4,598639456	TCF4.6910	1,29305136	YWHAG.1264	10,66749073
SPI1.940	3,178010471	TCF4.6910	1,160949868	YWHAG.1264	13,27131783
SPI1.940	4,389221557	TCF4.6910	1,138964578	YWHAG.1264	7,044642857
SPIB.858	0,984732824	TCF4.6841	0,183727034	YWHAG.3450	2,413934426
SPIB.858	0,937947494	TCF4.6841	0,243636364	YWHAG.3450	2,356481481
SPIB.858	1,178683386	TCF4.6841	0,209302326	YWHAG.3450	3,259668508
SPIB.3229	1,42	TCF4.6516	2,173076923	YWHAH.1123	1,034825871
SPIB.3229	1,101234568	TCF4.6516	2,018867925	YWHAH.1123	1,302777778
SPIB.3229	1,837037037	TCF4.6516	1,996212121	YWHAH.1123	1,819672131
SSRP1.1950	1,039900249	TRIP4.1099	3,129032258	YWHAH.553	1,33045977
SSRP1.1950	1,034666667	TRIP4.1099	2,097122302	YWHAH.553	1,100286533
SSRP1.1950	0,970515971	TRIP4.1099	1,93559322	YWHAH.553	1,130769231
SSRP1.810	1,556851312	TRIP4.625	1,329577465	YWHAH.1398	2,236220472
SSRP1.810	1,002320186	TRIP4.625	1,408571429	YWHAH.1398	1,828478964
SSRP1.810	1,829787234	TRIP4.625	1,378453039	YWHAH.1398	1,799363057
SSRP1.2285	0,882352941	TRIP4.585	1,854938272	YWHAQ.1842	3,272251309
SSRP1.2285	0,810185185	TRIP4.585	2,070469799	YWHAQ.1842	3,115207373
SSRP1.2285	0,805803571	TRIP4.585	2,343065693	YWHAQ.1842	2,82038835
SSRP1.945	1,578488372	TSC1.3801	1,645768025	YWHAQ.608	1,812286689
SSRP1.945	1,880794702	TSC1.3801	1,387362637	YWHAQ.608	1,722807018
SSRP1.945	2,189285714	TSC1.3801	1,529577465	YWHAQ.608	1,737201365
STAT6.1182	3,588541667	TSC1.2292	1,263803681	YWHAQ.1905	2,410480349
STAT6.1182	3,174603175	TSC1.2292	1,049479167	YWHAQ.1905	2,125506073
STAT6.1182	2,990338164	TSC1.2292	0,721757322	YWHAQ.1905	2,286245353
STAT6.2311	2,314136126	TSC1.4826	1,736363636	YWHAZ.1676	2,381132075
STAT6.2311	0,891891892	TSC1.4826	1,768253968	YWHAZ.1676	2,020547945
STAT6.2311	1,155172414	TSC1.4826	3,350230415	YWHAZ.1676	2,166666667
STAT6.2981	4,740506329	TSC1.1426	1,188235294	YWHAZ.2832	2,2109375
STAT6.2981	2,52173913	TSC1.1426	1,007075472	YWHAZ.2832	2,019230769
STAT6.2981	3,228971963	TSC1.1426	1,033678756	YWHAZ.2832	1,63880597
SUPT16H.1417	1,597315436	YWHAB.2958	2,128676471	YWHAZ.190	1,685358255
SUPT16H.1417	1,671480144	YWHAB.2958	1,938405797	YWHAZ.190	1,854014599
SUPT16H.1417	2,090566038	YWHAB.2958	1,930795848	YWHAZ.190	1,745980707
SUPT16H.1854	1,507122507	YWHAB.2997	2,120155039	SFN.1233	1,21
SUPT16H.1854	1,574285714	YWHAB.2997	1,705084746	SFN.1233	1,201995012
SUPT16H.1854	1,601173021	YWHAB.2997	1,727868852	SFN.1233	1,296103896
SUPT16H.4123	0,441941075	YWHAB.1187	0,979775281	SFN.1274	2,186046512
SUPT16H.4123	0,291845494	YWHAB.1187	1,227027027	SFN.1274	2,103846154
SUPT16H.4123	0,336601307	YWHAB.1187	0,979955457	SFN.1274	1,71835443
SUPT16H.2307	2,324723247	YWHAB.630	7,814814815		
SUPT16H.2307	3,12	YWHAB.630	7,267857143		

7. REFERENCES

Aitken A (2002) Functional specificity in 14-3-3 isoform interactions through dimer formation and phosphorylation. Chromosome location of mammalian isoforms and variants. Plant Mol Biol 50:993–1010.

Aitken A, Baxter H, Dubois T, Clokie S, Mackie S, Mitchell K, Peden A, Zemlickova E (2002) Specificity of 14-3-3 isoform dimer interactions and phosphorylation. Biochemical Society Transactions 30:351–360.

Anon (2008) Pillars Article: Generation of Antibody Diversity in the Immune Response of BALB/c Mice to Influenza Virus Hemagglutinin. :1–6.

Anon (2016) Diversity in the CDR3 Region of V. :1–9.

Aoufouchi S, Faili A, Zober C, D'Orlando O, Weller S, Weill J-C, Reynaud C-A (2008) Proteasomal degradation restricts the nuclear lifespan of AID. J Exp Med 205:1357–1368.

Basu U, Chaudhuri J, Alpert C, Dutt S, Ranganath S, Li G, Schrum JP, Manis JP, Alt FW (2005) The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. Nature 438:508–511.

Basu U, Meng F-L, Keim C, Grinstein V, Pefanis E, Eccleston J, Zhang T, Myers D, Wasserman CR, Wesemann DR, Januszyk K, Gregory RI, Deng H, Lima CD, Alt FW (2011) The RNA Exosome Targets the AID Cytidine Deaminase to Both Strands of Transcribed Duplex DNA Substrates. Cell 144:353–363.

Berek C, Milstein C (1987) Mutation drift and repertoire shift in the maturation of the immune response. Immunol Rev 96:23–41.

Berek C, Milstein C (1988) The dynamic nature of the antibody repertoire. Immunol Rev 105:5–26.

Betz AG, Milstein C, González-Fernández A, Pannell R, Larson T, Neuberger MS (1994) Elements regulating somatic hypermutation of an immunoglobulin kappa gene: critical role for the intron enhancer/matrix attachment region. Cell 77:239–248.

Betz AG, Rada C, Pannell R, Milstein C, Neuberger MS (1993) Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hot spots. Proc Natl Acad Sci USA 90:2385–2388.

Bransteitter, R., Pham, P., Calabrese, P., & Goodman, M. F. (2004). Biochemical Analysis of Hypermutational Targeting by Wild Type and Mutant Activation-induced Cytidine Deaminase. The Journal of Biological Chemistry, 279(49), 51612–51621.

Brar, S. S., Watson, M., & Diaz, M. (2004). Activation-induced Cytosine Deaminase (AID) Is Actively Exported out of the Nucleus but Retained by the Induction of DNA Breaks. The Journal of Biological Chemistry, *279*(25), 26395–26401.

Braselmann S, McCormick F (1995) Bcr and Raf form a complex in vivo via 14-3-3 proteins.

required to prevent mitotic catastrophe after DNA damage. Nature 401:616–620.

Chatterji M, Unniraman S, McBride KM, Schatz DG (2007) Role of activation-induced deaminase protein kinase A phosphorylation sites in Ig gene conversion and somatic hypermutation. JI 179:5274–5280.

Chaudhuri J, Khuong C, Alt FW (2004) Replication protein A interacts with AID to promote deamination of somatic hypermutation targets. Nature 430:992–998.

Conklin DS, Galaktionov K, Beach D (1995) 14-3-3 proteins associate with cdc25 phosphatases. Proc Natl Acad Sci USA 92:7892–7896.

Conticello SG, Langlois MA, Yang Z, Neuberger MS (2007) DNA Deamination in Immunity: AID in the Context of Its APOBEC Relatives. In: AID for Immunoglobulin Diversity, pp 37–73 Advances in Immunology. Elsevier.

Craparo A, Freund R, Gustafson TA (1997) 14-3-3 (epsilon) interacts with the insulin-like growth factor I receptor and insulin receptor substrate I in a phosphoserine-dependent manner. the journal of biological chemistry 272:11663–11669.

Crews S, Griffin J, Huang H, Calame K, Hood L (1981) A single VH gene segment encodes the immune response to phosphorylcholine: somatic mutation is correlated with the class of the antibody. Cell 25:59–66.

Crouch EE, Li Z, Takizawa M, Fichtner-Feigl S, Gourzi P, Montaño C, Feigenbaum L, Wilson P, Janz S, Papavasiliou FN, Casellas R (2007) Regulation of AID expression in the immune response. J Exp Med 204:1145–1156.

Dalal SN, Schweitzer CM, Gan J, DeCaprio JA (1999a) Cytoplasmic localization of human cdc25C during interphase requires an intact 14-3-3 binding site. Mol Cell Biol 19:4465–4479.

Dalal SN, Schweitzer CM, Gan J, DeCaprio JA (1999b) Cytoplasmic localization of human cdc25C during interphase requires an intact 14-3-3 binding site. Mol Cell Biol 19:4465–4479.

de Yébenes VG, Belver L, Pisano DG, González S, Villasante A, Croce C, He L, Ramiro AR (2008) miR-181b negatively regulates activation-induced cytidine deaminase in B cells. J Exp Med 205:2199–2206.

Dedeoglu F, Horwitz B, Chaudhuri J, Alt FW, Geha RS (2004) Induction of activation-induced cytidine deaminase gene expression by IL-4 and CD40 ligation is dependent on STAT6 and NFkappaB. Int Immunol 16:395–404.

Degner SC, Wong TP, Jankevicius G, Feeney AJ (2009) Cutting edge: developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. J Immunol 182:44–48.

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

Di Noia JM, Neuberger MS (2007) Molecular mechanisms of antibody somatic hypermutation. Annu Rev Biochem 76:1–22.

deaminating single stranded DNA. J Exp Med 197:1291–1296.

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 106:2758–2763.

Dorsett Y, McBride KM, Jankovic M, Gazumyan A, Thai T-H, Robbiani DF, Di Virgilio M, San-Martin BR, Heidkamp G, Schwickert TA, Eisenreich T, Rajewsky K, Nussenzweig MC (2008) MicroRNA-155 Suppresses Activation-Induced Cytidine Deaminase-Mediated Myc-Igh Translocation. Immunity 28:630–638.

Durandy A, Peron S, Taubenheim N, Fischer A (2006) Activation-induced cytidine deaminase: structure–function relationship as based on the study of mutants. Human Mutation 27:1185–1191.

Featherstone K, Wood AL, Bowen AJ, Corcoran AE (2010) The mouse immunoglobulin heavy chain V-D intergenic sequence contains insulators that may regulate ordered V(D)J recombination. the journal of biological chemistry 285:9327–9338.

Freed E, Symons M, Macdonald SG, McCormick F, Ruggieri R (1994) Binding of 14-3-3 proteins to the protein kinase Raf and effects on its activation. Science 265:1713–1716.

Fu H, Subramanian RR, Masters SC (2000) 14-3-3 proteins: structure, function, and regulation. Annu Rev Pharmacol Toxicol 40:617–647.

Garcia-Guzman M, Dolfi F, Russello M, Vuori K (1999) Cell adhesion regulates the interaction between the docking protein p130(Cas) and the 14-3-3 proteins. the journal of biological chemistry 274:5762–5768.

Garrett FE, Emelyanov AV, Sepulveda MA, Flanagan P, Volpi S, Li F, Loukinov D, Eckhardt LA, Lobanenkov VV, Birshtein BK (2005) Chromatin architecture near a potential 3' end of the igh locus involves modular regulation of histone modifications during B-Cell development and in vivo occupancy at CTCF sites. Mol Cell Biol 25:1511–1525.

Gazumyan A, Bothmer A, Klein IA, Nussenzweig MC, McBride KM (2012) Activation-Induced Cytidine Deaminase in Antibody Diversification and Chromosome Translocation. In, pp 167–190 Advances in Cancer Research. Elsevier.

Gazumyan A, Timachova K, Yuen G, Siden E, Di Virgilio M, Woo EM, Chait BT, Reina-San-Martin B, Nussenzweig MC, McBride KM (2011) Amino-terminal phosphorylation of activation-induced cytidine deaminase suppresses c-myc/IgH translocation. Mol Cell Biol 31:442–449.

Gearhart PJ, Bogenhagen DF (1983) Clusters of point mutations are found exclusively around rearranged antibody variable genes. Proc Natl Acad Sci USA 80:3439–3443.

Gonda H, Sugai M, Nambu Y, Katakai T, Agata Y, Mori KJ, Yokota Y, Shimizu A (2003) The balance between Pax5 and Id2 activities is the key to AID gene expression. J Exp Med 198:1427–1437.

Gossen M, Freundlieb S, Bender G, Müller G, Hillen W, Bujard H (1995) Transcriptional activation by tetracyclines in mammalian cells. Science 268:1766–1769

477:424–430.

Honjo T, Kinoshita K, Muramatsu M (2002) Molecular mechanism of class switch recombination: linkage with somatic hypermutation. Annu Rev Immunol 20:165–196.

Hsu SY, Kaipia A, Zhu L, Hsueh AJ (1997) Interference of BAD (Bcl-xL/Bcl-2-associated death promoter)-induced apoptosis in mammalian cells by 14-3-3 isoforms and P11. Mol Endocrinol 11:1858–1867.

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. Proc Natl Acad Sci USA 101:1975–1980.

Jankovic M, Robbiani DF, Dorsett Y, Eisenreich T, Xu Y, Tarakhovsky A, Nussenzweig A, Nussenzweig MC (2010) Role of the translocation partner in protection against AID-dependent chromosomal translocations. Proc Natl Acad Sci USA 107:187–192.

Jiang K, Pereira E, Maxfield M, Russell B, Goudelock DM, Sanchez Y (2003) Regulation of Chk1 includes chromatin association and 14-3-3 binding following phosphorylation on Ser-345. The journal of biological chemistry 278:25207–25217.

Jun G, Wing MK, Abecasis GR, Kang HM. (2015) An efficient and scalable analysis framework for variant extraction and refinement from population scale DNA sequence data. Genome research gr-176552.

Kanai F, Marignani PA, Sarbassova D, Yagi R, Hall RA, Donowitz M, Hisaminato A, Fujiwara T, Ito Y, Cantley LC, Yaffe MB (2000) TAZ: a novel transcriptional co-activator regulated by interactions with 14-3-3 and PDZ domain proteins. The EMBO Journal 19:6778–6791.

Kim S, Davis M, Sinn E, Patten P, Hood L (1981) Antibody diversity: somatic hypermutation of rearranged VH genes. Cell 27:573–581.

Kinoshita K, Tashiro J, Tomita S, Lee CG, Honjo T (1998) Target specificity of immunoglobulin class switch recombination is not determined by nucleotide sequences of S regions. Immunity 9:849–858.

Kobayashi M, Sabouri Z, Sabouri S, Kitawaki Y, Pommier Y, Abe T, Kiyonari H, Honjo T (2011) Decrease in topoisomerase I is responsible for activation-induced cytidine deaminase (AID)-dependent somatic hypermutation. Proc Natl Acad Sci USA 108:19305–19310.

Kumagai A, Yakowec PS, Dunphy WG (1998) 14-3-3 Proteins Act as Negative Regulators of the Mitotic Inducer Cdc25 in XenopusEgg Extracts. Mol Biol Cell 9:345–354.

Lam T, Thomas LM, White CA, Li G, Pone EJ, Xu Z, Casali P (2013) Scaffold functions of 14-3-3 adaptors in B cell immunoglobulin class switch DNA recombination. PLoS ONE 8:e80414–e80415.

Langmead B, Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. Nature 9:357-359.

Larijani M, Petrov AP, Kolenchenko O, Berru M, Krylov SN, Martin A (2007) AID associates with single-stranded DNA with high affinity and a long complex half-life in a sequence-independent manner. Mol Cell Biol 27:20–30.

microinjection of antisense morpholino oligos. Methods Mol Biol 518:31–41.

Lebecque SG (1990) Boundaries of somatic mutation in rearranged immunoglobulin genes: 5“ boundary is near the promoter, and 3” boundary is approximately 1 kb from V(D)J gene. J Exp Med 172:1717–1727.

Lee J, Krivega I, Dale RK, Dean A (2017) The LDB1 Complex Co-opts CTCF for Erythroid Lineage-Specific Long-Range Enhancer Interactions. CellReports 19:2490–2502.

Lee J, Kumagai A, Dunphy WG (2001) Positive regulation of Wee1 by Chk1 and 14-3-3 proteins. Hunt T, ed. Mol Biol Cell 12:551–563.

Li G, White CA, Lam T, Pone EJ, Tran DC, Hayama KL, Zan H, Xu Z, Casali P (2013) Combinatorial H3K9acS10ph histone modification in IgH locus S regions targets 14-3-3 adaptors and AID to specify antibody class-switch DNA recombination. CellReports 5:702–714.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R and 1000 Genome Project Data Processing Subgroup (2009) The Sequence alignment/map (SAM) format and SAMtools. Bioinformatics, 25, 2078-9.

Li H A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data (2011). Bioinformatics 27(21):2987-93.

Liu D, Bienkowska J, Petosa C, Collier RJ, Fu H, Liddington R (1995) Crystal structure of the zeta isoform of the 14-3-3 protein. Nature 376:191–194.

Liu M, Schatz DG (2009) Balancing AID and DNA repair during somatic hypermutation. Trends in Immunology 30:173–181.

Lopez-Girona A, Furnari B, Mondesert O, Russell P (1999) Nuclear localization of Cdc25 is regulated by DNA damage and a 14-3-3 protein. Nature 397:172–175.

Mai T, Pone EJ, Li G, Lam TS, Moehlman J, Xu Z, Casali P (2013) Induction of activation-induced cytidine deaminase-targeting adaptor 14-3-3γ is mediated by NF-κB-dependent recruitment of CFP1 to the 5“-CpG-3-”rich 14-3-3γ promoter and is sustained by E2A. J Immunol 191:1895–1906.

Marina-Zárate E, Pérez-García A, Ramiro AR (2017) CCCTC-Binding Factor Locks Premature IgH Germline Transcription and Restrains Class Switch Recombination. Front Immunol 8:751–10.

Masuyama N, Oishi K, Mori Y, Ueno T, Takahama Y, Gotoh Y (2001) Akt inhibits the orphan nuclear receptor Nur77 and T-cell apoptosis. the journal of biological chemistry 276:32799–32805.

Maul RW, Cao Z, Venkataraman L, Giorgetti CA, Press JL, Denizot Y, Du H, Sen R, Gearhart PJ (2014) Spt5 accumulation at variable genes distinguishes somatic hypermutation in germinal center B cells from ex vivo-activated cells. J Exp Med 211:2297–2306.

Maul RW, Saribasak H, Cao Z, Gearhart PJ (2015) Topoisomerase I deficiency causes RNA polymerase II accumulation and increases AID abundance in immunoglobulin variable

Hypermutation Is Limited by CRM1-dependent Nuclear Export of Activation-induced Deaminase. J Exp Med 199:1235–1244.

McBride KM, Gazumyan A, Woo EM, Barreto VM, Robbiani DF, Chait BT, Nussenzweig MC (2006) Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. Proc Natl Acad Sci USA 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.

McKean D, Huppi K, Bell M, Staudt L, Gerhard W, Weigert M (1984) Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. Proc Natl Acad Sci USA 81:3180–3184.

McKinsey TA, Zhang CL, Olson EN (2000) Activation of the myocyte enhancer factor-2 transcription factor by calcium/calmodulin-dependent protein kinase-stimulated binding of 14-3-3 to histone deacetylase 5. Proc Natl Acad Sci USA 97:14400–14405.

Methot SP, Litzler LC, Subramani PG, Eranki AK, Fifield H, Patenaude A-M, Gilmore JC, Santiago GE, Bagci H, Côté J-F, Larijani M, Verdun RE, Di Noia JM (2018) A licensing step links AID to transcription elongation for mutagenesis in B cells. Nature Communications:1–16.

Michaud NR, Fabian JR, Mathes KD, Morrison DK (1995) 14-3-3 is not essential for Raf-1 function: identification of Raf-1 proteins that are biologically activated in a 14-3-3- and Ras-independent manner. Mol Cell Biol 15:3390–3397.

Moore BW, Perez VJ. (1967) Specific acidic proteins of the nervous system. Physiological and Biochemical Aspects of Nervous Integration, ed. FD Carlson, pp. 343–59. Englewood Cliffs, NJ: Prentice-Hall

Morrison DK, Cutler RE (1997) The complexity of Raf-1 regulation. Curr Opin Cell Biol 9:174–179.

Muramatsu M, V. S. Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, 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. Chemistry 274(26), 18470–18476.

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

Muslin AJ, Tanner JW, Allen PM, Shaw AS (1996) Interaction of 14-3-3 with Signaling Proteins Is Mediated by the Recognition of Phosphoserine. Cell 84:889–897.

Muslin AJ, Xing H (2000) 14-3-3 proteins: regulation of subcellular localization by molecular interference. Cell Signal 12:703–709.

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

Transcription-coupled events associating with immunoglobulin switch region chromatin. Science 302:2137–2140.

Neuberger MS, Rada C (2007) Somatic hypermutation: activation-induced deaminase for C/G followed by polymerase η for A/T: Figure 1. J Exp Med 204:7–10.

Obsilová V, Silhan J, Boura E, Teisinger J, Obsil T (2008) 14-3-3 proteins: a family of versatile molecular regulators. Physiol Res 57 Suppl 3:S11–S21.

Odegard VH, Schatz DG (2006) Targeting of somatic hypermutation. Nat Rev Immunol 6:573–583.

Okazaki I-M, Hiai H, Kakazu N, Yamada S, Muramatsu M, Kinoshita K, Honjo T (2003) Constitutive expression of AID leads to tumorigenesis. J Exp Med 197:1173–1181.

Orthwein A, Patenaude A-M, 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.

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

Patenaude A-M, 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.

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 143:122–133.

Pefanis E, Wang J, Rothschild G, Lim J, Chao J, Rabadan R, Economides AN, Basu U (2015) Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. Nature 514:389–393.

Peng CY, Graves PR, Ogg S, Thoma RS, Byrnes MJ, Wu Z, Stephenson MT, Piwnica-Worms H (1998) C-TAK1 protein kinase phosphorylates human Cdc25C on serine 216 and promotes 14-3-3 protein binding. Cell Growth Differ 9:197–208.

Peng CY, Graves PR, Thoma RS, Wu Z, Shaw AS, Piwnica-Worms H (1997) Mitotic and G2 checkpoint control: regulation of 14-3-3 protein binding by phosphorylation of Cdc25C on serine-216. Science 277:1501–1505.

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

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

Petosa C, Masters SC, Bankston LA, Pohl J, Wang B, Fu H, Liddington RC (1998) 14-3-3zeta binds a phosphorylated Raf peptide and an unphosphorylated peptide via its conserved amphipathic groove. the journal of biological chemistry 273:16305–16310.

simulates somatic hypermutation. Nature:1–5.

Pham P, Afif SA, Shimoda M, Maeda K, Sakaguchi N, Pedersen LC, Goodman MF (2016) Structural analysis of the activation-induced deoxycytidine deaminase required in immunoglobulin diversification. DNA Repair 43:48–56.

Pozuelo Rubio M, Geraghty KM, Wong BHC, Wood NT, Campbell DG, Morrice N, Mackintosh C (2004) 14-3-3-affinity purification of over 200 human phosphoproteins reveals new links to regulation of cellular metabolism, proliferation and trafficking. Biochem J 379:395–408.

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. B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. (2014). B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. Cell 159(7), 1524–1537.

Rada C, Jarvis JM, Milstein C (2002) AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. Proc Natl Acad Sci USA 99:7003–7008.

Rajewsky K, Förster I, Cumano A (1987) Evolutionary and somatic selection of the antibody repertoire in the mouse. Science 238:1088–1094.

Ramiro AR, Jankovic M, Callen E, Difilippantonio S, Chen H-T, McBride KM, Eisenreich TR, Chen J, Dickins RA, Lowe SW, Nussenzweig A, Nussenzweig MC (2006) Role of genomic instability and p53 in AID-induced c-myc–Igh translocations. Nature 440:105–109.

Ramiro AR, Jankovic M, Eisenreich T, Difilippantonio S, Chen-Kiang S, Muramatsu M, Honjo T, Nussenzweig A, Nussenzweig MC (2004) AID Is Required for c-myc/IgH Chromosome Translocations In Vivo. Cell 118:431–438.

Ramiro AR, Stavropoulos P, Jankovic M, Nussenzweig MC (2003) Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. Nat Immunol 4:452–456.

Revy P (2000) Activation-Induced Cytidine Deaminase (AID) Deficiency Causes the Autosomal Recessive Form of the Hyper-IgM Syndrome (HIGM2). Cell:1–11.

Rittinger K, Budman J, Xu J, Volinia S, Cantley LC, Smerdon SJ, Gamblin SJ, Yaffe MB (1999) Structural Analysis of 14-3-3 Phosphopeptide Complexes Identifies a Dual Role for the Nuclear Export Signal of 14-3-3 in Ligand Binding. Molecular Cell 4:153–166.

Robbiani DF, Bothmer A, Callen E, Reina-San-Martin B, Dorsett Y, Difilippantonio S, Bolland DJ, Chen H-T, Corcoran AE, Nussenzweig A, Nussenzweig MC (2008) AID Is Required for the Chromosomal Breaks in c-myc that Lead to c-myc/IgH Translocations. Cell 135:1028–1038.

Robbiani DF, Bunting S, Feldhahn N, Bothmer A, Camps J, Deroubaix S, McBride KM, Klein IA, Stone G, Eisenreich TR, Ried T, Nussenzweig A, Nussenzweig MC (2009) AID Produces DNA Double-Strand Breaks in Non-Ig Genes and Mature B Cell Lymphomas with Reciprocal Chromosome Translocations. Molecular Cell 36:631–641.

Robbiani DF, Deroubaix S, Feldhahn N, Oliveira TY, Callen E, Wang Q, Jankovic M, Silva IT

Lymphoma. Cell 162:727–737.

Rothenfluh HS, Taylor L, Bothwell AL, Both GW, Steele EJ (1993) Somatic hypermutation in 5' flanking regions of heavy chain antibody variable regions. Eur J Immunol 23:2152–2159.

Sale JE, Neuberger MS (1998) TdT-accessible breaks are scattered over the immunoglobulin V domain in a constitutively hypermutating B cell line. Immunity 9:859–869.

Sayegh CE, Quong MW, Agata Y, Murre C (2003) E-proteins directly regulate expression of activation-induced deaminase in mature B cells. Nat Immunol 4:586–593.

Shen 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 280:1750–1752.

Sohail A, Klapacz J, Samaranayake M, Ullah A, Bhagwat AS (2003) Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. Nucleic Acids Research 31:2990–2994.

Steele EJ, Rothenfluh HS, Both GW (1992) Defining the nucleic acid substrate for somatic hypermutation. Immunol Cell Biol 70 (Pt 2):129–144.

Steinacker P, Schwarz P, Reim K, Brechlin P, Jahn O, Kratzin H, Aitken A, Wiltfang J, Aguzzi A, Bahn E, Baxter HC, Brose N, Otto M (2005) Unchanged survival rates of 14-3-3gamma knockout mice after inoculation with pathological prion protein. Mol Cell Biol 25:1339–1346.

Tashiro J, Kinoshita K, Honjo T (2001) Palindromic but not G-rich sequences are targets of class switch recombination. Int Immunol 13:495–505.

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.

Teng G, Papavasiliou FN (2007) Immunoglobulin Somatic Hypermutation. Annu Rev Genet 41:107–120.

Thomas D, Guthridge M, Woodcock J, Lopez A (2005) 14-3-3 protein signaling in development and growth factor responses. Curr Top Dev Biol 67:285–303.

Thomas-Claudepierre A-S, Schiavo E, Heyer V, Fournier M, Page A, Robert I, Reina-San-Martin B (2013) The cohesin complex regulates immunoglobulin class switch recombination. J Exp Med 210:2495–2502.

Thorson JA, Yu LW, Hsu AL, Shih NY, Graves PR, Tanner JW, Allen PM, Piwnica-Worms H, Shaw AS (1998) 14-3-3 proteins are required for maintenance of Raf-1 phosphorylation and kinase activity. Mol Cell Biol 18:5229–5238.

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

Tumas-Brundage K, Manser T (1997) The transcriptional promoter regulates hypermutation of the antibody heavy chain locus. J Exp Med 185:239–250.

the abundance of nuclear activation-induced deaminase. J Exp Med 208:2385–2391.

Victora GD, Nussenzweig MC (2012) Germinal Centers. Annu Rev Immunol 30:429–457.

Vuong BQ, Lee M, Kabir S, Irimia C, Macchiarulo S, McKnight GS, Chaudhuri J (2009) Specific recruitment of protein kinase A to the immunoglobulin locus regulates class-switch recombination. Nat Immunol 10:420–426.

Wang AH, Kruhlak MJ, Wu J, Bertos NR, Vezmar M, Posner BI, Bazett-Jones DP, Yang XJ (2000) Regulation of histone deacetylase 4 by binding of 14-3-3 proteins. Mol Cell Biol 20:6904–6912.

Wang H, Zhang L, Liddington R, Fu H (1998) Mutations in the Hydrophobic Surface of an Amphipathic Groove of 14-3-3ζ Disrupt Its Interaction with Raf-1 Kinase. the journal of biological chemistry 273:16297–16304.

Wang M, Rada C, Neuberger MS (2010) Altering the spectrum of immunoglobulin V gene somatic hypermutation by modifying the active site of AID. J Exp Med 207:141–153.

Wang M, Yang Z, Rada C, Neuberger MS (2009) AID upmutants isolated using a high-throughput screen highlight the immunity/cancer balance limiting DNA deaminase activity. Nat Struct Mol Biol 16:769–776.

Waterman MJ, Stavridi ES, Waterman JL, Halazonetis TD (1998) ATM-dependent activation of p53 involves dephosphorylation and association with 14-3-3 proteins. Nat Genet 19:175–178.

Weber JS, Berry J, Manser T, Claflin JL (1991) Position of the rearranged V kappa and its 5' flanking sequences determines the location of somatic mutations in the J kappa locus. JI 146:3652–3655.

Wei L, Chahwan R, Wang S, Wang X, Pham PT, Goodman MF, Bergman A, Scharff MD, MacCarthy T (2015) Overlapping hotspots in CDRs are critical sites for V region diversification. Proc Natl Acad Sci USA 112:E728–E737.

Willmann KL, Milosevic S, Pauklin S, Schmitz K-M, Rangam G, Simon MT, Maslen S, Skehel M, Robert I, Heyer V, Schiavo E, Reina-San-Martin B, Petersen-Mahrt SK (2012) A role for the RNA pol II-associated PAF complex in AID-induced immune diversification. J Exp Med 209:2099–2111.

Wilson TM, Vaisman A, Martomo SA, Sullivan P, Lan L, Hanaoka F, Yasui A, Woodgate R, Gearhart PJ (2005) MSH2–MSH6 stimulates DNA polymerase η, suggesting a role for A:T mutations in antibody genes. J Exp Med 201:637–645.

Wuerffel R, Wang L, Grigera F, Manis J, Selsing E, Perlot T, Alt FW, Cogné M, Pinaud E, Kenter AL (2007) S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. Immunity 27:711–722.

Xiao B, Smerdon SJ, Jones DH, Dodson GG, Soneji Y, Aitken A, Gamblin SJ (1995) Structure of a 14-3-3 protein and implications for coordination of multiple signalling pathways. Nature 376:188–191.

Xu Z, Fulop Z, Wu G, Pone EJ, Zhang J, Mai T, Thomas LM, Al-Qahtani A, White CA, Park S-

recombination. Nat Struct Mol Biol 17:1124–1135.

Yaffe MB (2002) How do 14-3-3 proteins work?-- Gatekeeper phosphorylation and the molecular anvil hypothesis. FEBS Lett 513:53–57.

Yaffe MB, Rittinger K, Volinia S, Caron PR, Aitken A, Leffers H, Gamblin SJ, Smerdon SJ, Cantley LC (1997) The Structural Basis for 14-3-3:Phosphopeptide Binding Specificity. Cell 91:961–971.

Yang J, Winkler K, Yoshida M, Kornbluth S (1999) Maintenance of G2 arrest in the Xenopus oocyte: a role for 14-3-3-mediated inhibition of Cdc25 nuclear import. The EMBO Journal 18:2174–2183.

Yeap L-S, Hwang JK, Du Z, Meyers RM, Meng F-L, Jakubauskaitė A, Liu M, Mani V, Neuberg D, Kepler TB, Wang JH, Alt FW (2015) Sequence-Intrinsic Mechanisms that Target AID Mutational Outcomes on Antibody Genes. Cell 163:1124–1137.

Zha J, Harada H, Yang E, Jockel J, Korsmeyer SJ (1996) Serine Phosphorylation of Death Agonist BAD in Response to Survival Factor Results in Binding to 14-3-3 Not BCL-XL. Cell 87:619–628.

Zhang L, Wang H, Liu D, Liddington R, Fu H (1997a) Raf-1 kinase and exoenzyme S interact with 14-3-3zeta through a common site involving lysine 49. the journal of biological chemistry 272:13717–13724.

Yu K1, Huang FT, Lieber MR. DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine. (2004). J. Biol. Chem. 279(8), 6496–6500.

Zhang SH, Kobayashi R, Graves PR, Piwnica-Worms H, Tonks NK (1997b) Serine phosphorylation-dependent association of the band 4.1-related protein-tyrosine phosphatase PTPH1 with 14-3-3beta protein. the journal of biological chemistry 272:27281–27287.

Zhou C, Saxon A, Zhang K (2003) Human activation-induced cytidine deaminase is induced by IL-4 and negatively regulated by CD45: implication of CD45 as a Janus kinase phosphatase in antibody diversification. JI 170:1887–1893.