



**HAL**  
open science

# Mécanismes impliqués dans la régulation de l'enzyme AID et de la réparation des dommages qu'elle induit à l'ADN

Jacques Moritz

► **To cite this version:**

Jacques Moritz. Mécanismes impliqués dans la régulation de l'enzyme AID et de la réparation des dommages qu'elle induit à l'ADN. Médecine humaine et pathologie. Université de Strasbourg, 2020. Français. NNT : 2020STRAJ107 . tel-04416342

**HAL Id: tel-04416342**

**<https://theses.hal.science/tel-04416342>**

Submitted on 25 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

École doctorale des sciences de la vie et de la santé  
Institut de Génétique et de Biologie Moléculaire et Cellulaire

**THÈSE** présentée par :  
**Jacques MORITZ**

soutenue le : **15 décembre 2020**

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**  
Discipline/ Spécialité : Aspects moléculaires et cellulaires de la biologie  
Immunologie

**Mécanismes impliqués dans la régulation  
de l'enzyme AID et de la réparation  
des dommages qu'elle induit à l'ADN**

**THÈSE dirigée par :**

**Dr. Bernardo REINA-SAN-MARTIN**

DR2, IGBMC, Université de Strasbourg

**RAPPORTEURS :**

**Pr. Karin TARTE**

PUPH, MICMAC, Rennes

**Dr. Ludovic DERIANO**

DR, Institut Pasteur, Paris

**AUTRES MEMBRES DU JURY :**

**Dr. Françoise DANTZER**

DR2, ESBS, Université de Strasbourg



## Remerciements

Je tiens tout d'abord à remercier chaleureusement les membres de mon jury de thèse, Karin Tarte, Ludovic Deriano et Françoise Dantzer. Merci d'avoir évalué mon travail.

Je remercie également Bernardo Reina-San-Martin, mon directeur de thèse, de m'avoir fait confiance en me prenant dans son labo. Tu as su être présent, à l'écoute et prodigueur de bons conseils quand il le fallait, tout en me laissant gagner de plus en plus d'autonomie avec le temps. Encore merci à toi.

Je remercie Stéphane Vincent et Ludovic Deriano pour avoir accepté de faire partie de mon jury de mi-thèse et pour m'avoir conseillé.

Mélanie, chaque jour qui passe, je mesure un peu plus la chance inouïe de t'avoir à mes côtés. Sérieuse et appliquée mais surtout dévouée pour ton travail et pour les autres, tu es la partenaire idéale. Aucun mot n'est assez fort pour te dire à quel point je te suis reconnaissant. C'est toi la meilleure !

Isa, le labo n'est plus tout à fait le même sans toi. Je me souviens encore de tous ces moments passés ensemble, de toute cette complicité et surtout de ta capacité à écouter et à donner les bons conseils. Tu es vraiment une personne en or.

Merci Arthur pour tous les conseils que tu m'as donné pour la fin de cette thèse. Même si nous avons passé relativement peu de temps ensemble, j'ai tout de suite senti ta gentillesse et ta franchise.

Merci à toi Vincent pour tous les bons moments de rigolade, pour ta fameuse galette et tes petits chocolats !

Merci Rocio pour ta bonne humeur et les bons moments passés ensemble.

Merci Léa pour m'avoir initié au merveilleux monde de la protéomique ! Tu as toujours été là pour m'aider quand il le fallait et répondre à toutes mes questions du début de ma thèse. Il est loin maintenant le temps où j'étais cantonné à un coin de ta paillasse !

Merci Ebe pour ton chaleureux accueil en Italie. Vivement qu'on puisse revenir te voir !

Je remercie également tous les stagiaires que j'ai encadré pour leur bonne humeur et pour l'aide sur les différents projets auxquels je les ai assignés, leur bonne humeur. Merci donc à Agata, Brenda, Senoussi, Martin, Diana et Zakaria.

Merci aux membres du labo Soutoglou pour tous les moments mémorables passés en salle culture, et en dehors. Mention spéciale à Duygu pour sa sincérité et sa gentillesse.

Merci au labo Ricci et tout particulièrement à Karl et Eric qui me rendent la vie agréable grâce à leur humour dévastateur et si particulier. Kévin, tu m'as été d'un grand soutien ces derniers temps et je t'en remercie.



Merci à tous les membres du labo Tomassetto pour leur bonne humeur et leur joie de vivre !  
Vivement la prochaine sortie dans un bar !

Un grand merci également à Claudine et Muriel, et pour toutes les heures en plus qu'on vous a fait rester ! Grâce à vous, c'est toujours un plaisir de faire de la cytométrie en flux, même quand les résultats ne suivent pas.

Merci également à Luc, Bastien et Frank, de vous être occupé de mes runs de mass spec, et d'avoir pris le temps de répondre à mes questions.

Je souhaite également remercier le ministère de l'enseignement supérieur et de la recherche, ainsi que la fondation ARC pour avoir financé ma thèse.

Un énorme merci à Paul et à Jean-Nicolas. Vous avez déserté le chemin que j'ai gravi aujourd'hui, quel effet ça vous fait ? Vivement qu'on puisse se retrouver comme avant, autour d'un jeu de société ou en rando.

Je remercie également toute ma belle-famille, Jacky, Marcelle, Sam, Flo, Jehane, Fanny et les autres. On passe toujours des moments très agréables en votre compagnie. Un merci tout particulier à Jacky et Marcelle pour leur gentillesse inégalée.

Enfin, un énorme merci à toute ma famille. Maman, Papa, Hélène, Thomas, Papi, Mamie et Seb. Si j'en suis là aujourd'hui, c'est forcément grâce à vous. Votre soutien est sans faille. Ce paragraphe ne vous rend pas hommage comme il se doit, mais je pense que vous savez ce que vous représentez à mes yeux.

# Table des matières

Liste des figures.....	1
Liste des abréviations.....	3
Préambule.....	6
<b>Introduction.....</b>	<b>7</b>
<b>I- La diversification du récepteur des cellules B (BCR).....</b>	<b>9</b>
1. Le récepteur des cellules B (BCR).....	9
2. L'assemblage du BCR.....	10
2.1. L'organisation des loci Ig.....	10
2.2. La diversité des Ac indépendante des antigènes.....	11
2.2.1. La recombinaison V(D)J.....	11
3. La diversification dépendante des antigènes.....	15
3.1. Les centres germinatifs.....	16
3.2. L'hypermutation somatique.....	18
3.3. La commutation isotypique.....	18
3.4. La conversion génique.....	21
<b>II- L'enzyme Activation-induced Cytidine Deaminase (AID).....</b>	<b>23</b>
1. La découverte d'AID et sa fonction.....	23
2. La structure d'AID.....	25
3. La régulation d'AID.....	26
3.1. Régulation de la transcription du locus <i>Aicda</i> .....	26
3.2. Régulation du niveau d'ARN messager d'AID.....	30
3.3. Régulation de la localisation d'AID.....	31
3.4. Régulation de l'activité d'AID pendant le cycle cellulaire.....	33
3.5. Modifications post-traductionnelles.....	33
4. Le ciblage spécifique de l'activité d'AID.....	35
4.1. Les motifs de séquence d'ADN.....	35
4.2. Les éléments en <i>cis</i> .....	36
4.3. La transcription.....	37
4.4. L'ARN issu de la transcription des régions S.....	39
4.4.1. Les R-loops.....	39
4.4.2. L'exosome à ARN.....	41
4.1.3. L'importance de l'épissage des ARN.....	41
4.1.4. Les structures G-quadruplex (G4).....	42
5. Le potentiel oncogénique d'AID.....	43
5.1. Les mutations hors-cible.....	44
5.2. Les translocations.....	44
5.3. L'expression ectopique d'AID.....	45
5.4. <i>Kataegis</i> .....	45

6. Le modèle de désamination de l'ADN.....	46
6.1. Uracil-DNA-glycosylase (UNG) et son rôle privilégié dans la CI et l'HMS.....	46
6.2. La génération de mutations induites par AID durant l'HMS.....	51
6.3. La génération de DSBs durant la CI.....	52
7. La réparation des DSBs durant la CI.....	53
7.1. La DDR.....	53
7.2 La réparation des DSBs.....	55
7.2.1. La réparation des DSBs par le C-NHEJ.....	56
7.2.2. La réparation des DSBs par le A-EJ.....	56
<b>Résultats.....</b>	<b>58</b>
<b>Chapitre 1 : Vers l'identification de nouveaux partenaires d'AID.....</b>	<b>59</b>
I- Génération et caractérisation du knock-in de BirA* dans <i>Aicda</i> .....	62
II- Mise en place du système BiOLD dans des cellules B primaires murines.....	67
III- Analyse des résultats de spectrométrie de masse.....	74
IV- Choix de gènes candidats à étudier.....	83
V- Discussion.....	88
BiOLD, une technique relativement nouvelle avec ses avantages et ses limites.....	88
Fonctionnalité des protéines de fusion.....	89
Analyse des différentes listes de protéines enrichies.....	90
AID-BirA* vs. BirA*-AID.....	90
AID-BirA* vs. AID-P2A-BirA*.....	91
AID-BirA* vs. AID <sup>ΔCter</sup> -BirA*.....	91
AID-BirA* +biotine vs. AID-BirA* -biotine.....	92
Conclusions sur le système BiOLD.....	92
Gènes candidats et perspectives.....	93
<b>Chapitre 2 : Etude du rôle de Fam72a dans la diversification des Ig.....</b>	<b>97</b>
I- Fam72a contrôle la balance entre la réparation de l'ADN mutagène et fidèle durant la diversification des anticorps.....	98
Manuscrit en cours de rédaction	
II- Résultats supplémentaires.....	126
1. Reconstitution de l'expression de mFam72a et mFam72a <sup>W125R</sup> dans les cellules Fam72a <sup>-/-</sup> .....	126
2. Reconstitution de l'expression des versions humaines hFAM72A-D dans les cellules Fam72a <sup>-/-</sup> .....	128
3. L'activité d'Ung est-elle modifiée en l'absence de Fam72a ?.....	128
III- Discussion.....	131
1. Ung2 et la commutation isotypique : une affaire de quantité ?.....	131
1.1. Une trop grande quantité d'Ung2 nuit-elle à la CI ?.....	131
1.2. Et une quantité trop faible ?.....	132

2. Fam72a et Ung2 dans la CI : d'autres régulations possibles.....	132
3. Et Ung1 dans tout ça ?.....	133
4. Fam72a et Ung2 dans l'HMS.....	134
5. Perspectives.....	135
5.1 Ung2 est-elle bien surexprimée en absence de Fam72a ?.....	135
5.2 Fam72a a-t-il un autre rôle que celui de dégrader Ung2 ?.....	135
5.3 Est-ce que le recrutement d'AID à S <sub>μ</sub> est altéré en absence de Fam72a ?.....	136
5.4 Est-ce que l'activité d'Ung est altérée en absence de Fam72a ?.....	136
5.5 Y-a-t-il moins d'uracile dans l'ADN en absence de Fam72a ?.....	136
5.6 La dégradation d'Ung2 est -elle dépendante du protéasome ?.....	137
5.7 Y-a-t-il plus ou moins de DSBs non réparées induites par AID en l'absence de Fam72a ?.....	137
5.8 Est-ce que Ung2 est supprimée dans les lignées de cellules cancéreuses humaines surexprimant Fam72a ?.....	137
5.9 Est-ce que la transcription des régions S est affectées dans les cellules B primaires Fam72a <sup>-/-</sup> ?.....	138
6. Conclusion.....	138
<b>Conclusion générale.....</b>	<b>139</b>
<b>Matériels et Méthodes.....</b>	<b>140</b>
<b>Références bibliographiques.....</b>	<b>144</b>

# Liste des figures

Figure 1: Structure du récepteur des cellules B.....	8
Figure 2: Organisation des loci Ig murins.....	8
Figure 3: Les étapes moléculaires clés de la recombinaison V(D)J.....	12
Figure 4: Le développement des cellules B.....	14
Figure 5: Mécanismes de diversification des anticorps.....	14
Figure 6: L'hypermutation somatique.....	19
Figure 7: La commutation isotypique.....	19
Figure 8: La conversion génique.....	22
Figure 9: Les domaines fonctionnels de la protéine Activation-Induced cytidine Deaminase (AID).....	24
Figure 10: Les mécanismes régulateurs d'AID .....	27
Figure 11: Le modèle de désamination de l'ADN par AID.....	47
Figure 12: Les domaines fonctionnels des isoformes de l'enzyme Uracil-DNA-Glycosylase (UNG).....	49
Figure 13: Formation et réparation des cassures double brin (DSBs) durant la CI.....	54
Figure 14: Purification et identification des partenaires protéiques d'AID à l'aide de la technologie BioID.....	61
Figure 15: Génération et criblage de clones CH12 knock-in pour l'expression de Flag-BirA*-AID. ....	63
Figure 16: Caractérisation fonctionnelle des clones CH12 knock-in BirA*-AID. ....	65
Figure 17: Stratégie et caractérisation des clones CH12 knock-in AID-BirA*.....	66
Figure 18: La technologie BioID dans des cellules B primaires de souris. ....	68
Figure 19: Les vecteurs pMX-PIE permettent une transduction rétrovirale efficace des cellules B primaires. ....	70

Figure 20: La déficience de CI des B primaires AID <sup>-/-</sup> peut être sauvée par l'expression d'AID-BirA* . . . . .	71
Figure 21: Les protéines de fusion sont bien exprimées et sont capables d'être purifiées avec leurs partenaires. . . . .	73
Figure 22: Comparaison des résultats obtenus grâce aux différentes protéines de fusion de BiolD. . . . .	75
Figure 23: Analyse statistique de l'enrichissement de protéines entre les conditions AID-BirA* +biotine et AID-BirA* -biotine. . . . .	77
Figure 24: Analyse statistique de l'enrichissement de protéines identifiées dans les conditions AID-BirA* +biotine, AID-P2A-BirA*, BirA*-AID et AID <sup>DCter</sup> -BirA* . . . . .	79-82
Figure 25: Enrichissement statistique de processus biologiques liés à AID des protéines enrichies dans AID-BirA* par rapport aux autres conditions. . . . .	84
Figure 26: Délétion et caractérisation de gènes candidats issus des cribles BiolD et CRISPR/Cas9. . . . .	86
Figure 27 : Reconstitution des cellules Fam72a <sup>-/-</sup> avec les formes souris, mFam72a et mFam72a <sup>W125R</sup> et les formes humaines hFAM72A-D. . . . .	127
Figure 28: Test d'activité d'Ung en absence de Fam72a. . . . .	129

## Liste des abbreviations

3'RR: region régulatrice en 3', 3' regulatory region  
53BP1: p53 binding protein  
Ac: anticorps  
A-EJ : voie alternative du NHEJ  
Ag: antigène  
AID: Activation-induced cytidine deaminase  
APE1: DNA-(apurinic/apyrimidic site) lyase 1  
APE2: DNA-(apurinic/apyrimidic site) lyase 2  
APLF: aprataxin and polynucleotide kinase like factor  
APOBEC: apolipoprotein B mRNA editing enzyme catalytic polypeptide-like  
ARNg: ARN guide  
ARN pol II: ARN Polymérase II  
APEX2: ascorbate peroxidase  
ASCC3: activating signal cointegrator 1 complex subunit  
ATM: ataxia telangiectasia mutated  
ADCC : cytotoxicité à médiation cellulaire dépendante des anticorps, antibody dependant cell-mediated cytotoxicity  
BATF: basic leucine zipper transcriptional factor ATF-like)  
BAFF: B-cell activating factor  
BioID:proximity-dependent biotin identification  
BirA:biotine ligase bactérienne  
BirA\* : biotine ligase mutante (R118G)  
Co-IP: co-immunoprécipitation  
BCL6 : B cell lymphoma 6 protein  
BCR : récepteur des cellules B, B cell receptor  
BER : réparation par excision de base, base excision repair  
BLIMP1: B lymphocyte-induced maturation protein 1  
BRIT1: BRCT-repeat inhibitor of TERT expression 1  
C : région constante  
Cas9 : CRISPR-associated protein-9 nuclease  
CD40L: CD40 Ligand  
CDR: complementarity-determining region  
CEBPs : CCAAT/enhancer-binding proteins  
CI : commutation isotypique  
CG : conversion génique  
C-NHEJ : voie Classique du NHEJ  
CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats  
CRM1: chromosomal maintenance1, exportin 1  
C-ter : extrémité carboxy-terminale  
CtIP : CtBP-interacting protein  
CTNBL1 : catenin Beta-like protein1  
D : segments géniques de diversité  
Dhx29: DExH-box helicase 29  
DBR1 : debranching enzyme1  
DDR : réponse aux dommages de l'ADN, DNA damage response  
DIVACs : diversification activators  
DLBCL : lymphome diffus à grandes cellules, diffuse large B-cell lymphoma  
DNAJA1: DnaH homolog subfamily A member 1

DNA-PKcs: DNA-dependent protein kinase catalytic subunit  
 DSB: cassure double brin à l'ADN, double stranded DNA breaks  
 DSIF: DRB Sensitivity Inducing Factor  
 DZ: zone sombre, dark zone  
 DTA : toxine diphtérique A  
 eEF1A: eukaryotic Elongation Factor 1 alpha  
 EGFP: enhanced green fluorescent protein  
 EXO1: exonuclease 1  
 FACT: facilitates chromatin transcription complex  
 FDC : cellules dendritiques folliculaires, follicular dendritic cells  
 FDR: false discovery rate  
 G4: G-quadruplex  
 GANP: Germinal center associated nuclear Protein  
 GCs: centre germinatifs, germinal centers  
 GLT: germline transcript  
 GO : ontologie génique, gene ontology  
 HMGB: high-mobility group box  
 HMCES: 5-Hydroxymethylcytosine binding, ES-cell-specific  
 HMS: hypermutation somatique  
 HMW: high molecular weight  
 hnRNPs: heterogeneous nuclear ribonucleoproteins  
 HR: homologous recombination  
 HSP90: heat shock protein-90  
 HSP40: heat shock protein-40  
 ICOSL: inducible T-cell costimulatory ligand  
 Id2: inhibitor of differentiation 2  
 Ig: immunoglobuline  
 IF : immunofluorescence  
 IgH : chaîne lourde des immunoglobulines, immunoglobulin heavy chain  
 IgL : chaîne légère des immunoglobulines, immunoglobulin light chain  
 IL-4: interleukin-4  
 IRF4: interferon regulatory factor 4  
 ITAM: immunoreceptor tyrosine-base activation motifs  
 J : segments géniques de jonction  
 Jmjd1c: Jumonji Domain Containing 1C  
 KAP1: KRAB domain-associated protein 1  
 LMW: light molecular weight  
 LPS: lipopolysaccharide  
 LZ: zone claire, light zone  
 MBD4: methyl-CpG binding domain 4  
 MDC1: mediator of DNA damage checkpoint protein 1  
 MLH1: MutL homolog 1  
 MHC : complexe majeur d'histocompatibilité  
 MS: mass spectrometry  
 MMP: mitochondrial processing peptidase  
 MMR : réparation des mésappariements, mismatch repair  
 MRN: Mre11, Rad50n Nbs1  
 MSH2: MutS protein homolog 2  
 MSH6: MutS protein homolog 6  
 Nbs1: nijmegen breakage syndrome protein 1  
 NES: signal d'export nucléaire nuclear export signal  
 NF-kB: nuclear factor NF-kappa-B



NHEJ : religation d'extrémités non-homologues, non-homologous end-joining  
 NK : natural killer  
 NLS : signal de localisation nucléaire, nuclear localization signal  
 N-ter : extrémité NH<sub>2</sub>-terminale  
 P2A: porcine teschovirus-1 2A  
 PAF: RNAP II-associated factor complex  
 PCNA: proliferating cell nuclear antigen  
 PCA : analyse par composante principale, principal component analysis  
 PCNA-Ub : ubiquitination spécifique de PCNA à la lysine 164  
 PKA: protein kinase A  
 PMS2: postmeiotic segregation increased 2  
 PP2A : protein phosphatase 2  
 Pre-BCR : précurseur du BCR  
 PTBP2: polypyrimidine tract-binding protein2  
 RAG: recombination-activating gene  
 Regg: proteasome activator complex subunit 3  
 RIF1: rap1-interacting factor 1 homolog  
 RNF: ring finger protein  
 RPA : Replication Protein A  
 RSS : séquences de signal de recombinaison, recombination signal sequences  
 S: région S, switch region  
 Spt5: suppressor of Ty homolog 5  
 SSBs: single stranded DNA breaks  
 SMUG1: single-strand-selective monofunctional uracil-DNA glycosylase 1  
 Stat6: signal transducer and activator of transcription 6  
 TAD: topologically associating domain  
 TCR: T cell receptor  
 TdT: terminal deoxynucleotidyl transferase  
 TDG: thymine-DNA glycosylase  
 TET: ten-eleven translocation  
 Tfh : cellules T auxiliaires folliculaire, T follicular helper  
 TGF-b: transforming growth factor b  
 TLR: Toll-like receptor  
 TSS : site d'initiation de la transcription, transcription start site  
 UNG: uracil-DNA glycosylase  
 V: région variable  
 WT: wild-type  
 WB: western blot  
 XBP-1: X-box binding protein 1  
 XLF: XRCC4-like factor  
 XRCC4: X-ray repair cross-complementing protein  
 YY1: Yin Yang 1

## Préambule

Chaque jour, notre corps doit faire face à des millions de microorganismes (bactéries, virus, parasites ou encore champignons), et dont certains peuvent entraîner des maladies. Afin de lutter contre ces dangers aussi divers que variés, nous avons hérité de systèmes de défense qui ont évolué depuis des millions d'années pour devenir de plus en plus sophistiqués et efficaces. Ces systèmes de défense sont regroupés sous le terme de système immunitaire qui est composé d'une grande variété de cellules et de molécules. Comme tous les autres mammifères, notre système immunitaire peut-être divisé en 2 parties : le système immunitaire inné et le système immunitaire adaptatif. Le système immunitaire inné, le plus primordial, permet de nous défendre efficacement contre la plupart des corps étrangers que nous rencontrons quotidiennement. C'est le premier à se mettre en place ; sa réponse, bien que spécifique, ne reconnaît en revanche que des domaines invariants des agents pathogènes.

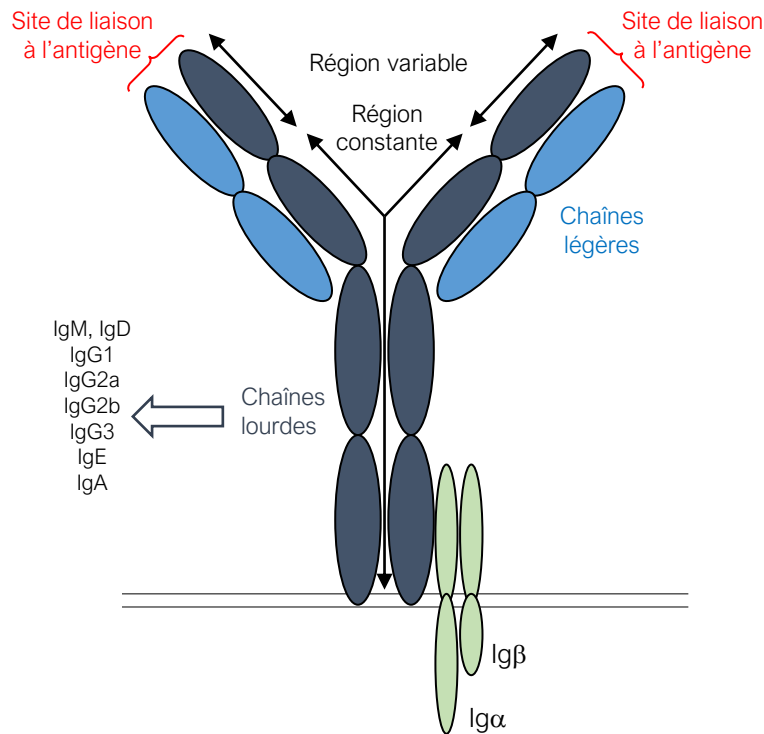
Lorsqu'il ne suffit plus, le système immunitaire inné passe le relais à l'adaptatif, constitué des lymphocytes T et B. Alors qu'une partie des cellules T (CD4+) orchestrent la réponse immunitaire, l'autre partie (CD8+) est capable de détruire des cellules contenant le pathogène ou des cellules cancéreuses. Les lymphocytes B ont pour rôle de produire des anticorps, capables de reconnaître, neutraliser et d'aider à détruire le pathogène.

Toutes ces cellules ont l'unique particularité de posséder chacune à leur surface un récepteur spécialisé avec une spécificité en théorie unique. L'ensemble de toutes les spécificités portées par toutes ces cellules peuvent donc potentiellement reconnaître et lutter contre n'importe quel corps étranger, avant même que celui-ci ne soit rencontré.

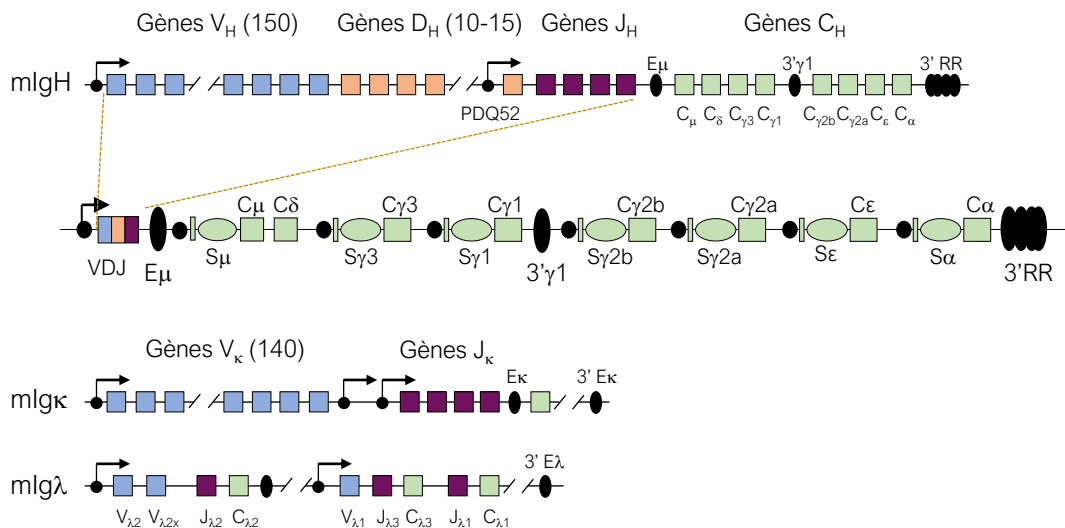
Dans le même temps, des lymphocytes T et B mémoires vont être générés. Ces cellules, qui sont l'autre particularité du système immunitaire adaptatif, permettent une réponse plus rapide et plus efficace en cas de réinfection par le même pathogène. De plus, lors d'une infection, les lymphocytes B sont capables de diversifier à nouveau l'anticorps qu'ils produisent afin de les rendre plus affins et plus efficaces, afin de générer une réponse hautement spécifique contre tout pathogène. Ces mécanismes étant dépendants de l'introduction de mutations dans l'ADN, ils sont potentiellement dangereux pour les cellules B s'ils sont mal régulés. Ainsi, de nombreux mécanismes de régulation existent pour protéger les cellules B contre ce potentiel oncogénique, tout en permettant la bonne diversification des anticorps qu'ils produisent.

C'est à ces mécanismes, dont beaucoup sont encore mal compris, que mon projet de thèse s'intéresse.

# INTRODUCTION



**Figure 1 : Structure du récepteur des cellules B.** Le récepteur des cellules B (appelé BCR) est constitué de deux chaînes lourdes et de deux chaînes légères comportant chacune une région variable et une région constante. Alors que l'association des régions variables d'une chaîne lourde et d'une chaîne légère forme le site de liaison de l'antigène, la région constante des chaînes lourdes détermine l'isotype et donc la fonction effectrice de l'anticorps. Le BCR est exprimé à la surface des cellules B et représente la version membranaire d'un anticorps. Il s'associe à l'hétérodimère  $Ig\alpha/Ig\beta$  pour permettre la transduction du signal suite à la reconnaissance d'un antigène, ce qui constitue son rôle principal.



**Figure 2 : Organisation des loci Ig murins.** Représentation schématique du locus IgH murin avant et après le mécanisme de recombinaison V(D)J, ainsi que des loci  $Ig\kappa$  et  $Ig\lambda$  murins. Le locus IgH murin ( $mIgH$ ) est composé de segments de gènes Variables (V; en bleu), de Diversité (D; orange), et de Jonction (J; violet) qui codent pour la région variable de l'Ig. Ces segments sont suivis par des unités transcriptionnelles codant pour les différents isotypes possibles, et formé notamment d'une région de switch ( $S_x$ ) et de l'exon codant pour l'isotype associé ( $C_x$ ). Les loci codant pour les chaînes légères ( $IgL$  ;  $Ig\kappa$  et  $Ig\lambda$ ) contiennent les gènes V, J et C qui sont organisés par famille de segments pour la chaîne  $\kappa$  et en cassettes distinctes pour la chaîne  $\lambda$ . Le nombre de segments géniques contenus dans chaque famille est indiqué entre parenthèses. Les promoteurs sont représentés en points noirs et les enhancers en formes ovales noires. Adapté de Cobb et al. 2006 et Schatz et al., 2011.

# I- La diversification du récepteur des cellules B

## 1. Le récepteur des cellules B et les anticorps

Les cellules B ont la particularité d'exprimer à leur surface un type de récepteur spécifique à la lignée B et capable de reconnaître un antigène (Ag), appelé récepteur des cellules B (BCR, de l'anglais B cell receptor) (**Figure 1**). Chaque cellule B n'exprime qu'une seule version du BCR qui lui est propre, lui conférant ainsi une spécificité en théorie unique. Le BCR ne diffère d'un anticorps (Ac) que par l'expression d'un petit domaine transmembranaire permettant de l'attacher à la surface d'un lymphocyte B. Le BCR est donc membranaire, alors qu'un Ac est sécrété par les lymphocytes B. Les Ac font partie de la superfamille des immunoglobulines (Ig) et sont constitués de deux chaînes légères et de deux chaînes lourdes identiques entre elles. Chez l'homme et la souris, les chaînes légères peuvent être soit de type  $\kappa$ , soit de type  $\lambda$ , et sont encodées par les loci IgK et IgL respectivement. Les chaînes lourdes sont encodées par un locus unique appelé IgH. Chaque chaîne lourde est liée de manière covalente à une chaîne légère et à l'autre chaîne lourde par des ponts disulfures. De plus, chaque chaîne (lourde ou légère) est constituée d'une région variable (V) et d'une région constante (C). Chaque région variable est constituée de 3 domaines hypervariables aussi appelés régions de détermination de la complémentarité (CDRs, de l'anglais complementarity-determining regions). Ensemble, les régions V d'une chaîne légère et d'une chaîne lourde s'associent pour former le site de reconnaissance à l'antigène, aussi appelé paratope. La région C des chaînes lourdes détermine, elle, l'isotype et donc les fonctions effectrices de l'anticorps, et notamment comment l'anticorps interagit avec d'autres cellules du système immunitaire pour éliminer l'Ag reconnu.

Enfin, le BCR est associé à un hétérodimère constitué d'Ig $\alpha$  et d'Ig $\beta$ , deux sous-unités possédant des « immunoreceptor tyrosine-based activation motifs » (ITAM). Ces motifs sont nécessaires à la transduction d'un signal intracellulaire permettant d'activer la cellule B lorsque le BCR reconnaît spécifiquement son Ag. Le BCR a également pour rôle d'internaliser l'Ag reconnu afin qu'il puisse être présenté aux cellules T helper grâce au complexe majeur d'histocompatibilité (MHC) de classe II.

Les Ac peuvent jouer plusieurs rôles primordiaux dans l'immunité : la neutralisation, afin d'empêcher les pathogènes d'entrer dans des cellules mais aussi de bloquer des toxines. Ils permettent également l'opsonisation, qui facilite la phagocytose de l'Ag reconnu, et la cytotoxicité à médiation cellulaire dépendante des anticorps (ADCC ; de l'anglais antibody dependant cell-mediated cytotoxicity) qui, similairement à l'opsonisation, facilite la lyse de cellules marquées par

des Ac. Les Ac peuvent également permettre l'activation du complément, un autre acteur moléculaire de l'immunité pouvant lyser des cellules.

## 2. L'assemblage du BCR

### 2.1. L'organisation génomique des loci Ig

Comme cela a été mentionné dans le paragraphe précédent, le BCR peut être encodé par 3 loci différents, IgK, IgL et IgH, regroupés sous le terme de loci Ig. Je détaillerai par la suite la structure des loci Ig murins puisque c'est sur ce modèle que se portent mes travaux.

Le locus IgH est situé sur le chromosome 12 et s'étend sur près de 3 mégabases (Mb) ([Chevallard et al., 2002](#)). Il est tout d'abord constitué, en 5', par 3 familles de segments géniques appelés gènes variables (V), gènes de diversité (D) et gènes de jonction (J), qui, une fois réassemblés, permettent de coder pour la région variable (**Figure 2**). On retrouve ainsi près de 150 segments V, 10 à 15 segments D et 4 segments J. Ceux-ci sont suivis par les exons C codant pour la région constante de l'anticorps. Il existe 8 exons C, à savoir dans l'ordre : C $\mu$ , C $\delta$ , C $\gamma$ 3, C $\gamma$ 1, C $\gamma$ 2b, C $\gamma$ 2a, C $\epsilon$  et C $\alpha$  codant respectivement pour les isotypes IgM, IgD, IgG3, IgG1, IgG2b, IgG2a, IgE et IgA. Chacun de ces exons (hormis C $\delta$ ) est compris au sein d'une unité transcriptionnelle constituée d'un promoteur sensible aux cytokines suivi d'un exon intermédiaire (I), d'une région dite de « switch » (S) puis de l'exon C. L'exon C $\delta$  est lui situé juste en aval de C $\mu$ , et son expression est permise par épissage alternatif.

Le locus IgH possède de plus 3 principaux éléments régulateurs : le promoteur/enhancer PDQ52 en 3' du dernier segment D, l'enhancer E $\mu$  situé dans la région intronique entre les segments J et C $\mu$ , et le super-enhancer (SE) nommé région régulatrice 3' (3'RR) localisé à l'extrémité 3' du locus et caractérisé par la présence de 7 sites hypersensibles à la DNaseI.

Les loci IgK et IgL sont situés sur les chromosomes 6 et 16 respectivement. Ils codent donc pour la chaîne légère des Ac et à ce titre, sont composés de segments géniques variables (V) et de jonction (J), codant pour la région variable suivis d'un exon C, codant pour la région constante. Ainsi, le locus IgK est constitué de 140 segments V et 4 segments J fonctionnels suivis par un unique exon C, alors que le locus IgL est constitué par 3 cassettes distinctes de segments V et J suivis de leur propre exon C ([Cobb et al., 2006](#)) (**Figure 2**).

L'organisation des loci Ig humains est semblable à celle des loci Ig murins mais possède néanmoins ses particularités : outre un nombre légèrement différent de segments géniques dans les loci IgK et IgV, les isotypes encodés par le locus IgH sont plus nombreux dûs à un événement

de duplication. En effet, chez l'humain, on retrouve dans l'ordre :  $C_{\mu}$ ,  $C_{\delta}$ ,  $C_{\gamma 3}$ ,  $C_{\gamma 1}$ , le pseudo-gène  $\psi C_{\epsilon}$ ,  $C_{\alpha 1}$ , le pseudo-gène  $\psi C_{\gamma}$ ,  $C_{\gamma 2}$ ,  $C_{\gamma 4}$ ,  $C_{\epsilon}$  et enfin  $C_{\alpha 2}$ . Du fait de cette duplication, on retrouve également deux 3'RR, le premier étant situé entre  $C_{\alpha 1}$  et  $C_{\phi\gamma}$ , et le deuxième à l'extrémité 3' du locus ([D'Addabbo et al., 2011](#)).

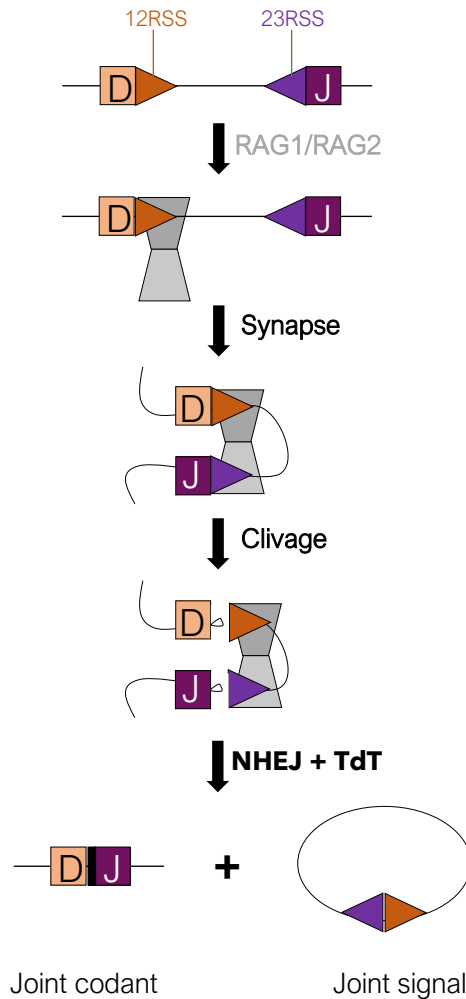
## 2.2. La diversification des anticorps indépendante des antigènes

L'ensemble des spécificités des BCRs exprimés par tous les lymphocytes B d'un individu, aussi appelé répertoire des BCRs, est estimé à  $10^{11}$ . Cette gigantesque diversité permet à un individu de reconnaître virtuellement n'importe quel pathogène. On peut alors se demander comment un tel niveau de diversité est atteint, alors même que ce nombre est bien supérieur au nombre de gènes présents dans un individu ? Le groupe de S. Tonegawa commença à résoudre cette énigme dès 1976, en observant que des fragments géniques sont recombinaisonnés au sein des loci Ig pour former une séquence continue dans les lymphocytes B matures mais pas dans les cellules embryonnaires ([Hozumi and Tonegawa, 1976](#)). Cette étude pose ainsi les bases du mécanisme appelé recombinaison V(D)J, capable de générer le répertoire primaire d'Ac de façon indépendante de la rencontre d'antigènes.

### 2.2.1. La recombinaison V(D)J

La recombinaison V(D)J est le mécanisme qui permet d'assembler des fragments géniques de la région variable des loci Ig afin de générer un récepteur fonctionnel mais aussi un répertoire diversifié de récepteurs antigéniques. Ce processus est donc nécessaire pour l'assemblage du BCR, mais également de son homologue trouvé chez les lymphocytes T, le TCR. Il a lieu lors du développement des lymphocytes B dans la moëlle osseuse, et n'est donc pas dépendant de la reconnaissance d'un antigène.

Cet assemblage va être le résultat d'une recombinaison aléatoire entre un segment V, un segment D et un segment J dans le locus IgH, et entre un segment V et un segment J dans les loci IgK et IgL. La recombinaison V(D)J est initiée par des cassures double brin (DSBs de l'anglais Double-Strand Breaks) à l'ADN générées par l'hétérodimère recombinaison-activating gene (RAG)1 et RAG2 à des séquences de signal de recombinaison (RSS ; de l'anglais recombination signal sequences) ([Oettinger et al., 1990](#); [Schatz et al., 1989](#)). Les RSS sont constitués d'une séquence héptamérique et d'une séquence nonamérique bien conservées et séparées par un espaceur moins bien conservé de 12 ou de 23 nucléotides ([Sakano et al., 1981](#)). Il existe donc 2 types de RSS appelées 12RSS et 23RSS en fonction de la longueur de cet espaceur. La recombinaison



**Figure 3 : Les étapes moléculaires clés de la recombinaison V(D)J.**

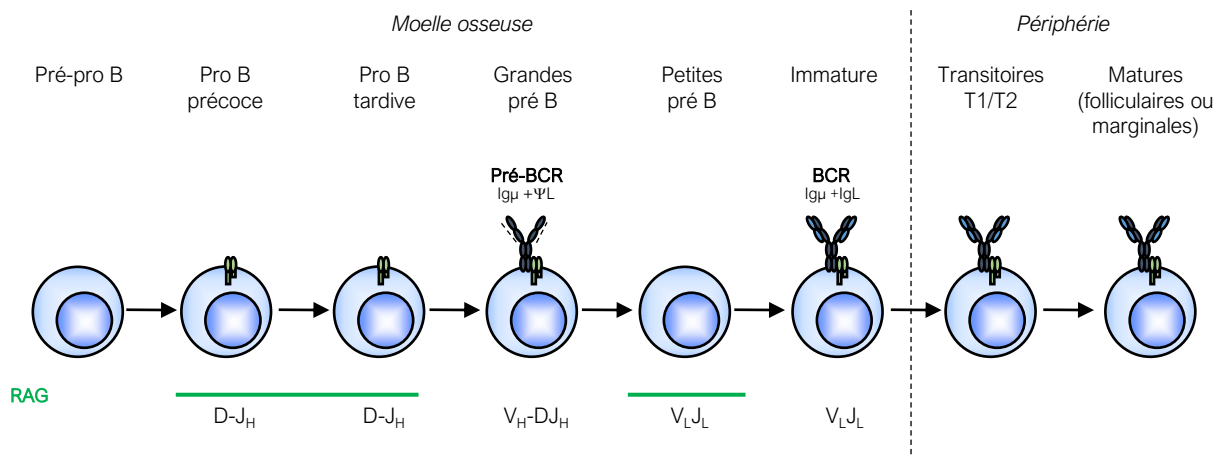
Les segments de gènes D et J sont bordés par des séquences signal de recombinaison (RSS) 12RSS et 23RSS, respectivement. Le complexe RAG formé de Rag1 et Rag2 (trapèzes gris) qui interagit avec des protéines de liaison à l'ADN va reconnaître une première RSS, puis capturer la seconde RSS selon la règle 12/23 et former une synapse. RAG génère une cassure double brin entre les segments géniques et les RSS. Ensuite, RAG participe à la réparation de ces cassures doubles brin avec les facteurs de la voie des jonctions non-homologues (NHEJ) pour réaliser la recombinaison. Le joint codant contient des nucléotides supplémentaires (rectangle noir) ajoutés par la polymérase terminale transférase (TdT), tandis que le joint signal est simplement ligué afin de former un épisode qui est éliminé du génome. Adapté de Ru, Zhang et al. 2018.



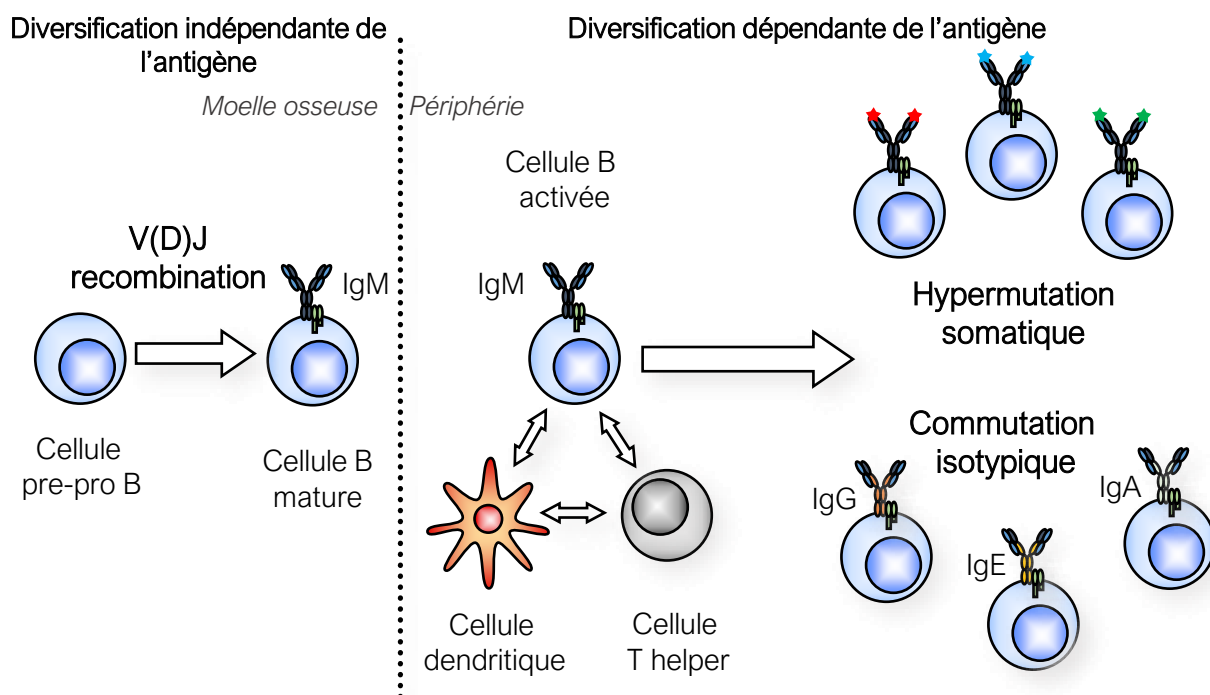
entre séquences RSS n'est possible que si elle suit une règle dite « 12/23 », c'est-à-dire entre un 12RSS et un 23RSS ([Akira et al., 1987](#); [Tonegawa, 1983](#)). En effet, dans le locus IgH, tous les segments V possèdent à leur bordure 3' un 23RSS alors que tous les segments D possèdent en 5' et en 3' un 12RSS alors que les segments J ont un 23RSS en 5'. Cette organisation structurale et la règle 12/23 permettent d'assurer que la recombinaison se fait entre des fragments géniques différents et dans la bonne orientation.

D'un point de vue mécanistique, la recombinaison V(D)J est un processus à plusieurs étapes qui est initié durant la phase G1 par le complexe RAG aidé par la protéine High-mobility group box (HMGB) 1 ([Ciubotaru et al., 2013](#)). RAG1 reconnaît tout d'abord une séquence RSS (12 ou 23) puis va générer une cassure simple brin précisément à la jonction entre la séquence codante et la RSS (**Figure 3**). Une RSS complémentaire (selon la règle 12/23) est ensuite « capturée » pour former un complexe apparié où cette RSS partenaire subit également une cassure simple brin. Ces cassures permettent de libérer un groupement 3'-hydroxyle qui attaque le brin opposé par une réaction de transestérification catalysée par le complexe RAG. Cela va alors générer deux structures en épingle à cheveux (une par fragment) appelées « coding ends » (CEs) d'une part, et des « signal ends » (SEs) avec des extrémités franches et contenant les 2 RSS de l'autre. Les structures en épingle à cheveux aussi bien que les extrémités franches seront réparées par la voie classique de jonction des extrémités non homologues (c-NHEJ ; de l'anglais classical Non-Homologous End Joining) afin d'ouvrir les structures en épingle à cheveux d'une part et de liguer les RSS entre eux d'autre part. Les RSS ainsi que la séquence qui les sépare seront alors éliminés du génome sous forme d'un épisode circulaire. Parmi les acteurs moléculaires impliqués dans cette étape, on retrouve la kinase ataxia telangiectasia mutated (ATM), les protéines Ku70 et Ku80 qui recrutent la sous-unité catalytique de DNA-dépendant protein kinase (DNA-PKcs), qui va alors phosphoryler Artemis, qui est responsable de l'ouverture des structures en épingle à cheveux ([Ma et al., 2002](#)). Une fois ouverte, Artemis peut également ajouter une série de nucléotides palindromiques appelés nucléotides P, mais est également capable de retirer quelques nucléotides de manière aléatoire ([Lafaille et al., 1989](#); [Ma et al., 2002](#)). La polymérase terminal-deoxynucleotidyltransferase (TdT) permet elle aussi d'augmenter la diversité générée par la recombinaison V(D)J en ajoutant des nucléotides aléatoires, nommés nucléotides N, aux extrémités des DSBs. Enfin, la réparation finale des DSBs se fait grâce à l'action coordonnée de X-ray repair cross-complementing protein 4 (XRCC4), XRCC4-like factor (XLF) et de l'ADN ligase IV ([Lescale and Deriano, 2017](#)).

La recombinaison V(D)J dicte les différents stades de développement des lymphocytes B au sein de la moelle osseuse. Ainsi, les premiers réarrangements se font au sein du locus IgH, entre un



**Figure 4 : Le développement des cellules B.** Les premiers précurseurs de cellules B engagés sont des cellules pré-pro-B. Le complexe RAG est exprimé dans les cellules pro-B et provoque la recombinaison de la chaîne lourde IgH, en démarrant par la jonction D-J<sub>H</sub> puis V<sub>H</sub>-DJ<sub>H</sub>. Le pré-BCR est exprimé à la surface des cellules pré-B, et est composé d'une chaîne lourde Ig<sub>μ</sub> qui a eu un réarrangement productif ainsi que d'une chaîne légère de substitution ΨL. Par la suite, ces cellules recombinent le locus de la chaîne légère V<sub>L</sub>J<sub>L</sub> pour exprimer un BCR fonctionnel au stade immature. Les cellules B vont ensuite migrer en périphérie où elle poursuivent leur maturation en cellules B folliculaires ou marginales. Adapté de Meffre et al. 2000.



**Figure 5 : Mécanismes de diversification des anticorps.** La génération du répertoire primaire des anticorps se produit dans la moelle osseuse, lors du développement des cellules B, grâce au mécanisme de recombinaison V(D)J. Les cellules B matures exprimant un BCR fonctionnel à l'issue de la recombinaison V(D)J migrent ensuite en périphérie. Suite à la reconnaissance de l'antigène appaïté et à l'activation médiée par les cellules T et par les cellules dendritiques, elles vont se diriger dans les organes lymphoïdes secondaires où elles vont diversifier leur répertoire d'anticorps par les mécanismes d'hypermutation somatique et de commutation isotypique.

fragment D et un fragment J au stade pro-B précoce, puisse entre un fragment V et le segment D-J réarrangé au stade pro-B tardif (Figure 4). Ce premier réarrangement, s'il est productif, permet la production d'une chaîne lourde, qui, associée avec une chaîne légère de substitution non arrangée, va permettre l'expression d'un pré-BCR (Melchers et al., 2000; Rajewsky, 1996). Ce pré-BCR va permettre la transduction d'un signal permettant qu'un réarrangement dans le locus IgL ou IgK entre un segment V et un segment J puisse se produire au stade petites cellules pré-B, afin d'aboutir à l'expression d'un BCR mature (Cobb et al., 2006). Ce processus est finement contrôlé par plusieurs mécanismes qui limitent l'expression de RAG1/2 à des stades précis du développement des lymphocytes, et qui permettent la bonne accessibilité aux RSS ainsi que le rapprochement de séquences éloignées dans le génome en modifiant l'architecture tridimensionnelle du locus en jeu pour pouvoir promouvoir une recombinaison efficace.

De par sa nature à générer des BCRs très diversifiés, la recombinaison V(D)J peut entraîner l'expression d'un BCR capable de reconnaître un antigène du soi, aussi appelé auto-antigène. Les cellules B portant des BCRs auto-réactifs sont potentiellement dangereuses pour l'organisme et sont donc la cible des mécanismes de la tolérance centrale. Le premier mécanisme, l'édition du récepteur, est le phénomène qui permet à un BCR auto-réactif de pouvoir être à nouveau réarrangé (avec les segments géniques restants) afin de tenter de le rendre non auto-réactif (Luning Prak et al., 2011). En cas d'échec, c'est l'autre allèle qui peut alors réaliser la recombinaison V(D)J. Si cet autre allèle a déjà été réarrangé, la cellule B va alors mourir par apoptose grâce au mécanisme de délétion clonale (Meffre and Wardemann, 2008; Nemazee, 2017; Rajewsky, 1996). Des défauts dans la tolérance centrale sont impliqués dans l'apparition de maladies auto-immunes et de certains désordres d'immunodéficience.

Il est considéré que la recombinaison V(D)J est en théorie capable de générer près de  $10^{13}$  BCRs différents (Calis and Rosenberg, 2014), et qu'un être humain n'exprime que 2% de cette diversité possible à un moment donné (Trepel, 1974).

### 3. La diversification des anticorps dépendante des antigènes.

Suite à la recombinaison V(D)J, les cellules B matures naïves exprimant un BCR fonctionnel et ne reconnaissant pas de molécules du soi vont migrer de la moelle osseuse vers la périphérie, et seront principalement localisées dans les organes lymphoïdes secondaires comme la rate et les ganglions (Figure 5). Là, elles pourront alors être activées par la reconnaissance d'un Ag présenté par des cellules dendritiques, et vont pouvoir alors l'internaliser et le présenter à des cellules T folliculaires helper (Tfh) grâce à leur complexe majeur d'histocompatibilité (CMH) de type II. Ces

Tfh ont ainsi l'opportunité de reconnaître l'antigène grâce à leur TCR et vont en retour permettre l'activation complète des lymphocytes B. Différents mécanismes et signaux sont mis en jeu lors de cette activation, et l'on peut ainsi citer les molécules de co-stimulation telles que CD40-CD40L, CD80/CD86 et CD28, ICOSL-ICOS, PD1L et PD1, et enfin l'action de l'IL-21, produite par les Tfh et qui joue également un rôle majeur dans l'activation des cellules B ([Ettinger et al., 2005](#); [Kuchen et al., 2007](#); [Parker, 1993a](#); [Parker, 1993b](#)). Suite à leur activation, les lymphocytes B vont alors proliférer et diversifier une nouvelle fois leur BCR par les mécanismes d'hypermutation somatique (HMS), de commutation isotypique (CI) et de conversion génique (CG). La majorité des cellules B subissant ces mécanismes le fera au sein de structures appelées centres germinatifs.

### 3.1 Les centres germinatifs

Les centres germinatifs (GCs ; de l'anglais germinal centers) sont des structures anatomiques que l'on retrouve suite à une infection au sein d'organes lymphoïdes secondaires, et plus particulièrement dans les centres des follicules des cellules B. C'est au sein des GCs que les lymphocytes B vont proliférer et subir les mécanismes de diversification de l'anticorps dépendants de l'antigène. Ce sont des structures dynamiques formées suite à la reconnaissance d'un Ag par des cellules B quiescentes, qui vont alors migrer en bordure du follicule pour recevoir des signaux de co-stimulation par des lymphocytes CD4+ ([Cyster, 2010](#); [Garside et al., 1998](#)). S'ensuit une prolifération des cellules B et la formation de clusters qui, à l'aide de cellules dendritiques folliculaires (FDCs ; de l'anglais follicular dendritic cells), formeront les GCs primitifs ([Mesin et al., 2016](#)).

Les centres germinatifs peuvent être délimités en 2 zones structurellement et fonctionnellement distinctes appelées zone sombre (DZ ; de l'anglais dark zone) et zone claire (LZ ; de l'anglais light zone). La DZ est la zone dans laquelle les lymphocytes B, alors appelés centroblastes, vont massivement proliférer tout en réalisant l'HMS et donc accumuler des mutations dans la région variable de leur BCR. La LZ contient des centrocytes, qui subissent une étape de sélection d'affinité de leur BCR pour un Ag par un processus appelé sélection clonale ([Victoria et al., 2010](#)). La sélection clonale fait en sorte que les cellules possédant les BCRs ayant le plus d'affinité vis-à-vis de l'Ag sont celles qui survivent, alors que celles possédant des BCRs avec une affinité plus faible meurent. Cette sélection se fait par l'interaction des centrocytes avec des FDCs et avec des cellules T helper folliculaires (Tfh ; de l'anglais T follicular helper). Alors que les FDCs ont notamment pour rôle de présenter l'Ag intact aux centrocytes pour que l'affinité des BCRs puisse être testée ([Heesters et al., 2014](#)), les Tfh sont capables de fournir des signaux de survie aux cellules B avec un BCR de haute affinité ([De Silva and Klein, 2015](#); [Victoria and Nussenzweig, 2012](#)).

Ainsi, les centrocytes sont capables de récupérer l'Ag présenté par les FDCs et de le présenter à leur tour aux Tfh, avec une efficacité dépendante de leur affinité pour cet Ag ([Batista and Neuberger, 2000](#)).

Parmi les signaux nécessaires à la survie des lymphocytes B, on retrouve notamment l'interaction CD40-CD40L ([Han et al., 1995](#)), promue par l'expression d'ICOSL par les lymphocytes B ([Liu et al., 2015](#)), mais aussi la cytokine BAFF majoritairement produite par les Tfh ([Goenka et al., 2014](#)). De même, il a été découvert que la force d'interaction entre les lymphocytes B et les Tfh ne déterminent pas seulement leur survie, mais également leur capacité à proliférer dans la DZ ([Gitlin et al., 2015](#); [Gitlin et al., 2014](#)).

En effet, les lymphocytes B transitent dans les GCs de façon cyclique : venant de la DZ où les lymphocytes B ont proliféré et généré des mutations dans leurs régions V par l'HMS, pouvant changer ainsi l'affinité de leur BCR pour l'Ag, ils transitent alors dans la LZ pour subir le mécanisme de sélection clonale de cellules B avec un BCR de haute affinité. Les lymphocytes B survivants vont ensuite repasser dans la DZ, et entrer dans un nouveau cycle de prolifération et d'HMS, et ainsi de suite. C'est ce processus qui permet l'accumulation de mutations rendant les BCRs de plus en plus affins pour l'Ag au fur et à mesure des cycles et qui est à la base de la maturation d'affinité. Ces transitions entre la DZ et la LZ sont notamment dues à la chimiokine CXCL12 exprimée par des cellules réticulaires (CRCs ; de l'anglais CXCL12-expressing Reticular Cells) qui attire et retient les centroblastes exprimant son récepteur CXCR4, alors que c'est la chimiokine CXCL13 exprimée par les FDCs et reconnue par le récepteur CXCR5 des lymphocytes B qui permet le passage vers la LZ ([Allen et al., 2004](#)). Les lymphocytes B de la LZ n'expriment ainsi que CXCR5, alors que ceux de la DZ expriment à la fois CXCR4 et CXCR5 ([Allen et al., 2004](#)).

Enfin, les Tfh vont également mener certains lymphocytes B à se différencier en plasmablastes et en cellules B mémoires et à sortir ainsi des GCs ([Victoria et al., 2010](#); [Zotos and Tarlinton, 2012](#)).

Les facteurs de transcription B cell lymphoma 6 protein (Bcl-6), requis pour la formation des GCs ([Basso et al., 2012](#)), et Paired box protein 5 (Pax5), un régulateur de l'identité des cellules B ([Revilla-I-Domingo et al., 2012](#)), inhibent tous deux la différenciation des cellules B en plasmablastes ([Recaldin and Fear, 2016](#)). Leur perte d'expression et l'induction concomitante de B lymphocyte-induced maturation protein 1 (BLIMP-1) par IRF4 et de X-box binding protein 1 (XBP-1) permet l'arrêt de la Cl, de l'HMS et de la prolifération de ces cellules, en entraînant leur différenciation en plasmablastes ([Nutt et al., 2015](#)). Les plasmablastes, sortis des GCs, se différencieront en plasmocytes alors capables de produire et de sécréter des anticorps affins et à

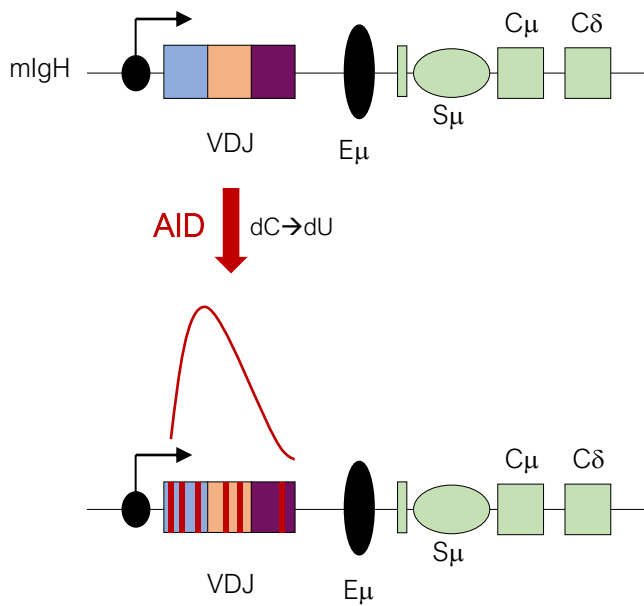
longue durée de vie en grande quantité, ce qui constitue le stade de différenciation ultime des cellules B.

### 3.2. L'hypermutation somatique

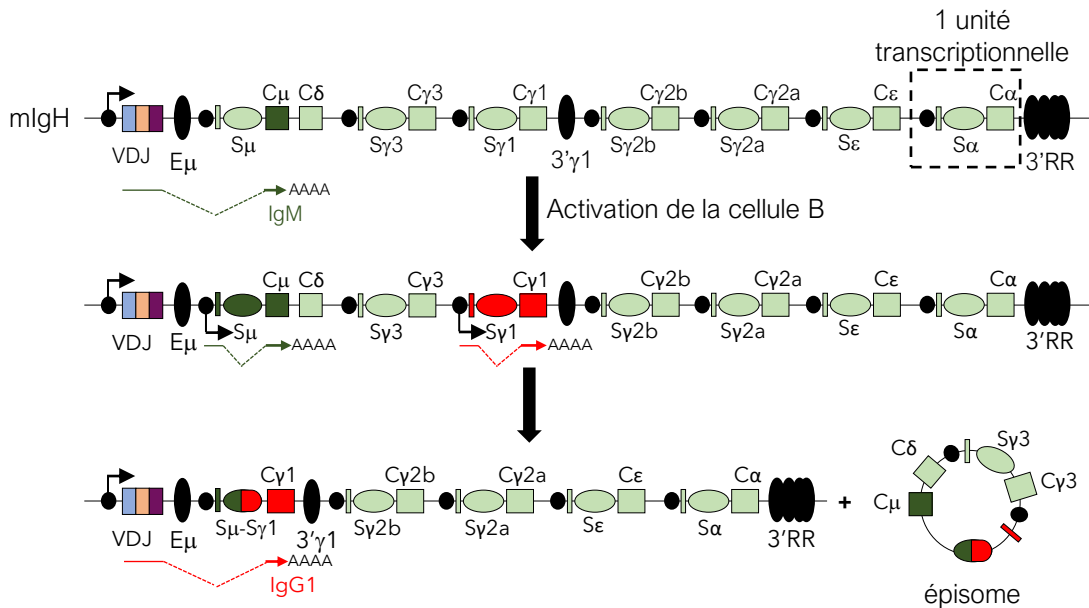
L'hypermutation somatique (HMS) est le processus qui permet l'introduction de mutations ponctuelles et plus rarement de petites insertions et délétions au sein des régions V des loci Ig, altérant ainsi l'affinité du BCR pour son antigène (**Figure 6**). Ces mutations ont lieu à une fréquence estimée de  $10^{-3}$  mutations par paire de bases (pb) et par division cellulaire ([Rajewsky et al., 1987](#)), ce qui est nettement plus élevé que le taux de mutation habituellement trouvé dans le reste du génome, estimé, lui, à  $10^{-9}$  mutations par pb et par division. La plupart de ces mutations est retrouvée à 150 pb en aval du promoteur et s'étend jusqu'à 1,5 ou 2 kb plus en aval ([Lebecque and Gearhart, 1990](#); [Rada and Milstein, 2001](#)). De plus, ces mutations s'accumulent préférentiellement dans les CDRs, qui sont en contact direct avec l'Ag, et assez peu dans les régions dites « framework » (FWRs ; de l'anglais framework regions), qui servent de support structural pour les CDRs ([Betz et al., 1993](#); [Rogozin and Kolchanov, 1992b](#); [Sharpe et al., 1991](#)). Il existe néanmoins des exceptions comme les anticorps neutralisants du virus de l'immunodéficience humaine (VIH) appelés bNAbs (de l'anglais broadly neutralizing antibodies) qui accumulent également des mutations dans les FWRs, permettant notamment d'augmenter leur flexibilité et de neutraliser efficacement le VIH ([Klein et al., 2013](#)). L'HMS est au centre d'un mécanisme appelé maturation d'affinité, qui permet à des anticorps de plus en plus affins d'être générés par les cellules B au cours d'une réponse immunitaire.

### 3.3. La commutation isotypique

La commutation isotypique (CI) est le mécanisme par lequel les lymphocytes B peuvent changer l'isotype exprimé par leurs BCRs ou les anticorps qu'ils produisent. Les lymphocytes B vont alors passer de l'expression d'IgM et/ou IgD à l'expression d'IgG, IgE ou IgA, sans toutefois modifier l'affinité pour l'Ag. La CI implique une réaction de recombinaison à longue distance se produisant au sein du locus IgH entre une région S donneuse ( $S_{\mu}$ ) et une région S acceptrice ( $S_x$ ) (**Figure 7**). Les régions S sont des régions répétées et non-homologues d'une longueur de 3 à 12 kilobases (kb), situées juste en amont d'un exon C. Ainsi, lors de la CI, des DSBs vont être générées dans chacune des régions S en jeu afin de pouvoir réaliser la recombinaison. Cette recombinaison va aboutir au rapprochement de la région variable et de l'exon C situé en aval de la région S acceptrice, placé dorénavant à la place des exons  $C_{\mu}$  et  $C_{\delta}$ . Tout comme pour la recombinaison



**Figure 6: L'hypermutation somatique.** L'hypermutation somatique permet la modification de la partie variable réarrangée (VDJ) des chaînes IgH et IgL des anticorps. Pour ce faire, des mutations ponctuelles, des petites insertions ou de petites délétions (traits rouges) y sont générées. La fréquence de mutations, qui est plus élevée à proximité du promoteur (rond noir), est représentée par la courbe rouge.



**Figure 7: La commutation isotypique.** La commutation isotypique est une réaction de recombinaison ayant cours au sein du locus IgH et ayant pour but de changer l'isotype des anticorps. Cette recombinaison est région spécifique car elle se produit entre deux régions aux séquences répétées appelées régions de switch (Sx). Ces régions de switch sont comprises au sein d'unités transcriptionnelles qui comprennent un promoteur inductible par des cytokines (rond noir), une région intronique (Ix) et un exon codant pour un nouvel isotype (Cx). Dans l'exemple montré ici, la recombinaison se déroule entre la région donneuse S $\mu$  et la région acceptrice S $\gamma$ 1. Les séquences comprises entre les deux régions de switch en jeu lors de la recombinaison vont être excisées du génome sous forme d'un épisome circulaire, ce qui fait que l'exon C $\mu$  est remplacé par C $\gamma$ 1. Cela va permettre l'expression de l'isotype IgG1, sans toutefois toucher à la région variable de l'anticorps.



V(D)J, la séquence génomique située entre les 2 régions S recombinées est excisée et perdue du génome sous forme d'un épisode circulaire. Ainsi, c'est à présent cet exon C qui sera transcrit à la suite des régions V, ce qui entraîne l'expression d'un BCR ou d'un Ac avec la même spécificité (les régions V n'ayant pas été touchées) mais avec un nouvel isotype. Un nouvel isotype permet de changer les fonctions effectrices de l'Ac en modifiant notamment sa localisation, sa durée de vie et la façon dont il interagit avec les autres cellules du système immunitaire et les molécules du complément, affectant son efficacité en fonction du type de pathogène contre lequel il faut lutter. Ainsi, le « choix » de l'isotype vers lequel les cellules B doivent réaliser la CI se fait grâce aux signaux qu'elles reçoivent des cellules T helper et de cellules dendritiques, signaux dépendant eux-mêmes du type de pathogène rencontré ([Stavnezer et al., 2008](#)). Parmi ces signaux, on retrouve notamment la liaison du récepteur CD40 avec son ligand CD40L exprimé par les T helper, mais aussi l'interleukine (IL)-4, l'IL-5, le transforming growth factor (TGF)- $\beta$  et l'interféron (IFN)- $\gamma$  par exemple. Chacun de ces signaux permet l'activation d'une cascade de signalisation intra-cellulaire aboutissant à l'expression ou à l'activation de facteurs de transcription qui vont se lier aux promoteurs d'une unité transcriptionnelle pour un isotype. Ainsi, chacun de ces promoteurs possède des éléments sensibles aux cytokines, et une combinaison spécifique de certains signaux aboutissant à la transcription du promoteur d'un isotype précis, ce mécanisme permet de guider la spécificité de la CI. La transcription à partir de ces promoteurs induit la production de transcrits germinaux (GLTs ; de l'anglais germline transcripts) dits stériles car ils ne codent pas pour une protéine ([Chaudhuri and Alt, 2004](#)). Ces transcrits sont néanmoins épissés afin d'éliminer la région intronique issue de la transcription des régions S, et ainsi joindre l'exon Ix à l'exon Cx. L'altération ou la suppression du promoteur d'un isotype entraîne une forte baisse de l'efficacité de CI vers cet isotype, ce qui indique que les transcrits stériles sont essentiels dans ce processus ([Bottaro et al., 1994](#); [Harriman et al., 1996](#); [Jung et al., 1993](#); [Zhang et al., 1993a](#)).

Les premiers anticorps sécrétés à la suite d'une première infection par un pathogène sont d'isotype IgM puisqu'ils sont sécrétés par des cellules n'ayant pas encore réalisé la CI. Ce sont souvent des anticorps de faible affinité que l'on retrouve généralement dans le sang. Ils ont toutefois la particularité de pouvoir s'assembler en pentamères stabilisés à l'aide de la chaîne J, ce qui leur confère une avidité plus grande, notamment si l'Ag est multivalent. De plus, les Ac IgM sont capables d'activer efficacement les molécules du complément. Les IgG sont monomériques et sont les anticorps les plus abondants dans le sang et dans les tissus. Ils ont une demi-vie relativement longue et peuvent être capables de neutraliser des toxines ou des virus. Ils sont également capables d'opsonisation et d'activer le complément et les cellules NK. Il existe plusieurs sous-classes d'IgG (IgG1 à IgG4 chez l'humain), la plus abondante étant IgG1. Les IgA se

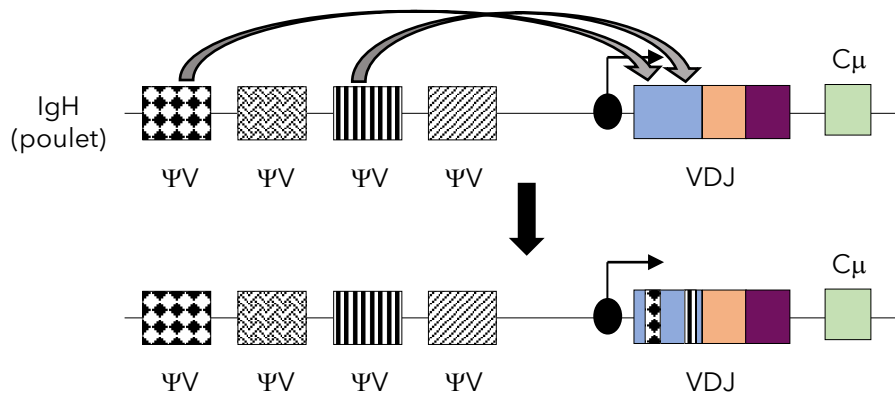


retrouvent principalement sous forme de dimères et ont la particularité de pouvoir passer la barrière épithéliale. On les retrouve ainsi dans les muqueuses et les sécrétions mais également dans l'intestin. Leur rôle consiste principalement à neutraliser des toxines ou des virus. Enfin, les IgE sont principalement associés aux mastocytes qu'on retrouve sous les épithéliums. Leur rôle primordial est de lutter contre les parasites extra-cellulaires, mais ils sont également impliqués dans l'hypersensibilité immédiate par l'activation des mastocytes qui relâchent alors de l'histamine ([Oettgen, 2000](#); [Wu and Zarrin, 2014](#)).

Les premières études décrivaient la CI comme un mécanisme se produisant très majoritairement dans les GCs ([Liu et al., 1996](#); [Shinkura et al., 1996](#)) et occasionnellement en dehors de ceux-ci ([Fagarasan et al., 2001](#); [Jacob et al., 1991](#); [Pape et al., 2003](#); [Toellner et al., 1998](#); [Toellner et al., 1996](#)). Cependant, ce paradigme est remis en cause par une étude récente ([Roco et al., 2019](#)). Il a ainsi été observé que les premiers GLTs apparaissent 1,5 jours après l'immunisation de souris, mais 2 jours avant la formation des GCs, au stade où se font les premières interactions entre les cellules B et T, et ce, quels que soient l'agent immunogène et l'adjuvant utilisés ([Roco et al., 2019](#)). De plus, l'apparition des GCs coïncide avec la baisse d'expression des GLTs. Enfin, l'analyse du profil transcriptomique de cellules individuelles de la LZ des GCs, sous-population dans laquelle on pensait que la CI se produisait, montre que la majorité de ces cellules n'expriment pas de GLTs, ce qui concorde avec les résultats précédents ([Roco et al., 2019](#)). Il semblerait donc que lors d'une réponse immunitaire, la CI se produise antérieurement à l'hypermutation somatique.

### 3.4. La conversion génique

La conversion génique (CG) est le processus grâce auquel certains animaux (principalement les oiseaux et des animaux de la ferme comme le lapin et le poulet par exemple) génèrent une diversité supplémentaire au sein des régions V, en plus de, ou à la place de l'HMS. Ce mécanisme a été découvert chez le poulet, où il a lieu au sein d'un organe lymphoïde primaire appelé bourse de Fabricius ([Reynaud, 1985](#); [Reynaud et al., 1987](#)). Ainsi, chez le poulet, la recombinaison V(D)J ne génère qu'une diversité très limitée, étant donné qu'il n'y a qu'un seul segment V, 15 segments D et un seul segment J dans le locus IgH, et à nouveau qu'un seul segment V et un seul J dans le locus IgL pouvant être recombinaison (les poulets ne possèdent pas de locus IgK) ([Reynaud et al., 1987](#)) (**Figure 8**). La CG permet donc de pallier à ce manque de diversité dans le répertoire primaire. Elle consiste en un échange non-réciproque d'une séquence d'ADN aléatoire entre des pseudogènes situés en amont des segments V dans chaque locus et les régions V déjà réarrangées à la suite du V(D)J. Ainsi, une séquence aléatoire pouvant aller jusqu'à 200 pb de



**Figure 8: La conversion génique.** La conversion génique est un mécanisme de diversification des anticorps se déroulant chez des espèces « de la ferme » comme le lapin ou encore le poulet. La conversion génique repose sur l'utilisation de pseudogènes ΨV pour diversifier la région IgV préalablement réarrangée. Adapté de Maizels et al., 2005 et de Tang et al., 2007.

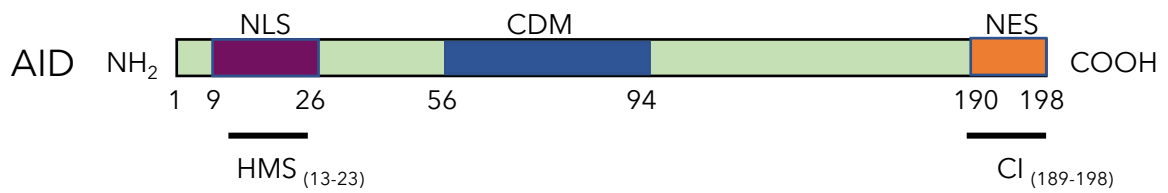
long et provenant des pseudogènes est insérée dans la région codante V pour générer de la diversité, par un mécanisme dépendant de la recombinaison homologue ([Helleday, 2003](#)). Si des cellules B perdent la capacité d'exprimer une Ig de surface (slg) à la suite de la CG, elles sont alors éliminées. Dans la plupart des espèces, ce répertoire secondaire d'anticorps peut être diversifié une nouvelle fois à l'aide de l'HMS dans les GCs ([Arakawa et al., 1996](#); [Tang and Martin, 2007](#)).

## II- L'enzyme Activation-induced cytidine deaminase (AID)

Les mécanismes qui viennent d'être évoqués (commutation isotypique, hypermutation somatique et conversion génique) ont la particularité d'être tous les trois dépendants de l'expression et de l'activité de la même enzyme, appelée activation-induced cytidine deaminase (AID). La découverte de l'implication d'AID dans l'initiation de ces processus, qui date d'il y a une vingtaine d'années, a permis de faire une percée dans la compréhension moléculaire des mécanismes de diversification des Ac dépendants des antigènes.

### 1. La découverte d'AID et de sa fonction

La découverte d'AID a grandement été facilitée par la mise en place d'un modèle cellulaire permettant d'étudier la CI. Cette lignée cellulaire, appelée CH12F3 (CH12), dérive d'un lymphome B murin et a été isolée pour sa capacité à pouvoir répondre au mix de stimulation TGF- $\beta$ , IL-4 et CD-40L en réalisant efficacement la CI vers IgA ([Nakamura et al., 1996](#)). Ainsi, un criblage par hybridation soustractive d'ADNc obtenu de CH12 stimulées et non-stimulées pour induire la CI a permis d'identifier AID comme étant exclusivement présente dans les CH12 stimulées ([Muramatsu et al., 1999a](#)). De plus, son expression semble limitée aux cellules B des GCs *in vivo* ([Muramatsu et al., 1999a](#)). La génération subséquente de souris déficientes pour AID a permis de découvrir qu'elle est à la fois nécessaire pour la CI et l'HMS ([Muramatsu et al., 2000](#)), alors que dans le même temps, l'équipe d'A. Durandy imputait le syndrome hyper-IgM de type II à une déficience d'AID ([Revy et al., 2000](#)). S'est alors posée la question de savoir quel est le rôle exact d'AID dans ces mécanismes. De par son appartenance à la famille des Apolipoprotein B mRNA Editing Catalytic Polypeptide-like (APOBECs) et de son homologie de séquence avec Apobec-1, il a d'abord été proposé qu'AID agisse sur l'ARN et notamment sur l'ARN messager (ARNm) d'un précurseur d'endonucléase qui, une fois édité, donnerait naissance à une endonucléase active encore non identifiée, et qui serait capable de couper dans les régions S et V des loci Ig



**Figure 9: Les domaines fonctionnels de la protéine Activation-Induced cytidine Deaminase (AID).** AID possède à son extrémité N-terminale un signal de localisation nucléaire (NLS) ainsi qu'un domaine requis pour l'hypermutation somatique (HMS). AID contient également un signal d'export nucléaire (NES) et un domaine requis pour la CI à son extrémité C-terminale. Enfin, son domaine catalytique (CDM ; cytidine deaminase motif) lui permettant de désaminer des cytosines en uraciles est compris entre les acides aminés 56 et 94. Les chiffres indiqués représentent les positions des acides aminés délimitant un domaine.

([Muramatsu et al., 2000](#)). Cependant, même si plusieurs études ont montré qu'AID est bien capable de se lier à l'ARNm ([Dickerson et al., 2003a](#); [Nonaka et al., 2009](#)), aucune n'a pu prouver qu'AID peut le modifier ([Bransteitter et al., 2003a](#); [Chaudhuri et al., 2003](#); [Fritz et al., 2013](#)). En revanche, peu de temps après sa découverte, il a été remarqué que la surexpression d'AID dans *E.coli* aboutit à l'apparition de mutations dans son ADN ([Petersen-Mahrt et al., 2002a](#)). Par la suite, d'autres études ont permis de confirmer qu'AID n'agit pas par l'intermédiaire d'un autre facteur pour initier l'HMS ([Martin et al., 2002](#); [Yoshikawa et al., 2002](#)), la CI ([Okazaki et al., 2002](#)) et la CG ([Arakawa and Buerstedde, 2004](#); [Harris et al., 2002](#)). Il existe aujourd'hui un nombre important de preuves apportées grâce à des études biochimiques et structurales d'AID, permettant d'affirmer qu'elle agit bien en désaminant l'ADN ([Chaudhuri et al., 2003](#); [Larijani et al., 2007](#); [Methot and Di Noia, 2017b](#); [Pham et al., 2016](#); [Rogozin and Diaz, 2004](#)).

## 2. La structure d'AID

Activation-induced Cytidine Deaminase (AID) est une petite protéine de 198 acides aminés qui appartient à la famille des apolipoprotéine B mRNA editing enzyme catalytic polypeptide-like (APOBECs). De fait, elle possède un domaine catalytique compris entre les acides aminés 56 à 94 et lui permettant la désamination de cytosines. Elle possède également un signal de localisation nucléaire (NLS ; de l'anglais nuclear localization signal) dans son domaine N-terminal (N-ter), et un signal d'export nucléaire (NES ; de l'anglais nuclear export signal) dans son domaine C-terminal (C-ter) ([Figure 9](#)).

Ces signaux NLS et NES sont compris au sein de domaines plus larges requis pour l'HMS et la CI et ne déterminent donc pas seulement la localisation sous-cellulaire d'AID ([Ito et al., 2004b](#); [McBride et al., 2004](#); [Patenaude et al., 2009](#)). En effet, 5 mutations ponctuelles dans la région N-ter éliminent l'activité initiateur d'HMS d'AID, sans toutefois affecter la CI ([Shinkura et al., 2004](#)), alors que la délétion de 10 acides aminés du domaine C-ter d'AID (AID<sup>Δ189-198</sup>) abolit la réaction de CI mais sans affecter l'HMS ([Barreto et al., 2003a](#); [Ta et al., 2003a](#)). Ces résultats révèlent également que l'initiation de l'HMS et de la CI est découplé. Il semblerait cependant que la région C-ter d'AID n'est pas essentielle pour son ciblage aux régions S, puisque le nombre et la répartition de mutations observées à S<sub>μ</sub> observés avec la version tronquée AID<sup>Δ189-198</sup> sont similaires à ceux observés avec la version sauvage d'AID ([Barreto et al., 2003a](#)). Néanmoins, le domaine C-ter pourrait être requis pour recruter des cofacteurs spécifiques à la CI ([Okazaki et al., 2002](#); [Ta et al., 2003a](#); [Yoshikawa et al., 2002](#)), alors que le domaine N-ter d'AID permettrait lui de lier des cofacteurs spécifiques à l'HMS ([Shinkura et al., 2004](#); [Ta et al., 2003a](#)).

AID appartient donc à la même famille de protéine qu'APOBEC1, APOBEC2 et APOBEC3. Les enzymes APOBECs sont capables de désaminer les cytosines dans l'ARN ou dans l'ADN et sont notamment impliquées dans des mécanismes de défense antiviraux. Lors de sa découverte, de par sa forte homologie avec APOBEC1 qui modifie l'ARN, il a été proposé qu'AID soit une enzyme d'édition de l'ARN ([Muramatsu et al., 1999b](#)). Depuis, nombre d'études à la fois biochimiques et génétiques ont montré qu'AID est une désaminase spécifique à l'ADN ([Bransteitter et al., 2003b](#); [Dickerson et al., 2003b](#); [Fritz et al., 2013](#); [Petersen-Mahrt et al., 2002b](#)).

### 3. La régulation d'AID

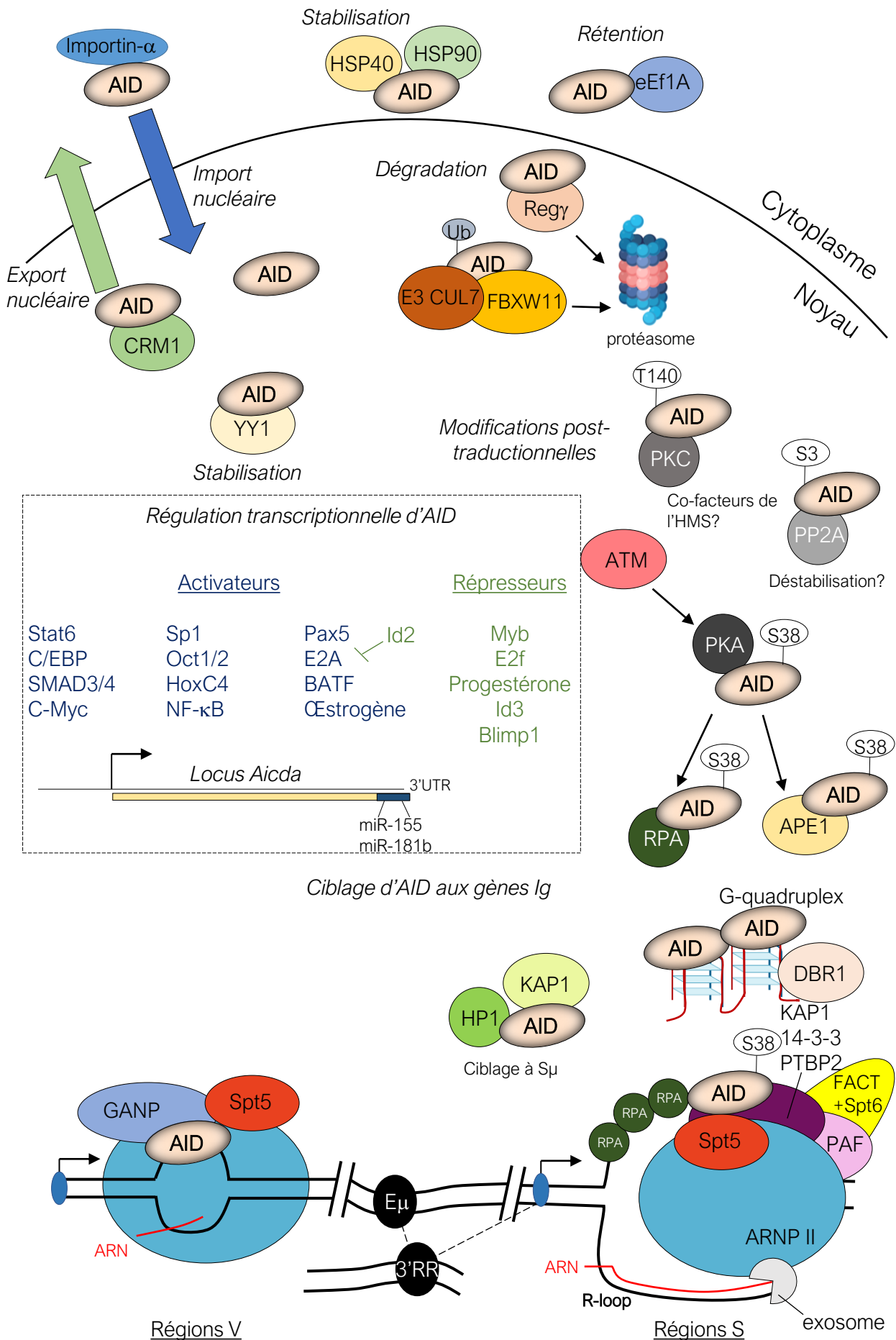
Il existe un ensemble de mécanismes permettant de contrôler la quantité, la disponibilité et l'activité d'AID. Plusieurs études ont montré que l'efficacité de la CI et de l'HMS est directement corrélée au niveau d'expression d'AID, comme l'atteste le phénotype des souris *Aicda*<sup>+/-</sup> qui voient leur niveau de CI et d'HMS divisés par deux ([Sernandez et al., 2008](#)). Cependant, on observe également chez ces souris 2 fois moins de translocations ([Takizawa et al., 2008](#)). A l'inverse, la surexpression d'AID est tumorigène ([Okazaki et al., 2003a](#); [Robbiani et al., 2009](#)), même si de faibles niveaux d'AID suffisent aussi à générer des lésions oncogéniques ([Pasqualucci et al., 2008](#); [Swaminathan et al., 2015](#)). Les mécanismes contrôlant le niveau et la localisation d'AID sont donc d'une importance primordiale pour mettre en place le juste équilibre qui permet la bonne diversification des anticorps, tout en limitant le risque d'une transformation oncogénique des lymphocytes B. Je vais présenter dans les prochains paragraphes ces différents mécanismes ([Figure 10](#)).

#### 3.1. Régulation de la transcription du locus *Aicda* :

Un des mécanismes par lequel passe la régulation de la quantité d'AID est la régulation de la transcription du locus *Aicda*. Au fil des études cherchant à identifier les facteurs capables d'influer sur cette régulation, quatre régions génomiques *cis*-régulatrices du locus *Aicda* ont été identifiées. On retrouve ainsi la région 4 située 8 kb en amont du site de démarrage de la transcription (TSS), la région 1 juste en amont du TSS, la région 2 située entre l'exon 1 et 2 et la région 3 située à 17 kb en aval du TSS ([Crouch et al., 2007](#); [Tran et al., 2010](#); [Yadav et al., 2006](#)).

Tout d'abord, afin de limiter l'expression d'AID à quelques sous-population cellulaires, il existe une forte répression transcriptionnelle médiée par des facteurs de transcription ubiquitaires E2f et c-Myb qui se lient à la région 2 ([Huong le et al., 2013](#); [Tran et al., 2010](#)). Les facteurs Inhibitor of differentiation (Id)2 et Id3 empêchent, eux, l'expression d'AID de manière indirecte en formant

Figure 10: Les mécanismes régulateurs d'AID.



**Figure 10: Les mécanismes régulateurs d'AID.** AID, du fait de son action mutagène, est finement contrôlée dans les lymphocytes B. L'expression d'AID est tout d'abord régulée au niveau transcriptionnel par des facteurs de transcription activateurs (Stat6, C/EBP, SMAD3/4, c-Myc, Sp1, Oct1/2, HoxC4, NF- $\kappa$ B, Pax5, E2A, BATF et le récepteur à l'œstrogène) et répresseurs (Myb, E2f, le récepteur à la progestérone et Id2) qui assurent une expression inductible ou restreinte d'AID dans les cellules B, suite à la rencontre de l'antigène et la activation du BCR. De plus, plusieurs miARNs (miR-155 et miR-181b) régulent négativement le niveau d'ARN messager d'AID.

Une fois exprimée, la quantité d'AID et son activité sont régulées par de nombreux cofacteurs. Ceux-ci participent notamment au contrôle de la localisation sous-cellulaire d'AID. Ainsi, CRM1 exporte activement AID du noyau alors que l'importine- $\alpha$  médie son import nucléaire. AID est stabilisée dans le cytoplasme par les protéines chaperones HSP40 et HSP90 et est retenu dans le cytoplasme par eEF1A. D'autres facteurs comme l'E3 CUL7, FBXW11 et REG $\gamma$  permettent sa dégradation dans le noyau par ubiquitination (Ub). Au contraire, l'interaction d'AID avec YY1 tend à la stabiliser. L'activité d'AID est contrôlée par des modifications post-traductionnelles, comme de la phosphorylation à S3, T140 et à S38. L'activation d'ATM suite aux premiers dommages générés par l'action d'ADN induit sa phosphorylation à la S38 par PKA, ce qui facilite son association avec RPA, facteur qui permet de stabiliser l'ADN simple brin durant la transcription et possiblement de cibler AID, et avec APE1, membre de la voie du BER.

La transcription active des loci Ig permet le ciblage d'AID aux régions S et IgV en exposant de l'ADN simple brin qui est le substrat d'AID. Le microenvironnement chromatinien du locus IgH inclut des interactions longues distances entre promoteurs (ovale bleu) et enhancers (E $\mu$  et 3'RR), requises pour une CI efficace. L'ARN polymérase II (ARNP II) effectue une pause durant la transcription des gènes cibles d'AID afin de recruter via les interactions impliquant les facteurs de pause/élongation, tels que Spt5 et l'exosome à ARN. Les séquences répétitives des régions S facilitent la formation de R-loops, ce qui favoriserait la pause de l'ARNP II et pourrait également cibler l'action d'AID en exposant de l'ADN simple brin sur de longues distances. D'autres complexes en lien avec la transcription permettent également de cibler l'action d'AID, comme par exemple l'exosome à ARN, le complexe FACT, PAF, et le facteur PTBP2. D'autres facteurs permettent de recruter et de retenir AID à certaines modifications d'histones présentes dans les régions S, tels que KAP1 et 14-3-3. Enfin, un modèle plus récent implique les structures secondaires G4, qui apparaissent dans l'ARN transcrit et dans l'ADN du brin non-transcrit, et qui semblent aussi requises pour le bon recrutement d'AID aux loci Ig.



des hétérodimères avec les protéines Pax5 et les E-protéines respectivement, inhibant ainsi leur fixation au locus *Aicda* ([Gonda et al., 2003](#); [Sayegh et al., 2003](#)). Enfin, une étude récente montre l'implication de Sirt1, une histone déacétylase, dans la régulation de l'expression d'AID ([Gan et al., 2020](#)). Sirt1 est exprimé dans les cellules B naïves et permet de désacétyler *Aicda*, et d'autres gènes cibles permettant d'activer les cellules B, tels ceux codant Dnmt1 et p65, une sous-unité de NF- $\kappa$ B, afin de réprimer leur transcription. Lors de l'activation des cellules B, Sirt1 est alors régulé négativement ce qui permet à terme d'augmenter l'expression d'AID et ainsi l'efficacité de CI et d'HMS ([Gan et al., 2020](#)).

Lors de l'activation des cellules B, plusieurs signaux vont permettre à celles-ci d'induire l'expression d'AID. Ces signaux sont notamment la liaison du récepteur de co-stimulation CD40 avec le récepteur CD154 des lymphocytes T, différentes cytokines telles que l'interleukine (IL) 4, le TGF- $\beta$  (de l'anglais Transforming Growth Factor  $\beta$ ) ou le LPS, le crosslinking de BCR ou encore l'activation de récepteurs Toll-like (TLR ; de l'anglais Toll-like receptors). Ces différents stimuli permettent l'activation de différents facteurs de transcription, et notamment de la voie canonique et non-canonique de NF- $\kappa$ B, de STAT6, de Smad3/4 ou encore de HoxC4 ([Dedeoglu et al., 2004](#); [Huong le et al., 2013](#); [Park et al., 2013](#); [Park et al., 2009](#); [Pone et al., 2012](#); [Tran et al., 2010](#)).

On retrouve ainsi des sites de liaison pour HoxC4, Sp1, Sp3, NF- $\kappa$ B, Stat6, CCAAT-enhancer-binding proteins (C/EBP) ainsi que pour c-myc dans la région 4 ([Dominguez-Sola et al., 2012](#); [Fernandez et al., 2013](#); [Huong le et al., 2013](#); [Tran et al., 2010](#)).

Les E-protéines et Pax5, facteurs plus spécifiques aux cellules B, peuvent stimuler l'activité transcriptionnelle du promoteur pour promouvoir l'expression d'AID en se liant à la région 2 ([Gonda et al., 2003](#); [Sayegh et al., 2003](#); [Yadav et al., 2006](#)).

De la même manière, le facteur de transcription Basic leucine zipper transcriptional factor ATF-like (BATF) joue un rôle activateur en se fixant sur la région 3, qui agirait comme un enhancer. Son absence induit une perte de l'expression d'AID mais il n'est cependant pas suffisant à lui seul pour induire l'expression d'AID ([Betz et al., 2010](#); [Huong le et al., 2013](#); [Ise et al., 2011](#)).

Trois autres enhancers situés plus loin du promoteur (jusqu'à 50 kb) ont également été identifiés en utilisant une technique nommée ChIA-PET ([Kieffer-Kwon et al., 2013](#)).

L'œstrogène et la progestérone jouent également un rôle dans cette régulation transcriptionnelle par l'intermédiaire de leurs récepteurs nucléaires. Le récepteur à l'œstrogène induit ainsi l'expression d'AID de façon directe en se fixant sur le locus *Aicda* ou de manière indirecte en activant le gène HoxC4 ([Mai et al., 2010](#); [Pauklin et al., 2009](#)). En revanche, le récepteur à la progestérone inhibe l'expression d'AID en se liant à la région 1 ([Pauklin and Petersen-Mahrt, 2009](#)). L'implication de ces hormones dans cette régulation pourrait expliquer

l'expression d'AID observée dans les oocytes et dans les cellules ES ([Incorvaia et al., 2013](#); [Morgan et al., 2004](#)) et avoir des implications dans l'auto-immunité ou l'oncogenèse liée aux hormones ([Incorvaia et al., 2013](#); [Pauklin and Petersen-Mahrt, 2009](#); [Pauklin et al., 2009](#)).

La répression du locus *Aicda* dans les cellules B naïves est également médiée par une hyperméthylation du promoteur ([Fujimura et al., 2008](#)). Suite à l'activation des lymphocytes B, le locus est déméthylé, notamment grâce à l'action des protéines ten-eleven-translocation (TET) recrutées par BATF ([Lio et al., 2019](#)). On observe alors un enrichissement des marques H3K4me3, H3K9ac et H3K14ac, ce qui permet d'augmenter l'accessibilité à la chromatine et d'induire l'expression d'AID ([Crouch et al., 2007](#); [Daniel et al., 2010](#); [Fujimura et al., 2008](#)).

Alors que son expression est très majoritairement réprimée dans les cellules non-B et les cellules B naïves, elle est fortement induite dans les cellules B activées puis à nouveau réprimée dans les plasmocytes ou les cellules B mémoires. En somme, c'est donc la balance entre des facteurs inhibiteurs et des facteurs activateurs qui dicte l'expression d'AID et la limite à quelques populations cellulaires. Alors que les facteurs inhibiteurs sont ubiquitaires, leur effet est contrebalancé lors de l'activation des lymphocytes B par l'expression de plusieurs facteurs activateurs ubiquitaires ou spécifiques aux cellules B et qui permettent l'induction de la transcription d'*Aicda* ([Tran et al., 2010](#)).

### 3.2. Régulation du niveau d'ARN messager d'AID :

La quantité d'ARN messager (ARNm) d'AID est contrôlée par au moins 6 microARN (miR) : miR-181b, miR-155, miR-361, miR-93 et miR-29b ([Basso et al., 2012](#); [Borchert et al., 2011](#); [Dorsett et al., 2008](#); [Recaladin et al., 2018](#); [Teng et al., 2008](#)). Ces miRs se fixent sur la partie 3' non traduite (UTR de l'anglais UnTranslated Region) de l'ARNm d'AID, empêchant de fait sa traduction. Ainsi, leur déficience est associée à une augmentation de la quantité protéique d'AID. De façon intéressante, l'absence de miR-155 est également associée à une forte augmentation de la fréquence des translocations entre IgH et c-myc, montrant un effet protecteur de ce miR pour le génome des lymphocytes B ([Dorsett et al., 2008](#); [Teng et al., 2008](#)). De plus, les cellules du lymphome de Burkitt dans lesquelles cette translocation est présente n'expriment pas le miR-155 ([Kluiver et al., 2006](#)). Il est aussi intéressant de noter que l'induction de la CI dans des lymphocytes B *in vitro* entraîne à la fois une forte expression d'AID et une augmentation de la quantité de miR-155, supposément pour contrôler finement la quantité d'AID ([Dorsett et al., 2008](#); [Teng et al., 2008](#)). Cependant, il a également été observé que Bcl-6, un répresseur transcriptionnel requis pour la mise en place des CG, régule négativement l'expression du miR-155 et de miR-361, induisant ainsi indirectement une hausse de la quantité d'AID dans les cellules des GC ([Basso et al., 2012](#)). Enfin, les miR-181b et miR-29b sont exprimés dans les cellules B

naïves et dans les cellules non-B pour éviter une expression basale d'AID mais sont réprimés lors de l'activation des lymphocytes B ([de Yebenes et al., 2008](#); [Dorsett et al., 2008](#); [Recaldin et al., 2018](#)).

### 3.3. Régulation de la localisation d'AID :

Les études cherchant à caractériser la localisation intracellulaire d'AID ont permis la découverte d'un constat singulier : bien qu'AID réalise son action dans le noyau, elle est très majoritairement située dans le cytoplasme, où elle ne semble pas y jouer de rôle ([Ito et al., 2004a](#); [Pasqualucci et al., 2004](#); [Rada et al., 2002a](#)). Cette répartition est due à une régulation complexe faisant intervenir rétention cytoplasmique, import et export actifs du noyau, ainsi qu'une stabilité protéique différente selon le compartiment cellulaire.

En effet, une fois traduite, AID est tout d'abord prise en charge par 2 protéines chaperones : Heat Shock Protein-90 (HSP-90) et DnaJ homolog subfamily A member 1 (DNAJA1). Ce complexe permet de stabiliser et de rendre AID mature ([Methot et al., 2015](#); [Orthwein et al., 2010](#); [Orthwein et al., 2012](#)). De fait, l'inhibition ou la déficience d'un de ces 2 facteurs aboutit à une baisse de la quantité d'AID et de son activité et donc à une baisse de l'efficacité de l'HMS et de la CI ([Montamat-Sicotte et al., 2015](#); [Orthwein et al., 2010](#); [Orthwein et al., 2012](#)). Une fois mature, elle est retenue dans le cytoplasme par un autre complexe de haut poids moléculaire contenant notamment eukaryotic Elongation Factor 1 alpha (eEF1A) ([Hasler et al., 2011](#); [Methot et al., 2015](#)). La stœchiométrie observée entre AID et eEF1A fait penser que ce complexe contient la majorité d'AID présent dans la cellule ([Hasler et al., 2011](#)). De plus, la rupture de ce complexe entraîne une augmentation de la CI et des translocations ([Methot et al., 2015](#)), appuyant l'hypothèse selon laquelle la majorité d'AID est mature et potentiellement active mais mise à l'écart de son substrat en étant retenue dans le cytoplasme.

AID circule constamment entre le cytoplasme et le noyau ([Brar et al., 2004](#); [Ito et al., 2004a](#); [McBride et al., 2004](#)). Bien qu'elle possède un poids moléculaire faible (24 kDa) lui permettant en théorie de diffuser passivement dans le noyau ([Macara, 2001](#)), elle est importée activement par l'importine- $\alpha$  qui reconnaît un signal de localisation nucléaire (NLS, de l'anglais Nuclear Localization Signal) bipartite situé en N-terminal d'AID ([Patenaude et al., 2009](#)). Plus récemment, une autre étude a également souligné l'importance d'autres résidus chargés positivement, répartis tout au long de la séquence protéique d'AID pour son import, sans que le mécanisme mis en jeu ne soit élucidé ([Hu et al., 2013](#)). De plus, il a été montré qu'AID pouvait être localisée dans les nucléoles, où l'on retrouve également une sous-unité du spliceosome appelée Beta-catenin-like protein 1 (CTNNBL1). Or, dans les cellules de poulet DT40, il a été suggéré que le facteur CTNNBL1 était requis pour l'import d'AID dans le noyau ([Ganesh et al., 2011](#)). Bien que l'interaction entre AID et CTNNBL1 ait initialement été identifiée dans un modèle murin ([Conticello](#)

[et al., 2008](#)), la déficience de CTNNB1 n'a en revanche pas d'impact significatif sur la CI dans les CH12 ([Han et al., 2010](#)), et qu'un impact modéré dans des cellules B murines primaires ([Chandra et al., 2013](#)), laissant suggérer que son rôle est limité à la GC et à l'HMS, ou que son rôle est redondant à d'autres protéines. Il a également été signalé que la surexpression de Germinal center Associated Nuclear Protein (GANP) entraîne l'accumulation d'AID dans le noyau ([Maeda et al., 2010](#)), laissant suggérer un rôle dans l'import d'AID, bien que cet effet puisse également être dû au fait que GANP est requis pour une HMS efficace en permettant de recruter AID aux loci Ig ([Kuwahara et al., 2004](#); [Singh et al., 2013](#)). De même, l'isoforme 3 de serine arginine protein splicing factor 1 (SRSF1-3) permet l'accumulation d'AID dans le noyau de cellules DT40 en interagissant avec le domaine C-ter d'AID ([Kawaguchi et al., 2017](#))

Plus récemment, une étude propose qu'AID soit capable de rentrer dans le noyau par de courtes pulsations d'une trentaine de minutes, coordonnées avec une accumulation nucléaire de p53, un régulateur de la réponse cellulaire aux dommages à l'ADN ([Le and Maizels, 2019](#)). Cependant, ces observations ont été faites dans des lignées non-B. Il reste néanmoins à confirmer si ce phénomène a également lieu dans les lymphocytes et à déterminer quelle est son importance vis-à-vis de l'HMS et de la CI.

AID est également piégée dans le noyau juste après la mitose, lorsque la membrane nucléaire se reforme ([Hu et al., 2013](#); [Ito et al., 2004a](#); [Methot et al., 2018](#); [Patenaude et al., 2009](#); [Wang et al., 2017](#)), ce qui est corrélé avec l'observation que la désamination d'AID et la réparation de l'ADN subséquente par le BER ou le MMR sont principalement réalisées dans la phase G1 du cycle cellulaire ([Faili et al., 2002](#); [Sharbeen et al., 2012](#)).

Une fois dans le noyau, la demi-vie d'AID dans le noyau est raccourcie pour limiter son potentiel oncogénique. Pour cela, AID est dégradée par le protéasome grâce à deux mécanismes dont l'un est médié par l'ubiquitine ligase E3 CUL7 et FBXW11 ([Luo et al., 2019](#)) en étant dépendant de l'ATP et de la polyubiquitinylation d'AID ([Aoufouchi et al., 2008](#)) et l'autre, médié par REG- $\gamma$  mais non dépendant de l'ubiquitine et de l'ATP ([Uchimura et al., 2011](#)). L'interaction d'AID avec le facteur de transcription Yin Yang 1 (YY1) permet en revanche de réguler l'accumulation d'AID dans le noyau en y stabilisant sa présence, probablement en empêchant son ubiquitination ([Zaprazna and Atchison, 2012](#); [Zaprazna et al., 2018](#)).

Enfin, AID est également exportée activement du noyau par CRM1 qui reconnaît un signal d'export nucléaire (NES, de l'anglais Nuclear Export Signal) situé en C-ter d'AID ([Brar et al., 2004](#); [Geisberger et al., 2009](#); [Ito et al., 2004a](#); [McBride et al., 2004](#)).

Il est intéressant de noter que l'inhibition de tous ces mécanismes n'entraîne pas toujours les mêmes effets sur AID. Alors que l'inhibition d'eEF1a entraîne une augmentation de la quantité

d'AID nucléaire, de CI et de translocations ([Methot et al., 2015](#)), la déficience de REG- $\gamma$  augmente également la quantité d'AID nucléaire, mais sans qu'on puisse observer plus de translocations ([Uchimura et al., 2011](#)). Au contraire, l'inhibition de HSP-90 réduit l'efficacité de la CI et de l'HMS ainsi que l'effet hors-cible d'AID ([Montamat-Sicotte et al., 2015](#); [Orthwein et al., 2010](#)). Ces différences pourraient découler du fait que ces mécanismes interviennent à différents moments de la « vie » d'AID. REG- $\gamma$  n'entraîne la dégradation d'AID qu'une fois son action accomplie, alors que la fraction d'AID libéré d'eEF1a est capable de cibler tout le génome ([Casellas et al., 2016b](#)).

### 3.4. Régulation de l'activité d'AID pendant le cycle cellulaire

L'activité d'AID semble majoritairement restreinte à la phase précoce G1 du cycle cellulaire ([Faili et al., 2002](#); [Khair et al., 2014](#); [Wang et al., 2017](#)), ceci étant probablement conséquent à l'action synergétique de plusieurs mécanismes tels que la régulation de la localisation en fonction de la phase du cycle cellulaire, sa phosphorylation, la disponibilité de ses co-facteurs ou encore la régulation de sa stabilité nucléaire par des cyclines ([He et al., 2015](#); [Lackey et al., 2013](#); [Le and Maizels, 2015](#)). Les mutations induites par AID, ainsi que les DSBs subséquentement induites, sont également réparées en G1 ([Petersen et al., 2001](#); [Schrader et al., 2007a](#); [Sharbeen et al., 2012](#); [Yamane et al., 2013](#)). Dans les cellules B déficientes en facteurs de réparation nécessaires à la réparation des DSBs (HR et NHEJ), celles-ci s'accumulent alors jusque dans les phases S, G2 et M ([Hasham et al., 2010](#); [Yamane et al., 2013](#)).

### 3.5. Modifications post-traductionnelles

On connaît à ce jour 7 sites de phosphorylation pour AID. Ceux-ci se situent sur les résidus sérine S3, thréonine T27, sérine S38, sérine S41, sérine S43, thréonine T140 et tyrosine Y184 ([Basu et al., 2005](#); [Gazumyan et al., 2011](#); [McBride et al., 2006](#); [McBride et al., 2008](#); [Pasqualucci et al., 2006](#)). La phosphorylation de ces résidus permet de réguler l'activité et le ciblage d'AID. Le résidu le plus étudié est la sérine en position 38 (S38) qui est phosphorylé par cAMP-dépendent protein kinase A (PKA) ([Basu et al., 2005](#)). Seule une petite fraction d'AID est phosphorylée à cette position (5 à 15%), mais c'est cette fraction qui contient la majorité d'AID associée à la chromatine et qui engendre de façon disproportionnée la plupart des mutations observées dans l'HMS ([McBride et al., 2006](#)). Ce phénomène peut s'expliquer par le fait que cette modification permet l'interaction d'AID avec Replication Protein A (RPA) et Apurinic/apyrimidic endonuclease (APE) 1 ([Basu et al., 2005](#); [Vuong et al., 2013](#)), deux protéines essentielles pour le ciblage et l'action d'AID. On observe de ce fait un défaut quasi-total de CI et d'HMS lorsque la sérine 38 est mutée en alanine (AID<sup>S38A</sup>) alors même que l'activité catalytique d'AID reste intacte. Ceci s'explique par la perte de sa capacité à générer des DSBs, ce qui concorde avec la perte d'interaction avec

APE1, qui joue un rôle critique dans ce processus ([Cheng et al., 2009](#); [Vuong et al., 2013](#)). On observe de ce fait une baisse significative du niveau d'HMS et de CI dans les cellules n'exprimant qu'AID<sup>S38A</sup> ou une version non fonctionnelle d'une sous-unité régulatrice de PKA ([Cheng et al., 2009](#); [McBride et al., 2008](#); [Vuong et al., 2009](#)). De la même façon, on constate une abolition totale de la CI ainsi qu'un défaut sévère de l'HMS dans des souris AID<sup>S38A/S38A</sup> MSH2<sup>-/-</sup>, mais, de façon intéressante, on observe des résultats similaires dans des souris AID<sup>S38A/S38A</sup> UNG<sup>-/-</sup>, ce qui indique qu'AID phosphorylée à la S38 est requise en l'absence d'Ung pour que le MMR puisse intervenir sur les mutations qu'elle induit, dans le cadre de l'HMS et de la CI ([Choi et al., 2020](#)).

PKA est en fait associée directement aux régions S lorsque celles-ci sont transcrites après activation des lymphocytes B pour réaliser la CI, ce qui suggère que c'est là qu'elle phosphoryle AID ([Vuong et al., 2009](#)). Cependant, une expérience a exposé qu'un mutant catalytiquement inactif d'AID mais capable de se lier aux régions S n'était pour autant pas phosphorylé, ce qui montre que l'activité d'AID est cruciale pour sa phosphorylation ([Vuong et al., 2013](#)). De plus, dans des cellules déficientes en Ung et Msh2 et dans lesquelles les dommages induits par AID sont donc incapables de générer des DSBs, on observe également une forte baisse du niveau de phosphorylation d'AID à S38. Conformément à ces résultats, la déficience d'ATM, une kinase clé dans la reconnaissance et dans la réparation des DSBs ([Bensimon et al., 2011](#)), entraîne aussi un défaut de phosphorylation d'AID ([Vuong et al., 2013](#)). Ces défauts sont cependant sauvés par l'induction artificielle de DSBs dans le génome à l'aide de rayons ionisants, ce qui indique que ce sont les dommages à l'ADN qui permettent la bonne phosphorylation d'AID. Ensemble, ces résultats ont donc permis l'élaboration d'un modèle où AID induit des mutations dans les régions S, entraînant l'apparition de DSBs à l'aide du BER et du MMR. Ces DSBs vont activer ATM, permettant elle-même à PKA de phosphoryler plus d'AID, qui sera alors capable de mieux s'associer à RPA et à APE1, et donc plus active. Dans ce modèle, AID est donc au centre d'une boucle de rétroaction positive qui renforce son activité et permet une forte densité de DSBs dans les régions S. Il permet aussi d'expliquer pourquoi la fréquence de DSBs induites par les mutations d'AID est beaucoup plus faible dans les gènes non-Ig que dans les régions S, où PKA est spécifiquement associée ([Vuong et al., 2009](#)). Néanmoins, le mécanisme permettant à ATM d'activer PKA reste à élucider.

Parmi les autres phosphorylations, celle à la thréonine T140 est médiée par la protéine kinase C *in vitro* ([McBride et al., 2008](#)). À nouveau, la mutation de ce résidu en alanine (AID<sup>T140A</sup>) n'affecte pas l'activité d'AID, mais semble plus affecter l'HMS que la CI, ce qui pourrait suggérer que cette phosphorylation joue un rôle dans l'interaction avec des facteurs spécifiques de l'HMS ([McBride et al., 2008](#)).

Contrairement aux autres modifications post-traductionnelles, la phosphorylation de sérine S3, catalysée par protéin phosphatase 2 (PP2A), permet de réguler négativement l'activité d'AID.

En effet, l'analyse du défaut de phosphorylation de ce résidu à l'aide d'un mutant montre une augmentation de l'efficacité d'HMS et de CI, ainsi qu'une augmentation des translocations entre IgH et c-myc ([Gazumyan et al., 2011](#)).

Bien que la phosphorylation de la tyrosine 184 d'AID ait été observée *in vivo* ([Basu et al., 2005](#); [McBride et al., 2006](#)), on ne connaît pas à ce jour l'impact physiologique de cette modification, puisqu'elle ne joue de rôle ni sur la compartimentalisation, ni sur l'activité d'AID ([Patenaude and Di Noia, 2010](#)).

## 4. Le ciblage spécifique de l'activité d'AID

Comment AID est-elle capable d'être ciblée et d'induire majoritairement des mutations aux loci Ig ? Les loci Ig possèdent-ils des particularités intrinsèques ou est-ce que l'action d'AID est ciblée à l'aide d'interactions avec des co-facteurs ? Les mécanismes de ciblage d'AID diffèrent-ils entre CI et HMS ? Toutes ces questions ont fait l'objet de nombreuses recherches mais ne sont toujours pas complètement résolues.

Plusieurs mécanismes ont été identifiés pour expliquer le ciblage d'AID majoritairement aux loci Ig mais également à de nombreux autres loci ([Yamane et al., 2011b](#)). Cependant, la contribution exacte de chacun de ces mécanismes, et la façon dont ils influent les uns sur les autres restent encore mal connues. Je détaille par la suite ces mécanismes ([Figure 10](#)).

### 4.1. Les motifs de séquence d'ADN

Avant la découverte d'AID, des études visant à étudier les mutations induites lors de l'HMS *in vivo* ont permis de mettre en lumière le fait que ces mutations ne sont pas aléatoires : on les retrouve notamment enrichies dans des séquences points chauds RGYW, où R = A/G, Y = C/T et W = A/T ([Rogozin and Kolchanov, 1992a](#)). Ce n'est qu'après la découverte d'AID que des études biochimiques ont permis d'appuyer ces résultats en montrant que l'efficacité de désamination d'une cytosine dépendait des nucléotides l'entourant ([Pham et al., 2003](#); [Yu et al., 2004](#)). Elles ont ainsi permis d'identifier le motif WRCY, complémentaire de RGYW, comme étant préférentiellement ciblé par AID ([Pham et al., 2003](#); [Yu et al., 2004](#)). Une autre étude a suggéré que les motifs DGYW / WRCH (où D = A/G/T et H = A/C/T) sont de meilleurs indicateurs des points chauds d'AID ([Rogozin and Diaz, 2004](#)). On retrouve ainsi incluse dans le motif DGYW, la séquence AGCT qui est notamment enrichie au sein des CDRs des régions variables et plus particulièrement encore au sein des régions S, avec l'avantage d'engendrer la même séquence



sur le brin complémentaire ([Hackney et al., 2009](#)). Comme cela a été mentionné auparavant, ce sont les protéines adaptatrices 14-3-3 qui jouent un rôle dans le recrutement d'AID aux séquences AGCT ([Xu et al., 2010](#)). Ces résultats sont également appuyés par des études sur la structure d'AID confirmant qu'AID contient une boucle de spécificité qui s'étend en dehors du site catalytique et qui est capable d'interagir avec 2 purines adjacentes à la cytosine ciblée ([Kohli et al., 2009](#); [Pham et al., 2016](#); [Wang et al., 2010](#)). Enfin, il semblerait que ce soit le nombre de motifs WRCY au sein d'une région S qui est le facteur important, plutôt que la densité ou la distance entre ces différents motifs ([Hwang et al., 2015](#); [Zhang et al., 2014a](#)).

Cependant, ces motifs ne semblent pas absolument requis pour la CI, puisque la délétion de la quasi-totalité de S $\mu$ , et donc des motifs WRCY, n'abolit pas complètement la CI ([Khamlichi et al., 2004](#)). De la même façon, une étude utilisant un allèle passager a montré que c'est la localisation de la région V au sein du locus IgH qui joue un rôle dans le ciblage d'AID plutôt que sa séquence proprement dite ([Yeap et al., 2015](#)).

#### 4.2. Les éléments en *cis*

Des éléments cis-régulateurs ont été identifiés au sein des loci Ig permettant le ciblage d'AID et promouvant donc les différents mécanismes de diversification des anticorps. Parmi ceux-ci, on retrouve des éléments conservés nommés DIVACs (de l'anglais diversification activators), importants pour la CG et l'HMS, et contenus au sein de séquences enhancer et « enhancer-like » des loci Ig ([Kim and Tian, 2010](#); [Kothapalli et al., 2008](#)). De façon intéressante, la juxtaposition d'un élément DIVAC à une séquence transcrite en dehors des loci Ig suffit à induire l'HMS dans celle-ci, ce qui montre que les DIVACs suffisent à cibler AID pour faire de l'HMS ([Blagodatski et al., 2009](#)). Plus récemment, ce concept a été étendu aux TADs (Topology Associated Domains), qui sont des domaines de la chromatine dans lesquels les séquences d'ADN interagissent plus entre elles qu'avec des séquences situées en dehors du TAD. Il a alors été observé que certains TADs seraient plus susceptibles à l'HMS que d'autres, et que cette susceptibilité dépendrait au moins en partie d'une activité DIVAC. En effet, un TAD à priori non ciblé par l'HMS le devient lorsqu'un DIVAC fort est inséré en son sein ([Senigl et al., 2019](#)).

Les DIVACs contiennent des sites de liaison pour une variété de facteurs de transcription de la famille d'Ikaros, des E-box ou de MEF2 notamment, et qui sont nécessaires pour promouvoir l'activité d'AID ([Buerstedde et al., 2014](#); [Dinesh et al., 2020](#); [Kohler et al., 2012](#); [Senigl et al., 2019](#)). Leur mécanisme d'action n'est pas clair, mais il semblerait que le ciblage d'AID nécessite une combinaison de ces sites de liaison ([Buerstedde et al., 2014](#)). Néanmoins, il se pourrait que les DIVACs soient transcrits de manière bidirectionnelle depuis les 2 brins d'ADN afin de produire de l'enhancer ARN (eARN) ([Senigl et al., 2019](#)). Or, d'autres études ont montré que la



transcription bidirectionnelle ainsi que la dégradation de l'eARN par l'exosome contribuent au ciblage d'AID ([Pefanis et al., 2014](#); [Pham et al., 2019](#)), pouvant expliquer, au moins en partie, l'effet des DIVACs.

La région régulatrice 3'RR située dans le locus IgH est également essentielle pour le ciblage d'AID. Outre son rôle majeur dans la régulation de la transcription du locus IgH et dans l'ouverture de celui-ci et donc dans la CI et l'HMS de manière générale ([Dunnick et al., 2009](#); [Dunnick et al., 2013](#); [Pinaud et al., 2011](#); [Rouaud et al., 2013](#); [Vincent-Fabert et al., 2010](#)), la région 3'RR joue également un rôle plus direct dans le recrutement d'AID aux régions V ([Rouaud et al., 2013](#)) et aux régions S ([Saintamand et al., 2015](#)). Plus récemment, il a même été montré que l'activité de super-enhancer de la région 3'RR aide également au recrutement d'AID par transvection, à savoir le recrutement d'AID sur l'autre allèle IgH grâce à une proximité inter-allélique ([Le Noir et al., 2017](#)).

### 4.3. La transcription

Un lien entre les mécanismes de diversification et la transcription fut établi avant même la découverte d'AID ([Hein et al., 1998](#); [Peters and Storb, 1996](#)). Les signaux reçus par les cellules B lors de leur activation permettent l'activation de promoteurs spécifiques entraînant la production de transcrits stériles appelés GLT essentiels à la CI. En effet, la délétion des promoteurs des gènes codant pour les différents isotypes entraîne une ablation de la CI vers ces mêmes isotypes ([Bottaro et al., 1994](#); [Jung et al., 1993](#); [Zhang et al., 1993b](#)).

Ces observations trouvèrent une explication lorsqu'une expérience de co-immunoprécipitation (co-IP) démontra l'interaction directe d'AID avec l'ARN polymérase II (ARN pol II) ([Nambu et al., 2003](#)). Un modèle émergea alors : la transcription par l'ARN pol II joue un rôle majeur dans le ciblage d'AID en permettant d'exposer l'ADN sous forme simple brin, qui est le substrat d'AID ([Chaudhuri et al., 2003](#); [Ramiro et al., 2003](#)).

A l'aide de CHIP, d'autres études ont observé une accumulation de l'ARN pol II aux régions S, et plus particulièrement de la version phosphorylée à la sérine 5 du domaine C-ter qui est associée à l'ARN pol II en pause ([Rajagopal et al., 2009](#); [Wang et al., 2009](#)). Cette accumulation est de plus corrélée avec les mutations induites par AID, ce qui a suggéré l'importance de la pause de l'ARN pol II dans le ciblage et l'activité d'AID ([Canugovi et al., 2009](#); [Wang et al., 2009](#)). L'ARN pol II peut notamment entrer en pause à proximité du promoteur (entre 20 à 50 pb en aval du TSS) ([Core and Lis, 2008](#)), lorsqu'elle rencontre des uraciles dans l'ADN précédemment générés par AID, ou encore en présence de structures secondaires telles que les R-loops et les G-quadruplex (évoquées plus en détail dans la prochaine section), situés plus à l'intérieur du gène ([Sun et al., 2013b](#)). L'identification du facteur Spt5 à l'aide d'un crible réalisé avec des short-hairpin ARN

(shARN) a confirmé cette corrélation ([Pavri et al., 2010](#)). Spt5 est une sous-unité du complexe DSIF (de l'anglais DRB Sensitivity Inducing Factor) dont le rôle est de stabiliser l'ARN pol II en pause. De plus, des données de ChIP-Seq pour AID et Spt5 ont montré qu'ils sont colocalisés sur l'ensemble du génome, notamment aux régions S et V des gènes Ig ([Maul et al., 2014](#); [Pavri et al., 2010](#)). En l'absence de Spt5, le recrutement d'AID aux gènes Ig mais aussi aux gènes non-Ig est fortement réduit ([Pavri et al., 2010](#)). Ces données indiquent que c'est Spt5, associé à la polymérase en pause, qui permet le recrutement d'AID aux gènes transcrits. Des études biochimiques ont révélé qu'AID était une enzyme inefficace *in vitro*, ce qui pourrait laisser penser que son association avec Spt5 et l'ARN pol II en pause est requise pour lui laisser assez de temps pour réaliser des déaminations ([King et al., 2015](#); [Larijani and Martin, 2012](#); [Pham et al., 2016](#)).

L'ARN pol II en pause peut être dégradée par le biais d'une mono-ubiquitination opérée par la ligase E3 Nedd4 ([Anindya et al., 2007](#)). Or, cette dernière a été identifiée comme étant requise pour une CI efficace, ce qui suggère que la dégradation de l'ARN pol II en pause est nécessaire pour qu'AID puisse accéder à l'ADN simple brin ([Sun et al., 2013a](#)). Il a été observé que la pause de la polymérase peut également entraîner une terminaison précoce de la transcription aux régions V, permettant également de fournir à AID son substrat ([Wang et al., 2014](#)). Enfin, le surenroulement négatif généré en aval de la polymérase lors de la transcription permet aussi de générer de l'ADN simple brin ([Parsa et al., 2012](#)).

Conformément à ce modèle, la pause de l'ARN pol II a également pu être observée lors de la transcription de gènes hors-cible d'AID, dans lesquels on retrouve des mutations proches du TSS ([Pavri et al., 2010](#); [Yamane et al., 2011b](#)).

Ainsi, bien que la transcription, et plus particulièrement la pause de la polymérase, joue un rôle majeur dans le ciblage d'AID, ces mécanismes ne permettent pas à eux seuls de prédire quels gènes sont mutés par AID. En effet, la pause proximale au promoteur de l'ARN pol II est observée dans de nombreux gènes ([Adelman and Lis, 2012](#)), et tous ne sont pas ciblés par AID ([Yamane et al., 2011b](#)). De plus, on ne retrouve pas toujours de mutations dans les gènes liés par AID. Ce phénomène peut découler du fait que le ciblage d'AID n'est pas forcément corrélé avec son activité ([Casellas et al., 2016b](#); [Matthews et al., 2014](#); [Yamane et al., 2011b](#)), ou bien que les mutations qu'elle induit sont réparées de façon fidèle dans certains gènes hors-cible ([Liu et al., 2008](#)). La question de savoir quel mécanisme déclenche la désamination d'AID n'est pas encore résolue. Cependant, l'étude de mutants d'AID dans lesquels 3 résidus arginines (R171, R174 et R178) sont mutés un par un a permis de démontrer l'importance de l'association d'AID au complexe d'élongation de la transcription ([Methot et al., 2018](#)). En effet, ces mutants n'ont pas d'activité catalytique diminuée et sont bien capables de s'associer à Spt5 et donc à l'ARN pol II

en pause, mais semblent incapables d'accéder à l'intérieur des gènes Ig car ils n'interagissent plus avec Spt6, un facteur associé à l'ARN pol II en élongation. Ces 3 résidus arginines font donc partie d'un même domaine fonctionnel d'AID qui lui permettrait d'accéder à l'intérieur des gènes ciblés au travers de l'élongation de la transcription ([Methot et al., 2018](#)).

Conformément à ces résultats, AID forme également un complexe avec des facteurs interagissant avec l'ARN pol II en cours d'élongation de la transcription, à savoir RNA polymerase II-associated factor (PAF), Spt6 et le complexe facilitates chromatin transcription (FACT) composé de Spt16 et SSRP1 ([Okazaki et al., 2011](#); [Willmann et al., 2012](#)). FACT permet notamment l'accumulation de la marque d'histone H3K4me3 aux régions S, et est requis pour une CI efficace ([Stanlie et al., 2010](#)). Tout comme Spt6, FACT permet aussi le remodelage de la chromatine pour faciliter la transcription ([Saunders et al., 2006](#)). Spt6, de par son interaction avec la méthyltransférase Set1, permet également le dépôt de la marque H3K4me3 aux régions S ([Begum et al., 2012](#)). Enfin, le complexe PAF permet lui, le recrutement d'AID aux loci Ig et sert de plateforme de recrutement pour FACT, Spt6, et Set1 ([Willmann et al., 2012](#)). Ces facteurs sont ainsi tous requis pour une CI efficace. Il reste néanmoins à déterminer plus en détail comment la transition entre la pause de la polymérase et son élongation affecte l'activité d'AID ([Feng et al., 2020](#)).

#### 4.4. L'ARN issu de la transcription des régions S

Ces dernières années, un nombre croissant d'études se sont penchées sur le rôle de l'ARN non-codant issu de la transcription des régions S pour expliquer qu'AID soit majoritairement ciblée au locus IgH lors de la CI. La particularité des régions S des mammifères réside dans leurs séquences riches en G dans le brin non-transcrit, ce qui permet la formation de structures secondaires telles que les R-loops et les G-quadruplex ([Pavri, 2017](#)).

##### 4.4.1. Les R-loops

Les R-loops sont des structures hybrides d'ADN/ARN se formant par complémentarité de séquence entre l'ARN généré lors la transcription et l'ADN du brin transcrit, laissant alors le brin non-transcrit sous forme d'ADN simple brin, permettant le ciblage d'AID aux régions S.

Il a été démontré que des « clusters » de G sont importants pour initier la formation de ces R-loops, et qu'une forte densité de G est importante pour leur élongation ([Roy and Lieber, 2009](#); [Zhang et al., 2014b](#)), éléments que l'on retrouve dans les régions S des mammifères. Les régions V en revanche, n'étant pas enrichies en G, ne peuvent pas former de R-loops ([Ronai et al., 2007](#)). Outre l'importance d'un fort enrichissement en G dans l'ARN, les cassures simple-brin ainsi que la transcription, en générant un surenroulement négatif de l'ADN en aval de la polymérase, qui

favorise la séparation des deux brins d'ADN, peut également favoriser la formation de R-loops ([Drolet et al., 2003](#); [Roy et al., 2010](#)). Il y a 30 ans, la transcription *in vitro* des différentes régions S avait déjà permis de montrer l'apparition d'hybrides ADN : ARN, avec l'hypothèse que ceux-ci puissent jouer un rôle dans la CI ([Daniels and Lieber, 1995](#); [Reaban and Griffin, 1990](#)). Ces expériences soulignaient de plus l'importance de l'orientation physiologique et non inversée de la séquence, que l'on sait être importante pour la formation de R-loops. En effet, la formation de R-loops est favorisée lorsque ce sont le brin d'ADN non-transcrit et l'ARN qui sont riches en G et donc le brin transcrit en C, plutôt que l'inverse ([Ratmeyer et al., 1994](#); [Roberts and Crothers, 1992](#)). Plus tard, il a été confirmé que des R-loops peuvent se former *ex vivo* dans des lymphocytes B primaires, permettant d'exposer jusqu'à plus d'1 kb d'ADN simple brin ([Yu et al., 2003](#)).

La contribution exacte des R-loops dans la CI fait cependant toujours débat. Une première étude montre que l'inversion de la séquence S $\gamma$ 1 et donc la transcription d'un ARN riche en C qui compromet la formation des R-loops aboutit à une baisse du niveau de CI vers IgG1 ([Shinkura et al., 2003](#)). Par la suite, d'autres études basées sur l'utilisation de la RNase H capable de dégrader les R-loops observent des résultats différents : alors qu'une étude montre un défaut de près de 50% de CI dans des CH12 exprimant la RNase H à l'aide d'une infection rétrovirale ([Wiedemann et al., 2016](#)), une autre ne montre pas d'effet sur la CI ([Parsa et al., 2012](#)). Enfin, une troisième étude utilisant des souris transgéniques ne montre pas non plus de défaut de la CI, mais observe en revanche une augmentation du nombre de mutations sur le brin transcrit, ce qui soutient l'hypothèse selon laquelle l'ARN des R-loops doit être dégradé pour qu'AID puisse désaminer le brin d'ADN transcrit ([Maul et al., 2017](#)). Il est cependant important de noter que même avec la surexpression de la RNase H, les R-loops ne sont pas complètement abolies, même si on estime que leur fréquence est réduite de 70% à S $\mu$  ([Maul et al., 2017](#)).

Une autre étude a remplacé la région variable d'une souris transgénique par une région dite « core » de S $\mu$  dans son orientation physiologique ou inversée. Le « core » S $\mu$  est une version raccourcie de cette région S, qu'il est plus facile de séquencer, ce qui a permis d'observer qu'il y a la même fréquence et la même signature de mutations d'AID que la séquence de la région « core » de switch, soit dans son orientation normale ou inversée ([Yeap et al., 2015](#)). Encore une fois, la région inversée n'étant supposément pas capable de former des R-loops, ces résultats remettent donc en cause le rôle des R-loops dans le ciblage d'AID. De plus, si l'on part de l'hypothèse que les R-loops sont importantes pour la CI, cela veut dire qu'elles jouent un rôle postérieur à AID, lors de l'induction des DSBs ou lors de leur réparation par exemple.

#### 4.4.2. L'exosome à ARN.

La découverte de l'implication du complexe de l'exosome à ARN dans le ciblage d'AID a permis de comprendre comment AID est capable de cibler à la fois le brin transcrit et le brin non-transcrit ([Basu et al., 2011](#)). Le rôle de ce complexe est de dégrader l'ARN en cours de transcription ainsi que les R-loops, permettant ainsi à AID d'accéder au brin transcrit et de le stabiliser sous forme simple brin ([Basu et al., 2011](#)). De plus, comme cela a déjà été mentionné auparavant, une étude suggère que l'ubiquitine ligase Nedd4 permet de dégrader l'ARN pol II accumulée aux régions S, ce qui permettrait au complexe de l'exosome d'accéder à la bulle de transcription pour y dégrader l'ARN naissant ([Anindya et al., 2007](#); [Sun et al., 2013a](#)). Enfin, l'exosome est également responsable de la dégradation de l'ARN non codant anti-sens issu des sites de transcription divergente et à ce titre, est donc impliqué dans le ciblage d'AID, y compris aux gènes hors-cible ([Pefanis et al., 2014](#)).

#### 4.4.3. L'importance de l'épissage des ARN

Le rôle de l'épissage dans la CI a été mis en évidence grâce à plusieurs études explorant l'effet de la délétion du site donneur de l'épissage d' $\gamma 1$ . Celles-ci ont en effet exposé une abolition de la CI vers IgG1 malgré une transcription active des régions S en absence du site donneur de l'épissage d' $\gamma 1$  ([Hein et al., 1998](#); [Lorenz et al., 1995](#)). Comme l'ont montré différentes expériences de co-immunoprécipitation, AID interagit de plus avec plusieurs protéines en lien avec l'épissage, telles que CTNNBL1, des membres de la famille des heterogeneous nuclear ribonucleoproteins (hnRNPs), le facteur polypyrimidine tract-binding protein 2 (PTBP2) ainsi que SRSF1-3 ([Conticello et al., 2008](#); [Hu et al., 2015](#); [Mondal et al., 2016](#); [Nowak et al., 2011](#); [Singh et al., 2020](#)).

Toutes ces protéines ne semblent cependant pas affecter le ciblage d'AID de la même façon, et parfois différemment entre les régions S et les régions V. Alors que hnRNPK est nécessaire pour l'activité d'AID aux régions S et aux régions V, hnRNPL n'est pas requis pour l'activité d'AID mais semble néanmoins essentielle à la CI dans le modèle CH12 ([Hu et al., 2015](#)). Le facteur d'épissage SRSF1-3 est quant à lui nécessaire à l'HMS dans les cellules DT40 en étant recruté aux régions V et en y régulant l'épissage des transcrits, ce qui permettrait à AID de faciliter son accession à l'ADN simple brin ([Kanehiro et al., 2012](#)). De plus, il est impliqué dans l'accumulation d'AID dans le noyau en interagissant avec son domaine C-terminal ([Kawaguchi et al., 2017](#)). SRSF1-3 interagit également avec la topoisomérase 1 (TOP1), pouvant empêcher cette dernière de résoudre les supertours négatifs induits dans l'ADN lors de la transcription et qui exposent l'ADN sous forme simple brin, bien que cette inhibition n'ait pas été démontrée ([Kumar Singh et al., 2019](#)). Enfin, une autre étude récente a montré que SRSF1-3 jouait un rôle plus indirect dans

l'HMS. En effet, il permettrait tout d'abord de contrôler l'expression d'UBN1, qui est associée au complexe HIRA et qui permet de placer le variant d'histone H3.3 aux promoteurs Ig ([Singh et al., 2020](#)). Ce variant d'histone a été caractérisé comme jouant un rôle dans l'accessibilité du locus Ig au ciblage d'AID en facilitant la génération d'ADN simple brin ([Aida et al., 2013](#); [Romanello et al., 2016](#)). SRSF1-3 permettrait également l'expression du facteur de transcription SATB2, alors capable de se lier aux régions d'attachement de la matrice (MARs de l'anglais Matrix Attachment Regions) des gènes Ig et ainsi faciliter la transcription en promouvant l'ouverture de la chromatine ([Singh et al., 2020](#)). PTBP2 est lui requis pour une CI efficace et de façon intéressante, son knockdown réduit la quantité d'AID associée aux régions S mais augmente celle associée à la région V adjacente ([Matthews et al., 2014](#)). Ainsi, il semblerait que différents facteurs permettent le recrutement d'AID aux régions V (comme SRSF1-3 et GANP) ou aux régions S (comme PTBP2 et les structures secondaires d'ADN/ARN), et que leurs actions pourraient orienter les cellules B vers la CI ou l'HMS dans les centres germinatifs.

Plus récemment, d'autres études ont permis de compléter ce modèle et de donner un rôle aux transcrits des régions S dans le ciblage d'AID lors de la CI. En effet, l'épissage des ARN issus de la transcription du locus IgH conduit à la génération d'une part, de transcrits matures et d'autre part, de transcrits des régions S excisés sous forme de lassos. Or, le knock-down de DBR1, une protéine capable de débrancher ces lassos, entraîne une baisse du recrutement d'AID aux régions S et une baisse de CI ([Zheng et al., 2015b](#)). De façon remarquable, ce phénotype peut être sauvé en exprimant de façon ectopique les transcrits des régions S, mais uniquement quand ceux-ci sont dans leur orientation physiologique et non pas quand celle-ci est inversée ([Zheng et al., 2015b](#)). Une fois les transcrits des régions S débranchés, ceux-ci peuvent alors former des structures secondaires appelées G-quadruplex (G4).

#### **4.4.4. Les structures G-quadruplex (G4)**

Les G-quadruplex (G4) sont des structures secondaires pouvant être adoptées aussi bien par l'ADN que par l'ARN. Ils sont constitués par l'empilement parallèle d'au moins 3 G-quartets, qui sont des plateaux planaires constitués par l'association de 4 résidus guanines stabilisée par un appariement Hoogsteen. Tout comme les R-loops, l'idée que ces structures puissent se former au sein des régions S n'est pas nouvelle ([Sen and Gilbert, 1988](#)). Il existe plusieurs hypothèses quant à la contribution des G4 pour le ciblage d'AID aux régions S. Tout d'abord, il a été proposé que les structures G4 se formant sur le brin non-transcrit aident à stabiliser les R-loops ([Duquette et al., 2004](#)). Puis, une autre étude a exposé que l'hélicase à ARN DEAD-box RNA helicase 1 (DDX1) aide la conversion des G-quadruplex en R-loops pour permettre le ciblage d'AID aux régions S ([Ribeiro de Almeida et al., 2018b](#)). Il reste néanmoins à expliquer comment AID arrive

à être transférée des G-quadruplex vers l'ADN simple brin. Deux récentes études ont révélé d'une part, qu'AID possède une préférence pour les substrats ramifiés tels que les ramifications retrouvées aux extrémités des structures G4 ([Qiao et al., 2017b](#)) et d'autre part, qu'AID nécessite le résidu G133 afin de reconnaître les structures G4. De fait, le mutant G133V est incapable de se lier aux régions S et d'initier la CI, ce qui appuie l'importance de la présence de G4 pour le ciblage d'AID ([Zheng et al., 2015b](#)). De plus, l'association d'AID aux G4 permettrait également son oligomérisation, ce qui augmenterait la densité de mutations générées par celle-ci ([Qiao et al., 2017b](#)). L'ensemble de ces travaux met donc en évidence un important mécanisme de ciblage d'AID aux régions S médié par l'ARN ([Qiao et al., 2017a](#); [Ribeiro De Almeida et al., 2018a](#); [Zheng et al., 2015a](#)). De plus, très récemment, un nouveau modèle de régulation de la CI a été proposé dans lequel, les protéines NME1 et NME2 agissent comme un couple coordonné inhibiteur-stimulateur pour réguler la CI. En effet, l'association des protéines NME aux régions S est favorisée par la formation de structure en G4 par le facteur stimulateur NME2, permettant ainsi la régulation de la CI ([Zheng et al., 2019](#)). Bien entendu, beaucoup de questions concernant le mécanisme précis de ces structures restent en suspens.

Malgré l'apport récent de nouveaux mécanismes en la matière, la raison pour laquelle AID mute préférentiellement les gènes Ig reste à éclaircir. Quoiqu'il en soit, le ciblage génomique d'AID est probablement dû à l'action combinée des mécanismes que nous venons d'évoquer et peut-être d'autres pas encore découverts. L'action de ces mécanismes est parfois redondante, mais permet, dans la majorité des cas, de protéger le reste du génome de l'activité d'AID.

## 5. Le potentiel oncogénique d'AID

Malgré l'existence des mécanismes de ciblage d'AID, celle-ci n'est en réalité pas seulement ciblée aux régions V et aux régions S. En effet, AID peut s'associer à 5 000 loci ([Yamane et al., 2011a](#)) dont seulement 450 gènes ont été trouvés mutés dans des cellules B normales ([Alvarez-Prado et al., 2018](#); [Methot et al., 2018](#)). Il a été révélé qu'AID peut effectivement entraîner la génération de DSBs et de mutations dans des gènes non-Ig ([Pasqualucci et al., 2001](#)), tels que Myc ([Robbiani et al., 2008](#)) et Bcl6 ([Pasqualucci et al., 1998](#); [Shen et al., 1998](#)), parmi bien d'autres. De fait, la mauvaise régulation du ciblage et/ou de l'action d'AID, ainsi que la réparation mutagène des mutations qu'elle induit favorisent le potentiel oncogénique d'AID. En effet, il a été montré que les mutations et les translocations apparaissant dans les gènes non-Ig du fait de l'action d'AID contribuent au développement de lymphomes dans les cellules B matures ([Nussenzweig and Nussenzweig, 2010](#); [Pasqualucci et al., 2008](#); [Rocha and Skok, 2013](#)).



## 5.1. Les mutations hors-cible

Il a été montré que des mutations dans plusieurs proto-oncogènes sont souvent retrouvées dans les lymphomes de cellules B matures ([Pasqualucci et al., 2001](#)). Ces mutations ont été rattachées à l'activité d'AID par le fait qu'elles sont dépendantes de la transcription, et qu'elles apparaissent préférentiellement aux paires de bases C/G (plus particulièrement au niveau des motifs WRC, ce qui correspond aux motifs ciblés préférentiellement par AID lors de l'HMS). Parmi les gènes touchés, on retrouve notamment Bcl6 ([Gaidano et al., 1997](#); [Migliazza et al., 1995](#); [Pasqualucci et al., 1998](#); [Shen et al., 1998](#)), Fas ([Muschen et al., 2000](#)), CD79a et CD79b ([Gordon et al., 2003](#)), aussi bien dans des cellules B tumorales ([Gaidano et al., 1997](#); [Migliazza et al., 1995](#)) que normales ([Pasqualucci et al., 1998](#); [Shen et al., 1998](#)). D'autres gènes, comme c-myc, Pax ou encore Pim1 ne sont mutés que dans les DLBCLs ([Pasqualucci et al., 2001](#)), ce qui montre qu'il existe deux niveaux pour protéger le génome des cellules B durant l'HMS ([Liu et al., 2008](#)). AID peut ainsi participer à la lymphomagenèse par l'introduction de mutations somatiques au niveau des gènes non-Ig. L'ensemble des gènes hors-cible d'AID a pu être révélé par ChIP-seq, ce qui a montré qu'AID est recrutée à de nombreux gènes dans les cellules B activées ([Yamane et al., 2011a](#)). Une étude plus récente a mis en avant près de 275 gènes ciblés par AID ([Alvarez-Prado et al., 2018](#)). Ces études démontrent bien qu'AID a la capacité d'induire des mutations sur l'ensemble du génome, et que celles-ci peuvent menacer l'intégrité du génome des cellules B.

## 5.2. Les translocations

Avant même qu'AID soit découverte, des études génétiques ont révélé la présence de translocations entre les gènes Ig et d'autres gènes dans les lymphomes B ([Küppers and Dalla-Favera, 2001](#)). C'est en réalité la réparation aberrante des DSBs induites par l'action d'AID qui sont la cause de l'apparition de ces translocations. Ainsi, des translocations entre le locus IgH et des proto-oncogènes sont la signature de bon nombre de lymphomes B. Parmi eux, on peut notamment citer le lymphome de Burkitt dans lequel on retrouve des translocations entre IgH et c-myc, mais également les lymphomes folliculaires caractérisés par des translocations entre IgH et Bcl2 ([Küppers, 2005](#)). Or, des études montrent que c'est bien AID qui est à l'origine des cassures générées dans c-myc ([Ramiro et al., 2004](#); [Robbiani et al., 2008](#)) et que les mutations qu'elle génère nécessite l'action d'UNG pour réparer les mésappariements et générer les translocations ([Ramiro et al., 2006](#)). Il est cependant intéressant de noter qu'UNG, ainsi que MSH2/MSH6, semblent jouer également un rôle protecteur du génome des lymphocytes B des GC. En effet, de nombreuses études ont montré que la fréquence de mutations observée est plus élevée en l'absence combinée de ces facteurs ([Alvarez-Prado et al., 2018](#); [Liu et al., 2008](#); [Methot and Di Noia, 2017a](#); [Pérez-Durán et al., 2012](#)). Cette hypothèse du rôle protecteur d'UNG a



notamment fait l'objet de récentes études qui ont pu montrer qu'en effet, UNG protège globalement les cellules B d'une transformation cancéreuse ([Delgado et al., 2020](#); [Safavi et al., 2020](#)). Ainsi, ce double rôle, à la fois promoteur de mutations dans le cadre de la diversification des anticorps mais aussi protecteur du reste du génome reste encore mal compris. Ce phénomène est au cœur de la question qui cherche à savoir ce qui pousse à une réparation mutagène ou à une réparation fidèle, et reste une des questions clés pour la compréhension des mécanismes de diversification des anticorps.

### 5.3. L'expression ectopique d'AID

Il a été observé que la dérégulation d'AID pouvait être une cause de la tumorigenèse. En effet, peu de temps après la découverte d'AID, une étude montrait déjà que l'expression constitutive et ubiquitaire d'AID dans un modèle murin entraîne l'apparition de lymphomes T, d'adénocarcinomes et de micro-adénomes pulmonaires ([Okazaki et al., 2003b](#)). De manière intéressante, ces souris ne développent pas de lymphomes B, soit parce qu'il existe des mécanismes spécifiques protégeant les cellules B de l'action d'AID, soit parce que ces souris meurent avant d'avoir pu en développer ([Okazaki et al., 2003b](#)). Des études postérieures ont permis d'appuyer la 1<sup>ère</sup> hypothèse, en montrant que la surexpression spécifique d'AID dans des cellules B ne suffit pas à entraîner l'apparition de lymphomes B ([Muto et al., 2006](#)), mais entraîne néanmoins l'apparition d'un nombre important de mutations dans certains gènes non-Ig ([Delgado et al., 2020](#); [Robbiani et al., 2009](#)). De plus, l'expression ectopique d'AID est corrélée à plusieurs cancers tels ceux de l'estomac, du pancréas et du foie notamment ([Endo et al., 2008](#); [Kou et al., 2007](#); [Matsumoto et al., 2007](#); [Sawai et al., 2015](#)). Enfin, AID est également impliquée dans la pathogenèse de certains et notamment le DLBCL ([Lenz et al., 2007](#)).

### 5.4. *Kataegis*

Grâce au séquençage génomique d'un grand panel de plusieurs types de tumeurs humaines, un nouveau profil de mutations a été découvert impliquant AID et APOBEC3 ([Alexandrov et al., 2013](#); [Burns et al., 2013](#); [Lada et al., 2012](#); [Nik-Zainal et al., 2012](#); [Taylor et al., 2013](#)). Il a ainsi été observé qu'un grand nombre de mutations peuvent s'observer dans des régions génomiques relativement petites (inférieures à 10 kb) à la suite d'événements appelés tempêtes de mutations ou « *kataegis* ». Ces mutations ont été rapprochées de l'activité d'AID et d'APOBEC par le fait qu'elles consistent majoritairement en des transitions C→T situées sur le même brin d'ADN, ce qui correspond bien au mode d'action processif d'AID. De plus, on retrouverait ces mutations principalement dans les points chauds d'AID de séquence WRCY ([Alexandrov et al., 2013](#)). Après la génération éventuelle de DSBs à la suite de la réparation de ces mutations par le BER ou le

MMR, la résection des brins qui s'ensuivrait permettrait d'exposer encore plus d'ADN simple brin, le substrat d'AID, ce qui amplifierait les mutations observées ([Casellas et al., 2016a](#); [Kruhlik et al., 2007](#); [Shanbhag et al., 2010](#)). Ces *kataegis* sont également à l'origine de réarrangements chromosomiques importants, comme des translocations ou pire encore, l'éclatement et le réarrangement de chromosomes entiers, phénomène appelé chromothripsis et que l'on retrouve habituellement dans les cellules cancéreuses ([Stephens et al., 2011](#)).

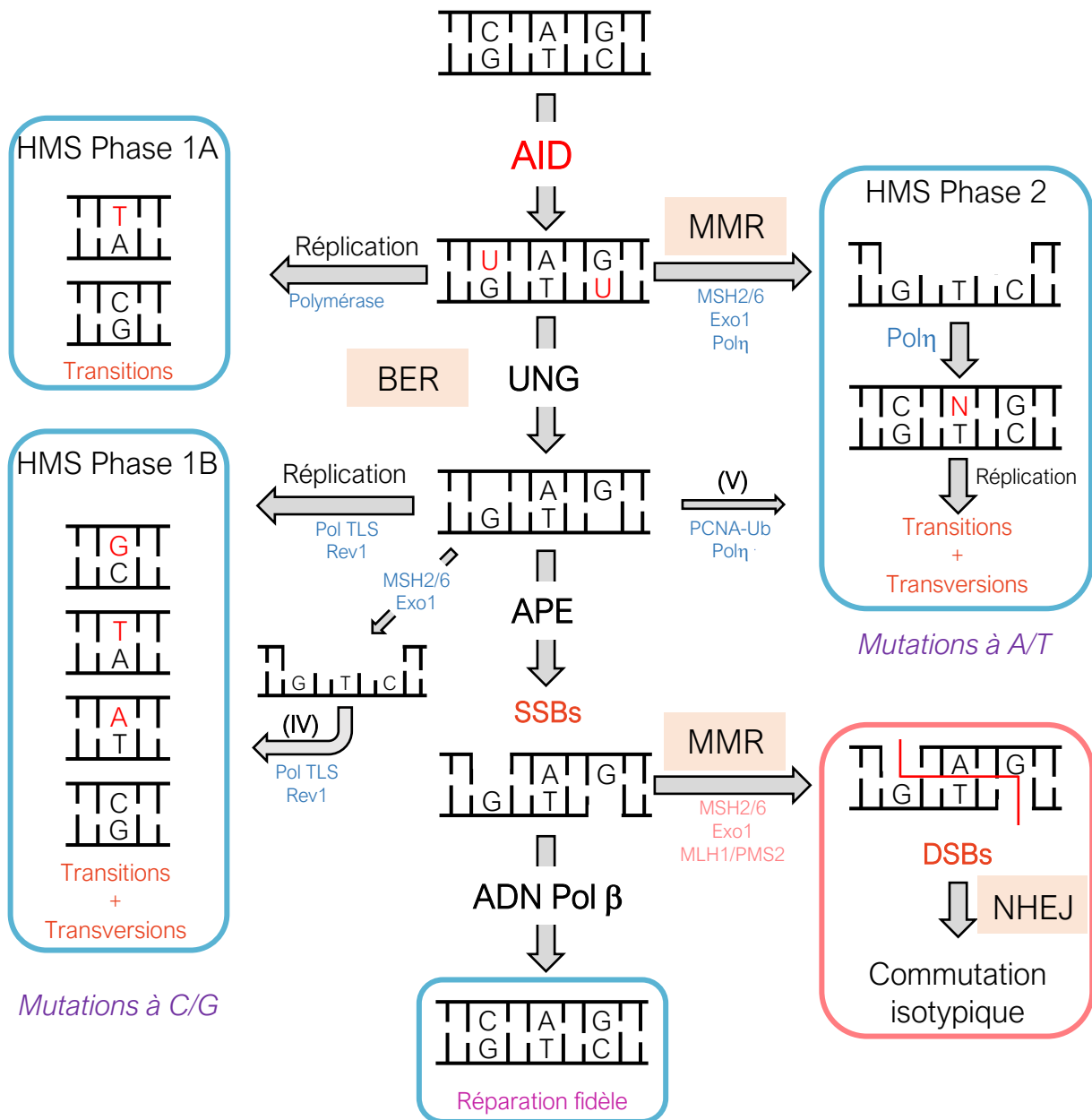
Quoiqu'il en soit, malgré le potentiel oncogénique fort d'AID et sa capacité à générer des mutations et des cassures double brin délétères pour les cellules B, on retrouve nettement plus de mutations et de cassures au niveau des loci Ig qu'ailleurs dans le génome. Cela indique bien qu'il existe des mécanismes de régulation permettant de limiter ce potentiel oncogénique en réparant fidèlement les mutations induites par AID en dehors des loci Ig.

## 6. Le modèle de désamination de l'ADN

Un modèle permettant d'expliquer comment les mêmes désaminations induites par AID dans l'ADN peuvent conduire à la fois à l'apparition de mutations lors de l'HMS et de DSBs dans le cadre de la CI a pour la première fois été proposé par le groupe de M. Neuberger peu de temps après la découverte d'AID ([Petersen-Mahrt et al., 2002a](#)). Ce modèle implique ainsi à la fois des éléments de la réparation par excision de base (BER, de l'anglais Base Excision Repair) et des facteurs de la voie de réparation des mésappariements (MMR; de l'anglais mismatch repair) pour la réparation des uraciles introduits dans l'ADN par AID (**Figure 11**). Ces voies de réparation, habituellement fidèles, sont ici détournées pour induire des mutations ou des DSBs lors de la diversification des Ig. Le profil de mutation associé à la perte de l'enzyme UNG a contribué à disséquer la mutagenèse et les mécanismes de réparation de l'ADN qui résultent des mutations induites par AID et qui conduisent à la recombinaison et aux mutations pour la CI et l'HMS respectivement ([Petersen-Mahrt et al., 2002a](#); [Rada et al., 2002b](#)).

### 6.1. Uracil-DNA-glycosylase (UNG) et son rôle privilégié dans la CI et l'HMS

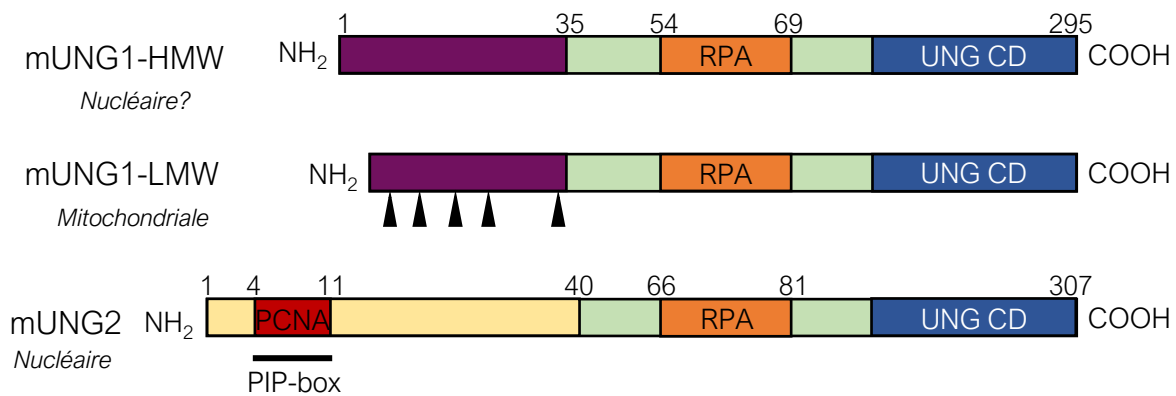
La désamination des C en U étant le dommage aux bases le plus courant, il existe plusieurs glycosylases capables de reconnaître et d'exciser un U dans l'ADN chez les mammifères : UNG, single-strand selective monofunctional uracil DNA glycosylase (SMUG1), thymine DNA glycosylase (TDG) et methyl-CpG-binding domain protein 4 (MBD4) ([Svilar et al., 2011](#)). Elles diffèrent entre elles notamment par leur spécificité de substrat : alors que TDG et MBD4 agissent



**Figure 11: Le modèle de désamination de l'ADN par AID.** AID convertit des cytosines (C) en uraciles (U) dans l'ADN des régions V et des régions S des loci Ig. La réplication sans réparation préalable des U générés peut entraîner l'apparition de mutations de type transitions (HMS phase 1A). Les U générés dans l'ADN peuvent être reconnus par la voie de réparation des mésappariements (MMR), qui génère alors une brèche dans un des deux brins qui sera remplie par la polymérase non-fidèle η, qui introduit des transitions et des transversions principalement aux bases A/T (HMS phase 2). L'uracile peut également être reconnu et glycosylé par UNG, générant ainsi un site abasique pouvant être répliqué par des polymérases translésionnelles (TLS) dont Rev1, ce qui peut aboutir à l'apparition de transitions et de transversions (HMS phase 1B). Ce site abasique peut également être pris en charge par une endonucléase apurinique/aprimidique (APE), qui va alors créer une cassure simple brin (SSB) pouvant être réparée de façon fidèle à l'aide de l'ADN polymérase β. La voie IV, autrement appelée voie hybride UNG+MMR est dépendante des TLS pour générer des transversions vers C/G. La voie V ou bien la voie nc-long patch BER est dépendante d'UNG2 de PCNA-Ub et de la polymérase η. Cette cassure simple brin peut enfin également être convertie en cassure double brin (DSB) à extrémités sortantes si une autre SSB proche située sur le brin opposée est également présente. Cette DSB est prise en charge par le NHEJ pour réaliser la recombinaison nécessaire à la commutation isotypique.

sur de l'ADN double brin, SMUG1 retire les U de l'ADN sous forme simple brin, et UNG est active à la fois sur l'ADN simple brin et l'ADN double brin ([Stratigopoulou et al., 2020](#)). Il s'est très vite avéré que c'est UNG, qui, des 4 glycosylases, est responsable d'exciser les U dans l'ADN lors de la CI et de l'HMS. Cela a notamment été montré par le fait que lorsqu'AID est surexprimé dans *E.coli*, on retrouve nettement plus de mutations dans l'ADN lorsqu'UNG est déficiente plutôt qu'active ([Petersen-Mahrt et al., 2002a](#)), et plus tard confirmé à l'aide d'un modèle murin déficient en UNG dans lequel la CI est fortement réduite et l'HMS affectée ([Rada et al., 2002c](#)). De fait, il a également été montré que MBD4 et TDG ne sont pas impliquées dans les mécanismes de CI et d'HMS ([Bardwell et al., 2003](#); [Imai et al., 2003b](#)). En revanche, SMUG1 pourrait jouer un rôle mineur dans ces processus ([Di Noia et al., 2006](#); [Dingler et al., 2014](#); [Doseth et al., 2012](#)), et même partiellement sauver la déficience d'UNG lorsqu'elle est surexprimée ([Di Noia et al., 2007a](#)). Il est cependant important de noter que l'activité de SMUG1 semble être réduite lors de l'induction des cellules B ([Di Noia et al., 2006](#)).

Il existe en réalité 2 promoteurs dans le gène UNG, ce qui, couplé à de l'épissage alternatif, aboutit à l'expression de 2 isoformes appelées UNG1 et UNG2 ([Nilsen et al., 1997](#); [Otterlei et al., 1998](#)). Ces 2 isoformes diffèrent par leur domaine N-ter mais possèdent le même domaine catalytique et le même domaine d'interaction avec RPA ([Nagelhus et al., 1997](#)). Ainsi, UNG1 a été décrite jusqu'à peu comme étant exclusivement mitochondriale de par la présence d'un signal de localisation mitochondrial (MLS) dans son domaine N-ter unique, alors qu'UNG2 est nucléaire car elle possède un NLS en N-ter ([Otterlei et al., 1998](#)) (**Figure 12**). Ainsi, il était communément admis que c'était l'isoforme UNG2 qui jouait un rôle clé dans la CI et l'HMS. Cependant, une étude récente vient remettre en cause cette assomption, car il a été observé que la déficience spécifique d'UNG2 n'entraîne pas de défaut de CI dans des CH12, prétendument car un variant d'UNG1 serait capable de sauver le défaut de CI entraîné par la déficience spécifique d'UNG2 ([Sarno et al., 2019b](#)). Ainsi, alors qu'aucun des knock-out (KO) individuels d'UNG1 ou d'UNG2 n'entraîne pas de baisse significative de CI, c'est la double délétion d'UNG1 et d'UNG2 qui entraîne une ablation presque totale de l'efficacité de CI ([Sarno et al., 2019b](#)). De fait, les études précédentes permettant de désigner le rôle primordial d'UNG dans les mécanismes de diversification des Ig ont utilisées des modèles murins où un exon commun à UNG1 et UNG2 était supprimé, entraînant ainsi leur double déficience ([Imai et al., 2003b](#); [Rada et al., 2002c](#)). L'expression d'UNG1 aboutit effectivement à l'apparition de 2 variants suite à l'action de la mitochondrial processing peptidase (MPP) qui est capable de cliver le domaine N-ter d'UNG1 ([Bharati et al., 1998](#)). Ces 2 variants ont donc des poids moléculaires différents et sont de fait appelés UNG1-HMW (de l'anglais high molecular weight) qui est la forme non processée et UNG1-LMW (de l'anglais low molecular weight) qui est le variant clivé par MPP. UNG1-HMW est ainsi présente dans des extraits



**Figure 12: Les domaines fonctionnels des isoformes de l'enzyme Uracil-DNA-Glycosylase (UNG).** La transcription du locus UNG permet l'expression de 2 isoformes appelées UNG1 et UNG2. Ces deux isoformes varient par leur séquence N-terminale (rectangle violet pour UNG1 et jaune pour UNG2). UNG1 existe sous la forme de deux variants : un de haut poids moléculaire appelé UNG1-HMW (high molecular weight, 31 kDa) et un de plus faible poids moléculaire, UNG1-LMW (low molecular weight, 26 kDa). La séquence N-ter du variant UNG1-LMW est ciblée par un processus protéolytique conduit par la MPP (mitochondrial processing peptidase). UNG1 et UNG2 ont en commun un site de liaison à RPA (résidus 54 à 69) ainsi que le domaine catalytique UNG (UNG CD) en C-ter de la protéine. Une étude récente a montré que le variant UNG1-HMW se trouve également dans le noyau, tout comme UNG2, et pourrait contribuer à la CI (Sarno et al. 2019). UNG2 possède néanmoins un site de liaison à PCNA, appelé PIP-box, dans sa partie N-ter et que ne possède donc pas UNG1. Adapté de Sarno et al. 2019.

nucléaires de CH12 et constitue jusqu'à 25% de la quantité d'UNG nucléaire, et ce serait donc le variant capable d'exciser des U dans l'ADN nucléaire afin de promouvoir la CI ([Sarno et al., 2019b](#)). Son rôle précis en tant qu'acteur à part entière ou substitut compétent en cas de déficience d'UNG2 reste néanmoins à déterminer.

UNG2 peut interagir avec la sous-unité RPA2 grâce à 2 sites de liaison différents, dont l'un n'est pas encore totalement localisé ([Otterlei et al., 1999](#); [Rodriguez et al., 2017](#); [Weiser, 2020](#)). Elle peut également interagir avec PCNA par un motif PIP-box situé dans son domaine N-ter entre les aa 4 et 11 ([Ko and Bennett, 2005](#); [Rodriguez et al., 2017](#)). Il se pourrait que la liaison d'UNG2 avec RPA2 et PCNA régule son recrutement et/ou son activité au niveau de l'ADN simple brin et double brin respectivement, étant donné que ces facteurs y sont majoritairement associés ([Rodriguez et al., 2017](#); [Sarno et al., 2019b](#); [Torseth et al., 2012](#)). La version humaine d'UNG2 (hUNG2) peut également être régulée par des modifications post-traductionnelles telles que la phosphorylation. Ainsi, la phosphorylation de la thréonine 6 ou de la tyrosine 8 d'hUNG2 peut inhiber son interaction avec PCNA sans toutefois affecter son interaction avec RPA ([Weiser et al., 2017](#)). La phosphorylation de la thréonine 6 et de la thréonine 126 d'hUNG2 entraîne une augmentation de son activité glycosylase *in vitro* ([Lu et al., 2004](#)). Ainsi, la protéine phosphatase 1D (PPM1D) inhiberait le BER en déphosphorylant la thréonine 6 d'hUNG2, possiblement afin de faciliter son inactivation après que l'ADN ait été réparé ([Lu et al., 2004](#)). Cependant, cette conclusion est contredite par des études montrant que la forme non phosphorylée d'hUNG2 est catalytiquement bien active ([Hagen et al., 2008b](#); [Kavli et al., 2002](#); [Torseth et al., 2012](#)). L'expression d'UNG2 est elle aussi régulée tout au long du cycle cellulaire : étant habituellement associée à la machinerie de réplication ([Otterlei et al., 1999](#)), son expression culmine en phase S, mais elle est également présente en phase tardive G1 ([Hagen et al., 2008a](#)), ce qui est en accord avec son rôle dans la réparation des dommages induits par AID pendant cette même phase ([Faili et al., 2002](#); [Schrader et al., 2007b](#); [Sharbeen et al., 2012](#); [Wang et al., 2017](#)). Enfin, l'expression et l'activité d'UNG2 sont fortement induites dans des CH12 et des cellules B primaires murines stimulées pour induire la CI ([Di Noia et al., 2006](#); [Doseth et al., 2011](#); [Shalhout et al., 2014](#)). L'importance de ces augmentations n'est néanmoins pas connue et les mécanismes mis en jeu pour y contribuer non plus. De façon intéressante, il est néanmoins possible de sauver le phénotype de déficience de la CI dans des cellules B Ung<sup>-/-</sup> auxquelles on fait surexprimer des mutants d'UNG2 ne possédant plus qu'environ 1% de leur activité résiduelle ([Begum et al., 2007](#); [Begum et al., 2004](#); [Di Noia et al., 2007b](#)). Ces observations ont mené à l'élaboration d'une hypothèse selon laquelle UNG2 jouerait un rôle non dépendant de son activité de glycosylation dit non-canonique, et servirait ainsi de plateforme de recrutement pour différents facteurs de

réparation ([Begum, 2004](#); [Begum et al., 2007](#); [Yousif et al., 2014](#)). Cependant, des preuves formelles permettant d'étayer ce modèle manquent encore.

## 6.2. La génération de mutations induites par AID durant l'HMS

Plusieurs voies de réparation différentes sont mises en jeu lors de l'HMS afin de prendre en charge les uraciles introduits dans l'ADN par AID. Toutes ces voies ne produisent pas les mêmes types de mutations, et c'est pour cette raison que l'HMS a été découpée en plusieurs phases dans le modèle de désamination de l'ADN ([Petersen-Mahrt et al., 2002a](#); [Rada et al., 2004a](#)) (**Figure 11**). Tout d'abord, il se peut que les uraciles ne soient pas reconnus et réparés avant qu'il y ait réplication de l'ADN par une ADN polymérase lors de la phase S. Dans ce cas, l'uracile est reconnu comme une thymine par la machinerie de réplication qui place alors une adénine sur le brin néoformé, ce qui aboutit à l'apparition de transitions C→T et G→A et correspondant à la phase 1A de l'HMS ([Krokan et al., 2002](#)). Dans la phase 1b, la réplication ayant lieu sur le site abasique créé par UNG peut entraîner des transitions et des transversions aux bases C et G qui sont médiées par les ADN polymérases translésionnelles (TLS) dont Rev1 ([Krijger et al., 2013](#)). Cette dernière, en tant que désoxycytidine transférase, introduit la majorité des mutations de type transversions aux paires de bases C/G en ajoutant préférentiellement un C en face du site abasique ([Jansen et al., 2006](#)).

Les uraciles générés par AID peuvent également être reconnus par les acteurs de la voie non canonique de réparation des mésappariements (ncMMR). Cette voie de réparation aboutit à la phase 2 de l'HMS qui permet l'apparition de mutations aux paires de bases A/T. Elle débute par la reconnaissance du mésappariement dU :dG par l'hétérodimère MSH2/MSH6, qui sert alors de plateforme de recrutement à d'autres facteurs de réparation. Ainsi, un autre hétérodimère, constitué par PMS2 et MLH1, va alors créer une cassure dans un des 2 brins d'ADN afin de permettre l'excision de celui-ci par l'exonucléase Exo1 ([Bardwell et al., 2004](#); [Frieder et al., 2009b](#)). Il est important de noter que le brin excisé n'est pas forcément celui contenant l'uracile. Cependant, plusieurs études ont montré que PMS2 et MLH1 ne sont pas indispensables pour l'HMS ([Chahwan et al., 2012](#); [Ehrenstein et al., 2001](#); [Frey et al., 1998](#); [Quy H. Phung et al., 1999](#); [Van Oers et al., 2010](#); [Zanotti and Gearhart, 2016](#)). Cela fait ainsi écho au fait qu'UNG2 serait capable de créer cet espace nécessaire à l'entrée et à l'action d'Exo1, pouvant ainsi compenser la perte de MLH1 et/ou de PMS2. ([Frieder et al., 2009b](#); [Girelli Zubani et al., 2017](#)). Une brèche va donc être générée par Exo1 ([Bardwell et al., 2004](#); [Frieder et al., 2009b](#)) puis va être à nouveau remplie préférentiellement par l'ADN polymérase translésionnelle  $\eta$  qui insère majoritairement des nucléotides T aux motifs WA/TW ([Delbos et al., 2007](#); [Rogozin et al., 2001](#); [Zeng et al., 2001](#)). C'est cette polymérase qui permet donc de générer des mutations aux paires de bases A/T. Elle



est recrutée grâce à la mono-ubiquitination spécifique de PCNA (Proliferating cell nuclear antigen) à la lysine 164 ([Langerak et al., 2007](#); [Peña-Diaz et al., 2012](#); [Roa et al., 2008](#); [Zlatanou et al., 2011](#)), qui est alors appelé PCNA-Ub.

Enfin, deux autres voies ont été découvertes et mettent en œuvre à la fois des acteurs du BER et des acteurs du ncMMR. Ainsi, la 4<sup>ème</sup> voie nécessite plusieurs éléments pour pouvoir se produire : tout d'abord, il est nécessaire d'avoir 2 uraciles proches l'un de l'autre et l'action combinée de MSH2/MSH6 et d'UNG pour les réparer ([Frieder et al., 2009b](#); [Krijger et al., 2009](#)). La brèche générée suite à l'action d'Exo1, si elle est générée sur le brin opposé au site AP engendré par l'action d'UNG, permet aux polymérases translésionnelles de réaliser des transversions à G/C, en intégrant une base en face du site AP ([Krijger et al., 2013](#)). Enfin la dernière voie implique la voie de réparation longue du BER (lpBER, de l'anglais long patch BER) et a été découverte dans des souris déficientes en MSH2 ([Rada et al., 2004b](#)). Elle nécessite ainsi l'action d'UNG, de PCNA-Ub et de la polymérase  $\eta$  et permet l'apparition de 8 à 20% des mutations aux paires de base A/T ([Ijspeert et al., 2019](#); [Pilzecker and Jacobs, 2019](#)).

### 6.3. La génération de DSBs durant la CI

Tout comme pour l'HMS, les uraciles générés par AID ont la possibilité d'être reconnus par les voies du BER et/ou du MMR. La CI requiert ainsi l'induction de DSBs par ces voies au sein des régions S pour pouvoir réaliser la recombinaison intra-locus nécessaire au changement d'expression de l'isotype ([Figure 11 & 13](#)).

Ainsi, UNG, aussi bien qu'APE1, sont tous deux requis pour la CI ([Imai et al., 2003a](#); [Kavli et al., 2005](#); [Rada et al., 2002c](#); [Schrader et al., 2005a](#)) et leur déficience entraîne une importante baisse du nombre de DSBs dans les régions S ([Imai et al., 2003a](#)). Dans la même logique, la déficience des acteurs du ncMMR entraîne un défaut de CI, bien qu'à des degrés variables selon les facteurs touchés ([Bardwell et al., 2004](#); [Martomo et al., 2004](#); [Péron et al., 2008](#); [R.Ehrenstein and S.Neuberger, 1999](#); [Ranjit et al., 2011](#); [Schrader et al., 1999](#)). Pour confirmer l'apport de ces deux voies, il est à noter qu'une déficience combinée dans des facteurs clés du BER et du MMR entraîne une abolition complète de la CI ([Xue et al., 2006](#)). Le BER est cependant la voie primordiale de la réparation des uraciles dans le cadre de la CI, et Rev1 joue notamment un rôle dans le recrutement d'UNG aux régions S ([Zan et al., 2012](#)). Comme cela a déjà été mentionné, les uraciles peuvent également être pris en charge par MSH2/MSH6, qui recrutent alors MLH1 et de PMS2, puis l'action d'Exo1 va permettre de générer des cassures simple brin, qui peuvent être converties en cassures double brin si deux d'entre elles sont proches et sur les 2 brins opposés de l'ADN. De façon intéressante, Exo1 peut également générer des DSBs à la suite de la



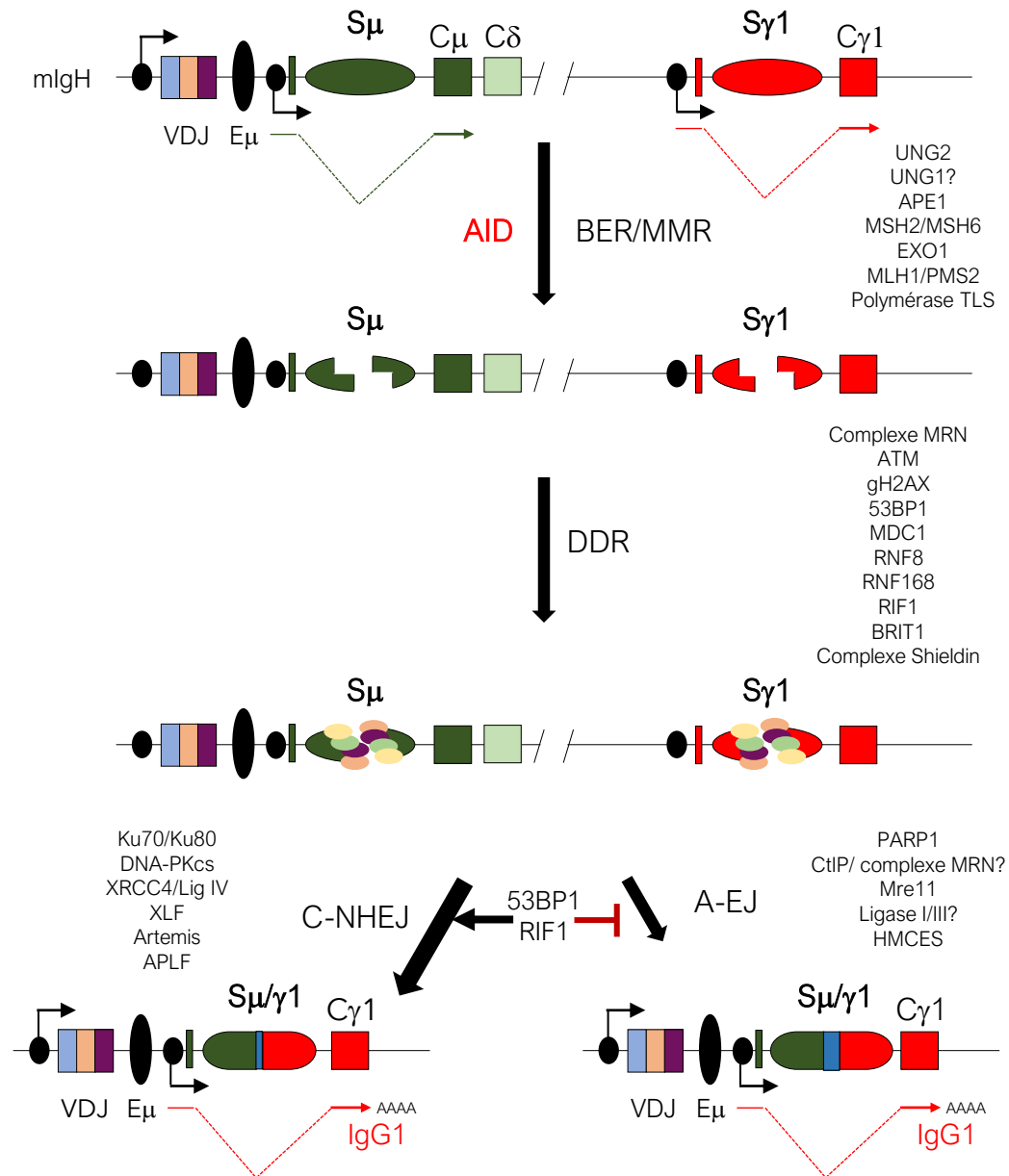
réparation initiée par UNG et APE1, en grignotant un des deux brins à partir de la SSB générée par APE1 jusqu'à arriver à une SSB sur le brin opposé, ce qui permet de créer une DSB avec des extrémités sortantes en 5' ([Schrader et al., 2005b](#)). Ces dernières peuvent également être remplies par les polymérase translésionnelles  $\eta$  et  $\theta$  ([Schrader et al., 2005b](#)). Quoiqu'il en soit, les DSBs activent la voie de réponse des dommages à l'ADN (DDR), qui va alors orienter la réparation de ces DSBs vers la voie classique ou alternative de réparation des extrémités non-homologues (C-NHEJ et A-EJ respectivement), afin de permettre la recombinaison longue distance entre 2 DSBs situées dans les régions S en jeu lors de la CI ([Casellas et al., 1998](#); [Eccleston et al., 2011](#); [Manis et al., 1998](#)).

## 7. La réparation des DSBs durant la CI

### 7.1. La DDR

Peu de temps après la découverte d'AID, il a été montré que des foyers de réparation incluant  $\gamma$ H2AX et Nbs1 pouvaient être observés au niveau des loci IgH des cellules B des GC réalisant la CI. Ceci montre que les dommages à l'ADN et la CI sont liés. De plus, ces foyers sont dépendants de l'expression d'AID, qui joue donc un rôle antérieur à leur formation ([Petersen et al., 2001](#)).

En réalité, lors de l'induction des DSBs, les premiers facteurs à intervenir sont 53BP1, le complexe MRN formé par Mre11, Rad50 et Nbs1, ATM et l'histone H2AX. Va alors se mettre en place une cascade de phosphorylation orchestrée par ATM, qui va notamment phosphoryler l'histone H2AX sur la sérine 139, ce qui constitue la forme plus couramment appelée  $\gamma$ H2AX ([Boboila et al., 2012b](#)). ATM va également phosphoryler 53BP1, un facteur essentiel pour la réparation des DSBs. 53BP1 une fois phosphorylé va pouvoir recruter Rif1 et le complexe Shieldin formé Shld1, Shld2, Shld3 et Rev7, qui, ensemble, vont protéger les extrémités des brins d'ADN d'une résection qui orienterait la réparation vers l'A-EJ ou la recombinaison homologues ([Ghezraoui et al., 2018](#); [Gupta et al., 2018](#); [Liang et al., 2020](#); [Noordermeer et al., 2018](#)). Ainsi, il a été montré que 53BP1 et Rif1 inhibent la réparation par la voie du A-EJ ([Bothmer et al., 2011](#); [Bothmer et al., 2010](#); [Nicolas et al., 2018](#)). Parmi les autres facteurs impliqués, on peut citer MDC1 ([Lou et al., 2006](#); [Stucki et al., 2005](#)) et Nbs1 du complexe MRN ([Falck et al., 2005](#); [Kobayashi et al., 2002](#)). MDC1 une fois phosphorylée va permettre le recrutement de l'E3 ubiquitine ligase RNF8, qui va alors



**Figure 13: Formation et réparation des cassures double brin (DSBs) durant la CI.** Durant la CI, la transcription germinale des régions S donneuse ( $S_{\mu}$ ) et acceptrice (ici  $S_{\gamma 1}$ ) mène à l'exposition de l'ADN sous forme simple brin qui est le substrat d'AID. AID va donc induire la désamination des cytosines dans l'ADN, ce qui entraîne l'apparition de mésappariements dU:dG. Ces mésappariements sont traités par des voies détournées du BER et MMR qui vont permettre la formation de DSBs dans les régions de switch ( $S_x$ ), nécessaires à la recombinaison se produisant lors de la CI. Les DSBs générées vont être tout d'abord reconnues par la voie de réponse des dommages à l'ADN (DDR), puis réparées et liguées par les voies des jonctions non-homologues classique (C-NHEJ) ou bien alternative (A-EJ).

ubiquitiner l'histone H2A, ce qui permet également de recruter RNF168 ([Huen et al., 2007](#); [Kolas et al., 2007](#); [Mailand et al., 2007](#)). Ce dernier est particulièrement important pour la stabilisation de 53BP1 aux DSBs ([Doil et al., 2009](#)).

Tous les acteurs qui viennent d'être cités s'accumulent à proximité de la cassure au sein de structures appelées foci. L'implication de ces facteurs dans la CI a été découverte suite à des inactivations réalisées dans des modèles murins. Ainsi, parmi les facteurs clés identifiés, on retrouve 53BP1 ([Dudley et al., 2002](#); [Manis et al., 2004b](#); [Reina-San-Martin et al., 2007](#); [Ward et al., 2004](#)), ATM ([Pan-Hammarstrom et al., 2003](#); [Pan et al., 2002](#); [Reina-San-Martin et al., 2004](#)), H2AX ([Celeste et al., 2002](#); [Franco et al., 2006](#); [Reina-San-Martin et al., 2003](#)), les facteurs Mre11, Rad50 et Nbs1 formant le complexe MRN ([Dinkelmann et al., 2009a](#); [Kracker et al., 2005](#); [Lahdesmaki et al., 2004](#); [Reina-San-Martin et al., 2005](#)) ainsi que MDC1 ([Lou et al., 2006](#)) ou les E3 ubiquitines ligases RNF8 et RNF168 ([Li et al., 2010a](#); [Ramachandran et al., 2010](#); [Santos et al., 2010](#)). L'absence ou la déficience de ces facteurs entraînent un défaut de CI, bien qu'à des degrés variables, allant ainsi de 30 à 90% de réduction de l'efficacité par rapport à des cellules normales. De plus, ces déficiences augmentent également le risque d'observer des translocations ou autres anomalies chromosomiques ([Bassing et al., 2003](#); [Celeste et al., 2002](#)). On peut également noter la découverte de l'implication de BRIT1, qui, grâce à son domaine BRCT, est recruté au locus IgH indépendamment de H2AX et pourrait servir de plateforme de recrutements à d'autres facteurs de réparation encore inconnus ([Yen et al., 2017](#)).

## 7.2. La réparation des DSBs

Il existe 4 voies majoritaires de réparation des DSBs : la recombinaison homologue, le NHEJ, l'A-EJ (Alternative End Joining) et l'hybridation simple brin (SSA, de l'anglais Single-strand annealing) ([Hustedt and Durocher, 2016](#)). La recombinaison homologue nécessitant une homologie sur une longue séquence, son utilisation est limitée à la phase S et G2 du cycle cellulaire lorsque la chromatide sœur du chromosome touché par la DSB est présente. En revanche, les autres voies de réparation ont lieu tout au cours du cycle cellulaire. Or, il se trouve que la grande majorité des DSBs générées lors de la CI se produisent en G1, ce qui exclut l'utilisation de la HR ([Chaudhuri et al., 2007](#)). De plus, il a été observé que la réparation des DSBs induites lors de la CI présentent des jonctions directes ou avec des homologies courtes (moins de 4 nucléotides) ([Boboila et al., 2010](#); [Yan et al., 2007](#)), ce qui a permis de déterminer que c'est le c-NHEJ qui est la voie majoritaire de réparation utilisée lors de la CI ([Alt et al., 2013](#); [Boboila et al., 2012a](#)).

### 7.2.1. La réparation des DSBs par le C-NHEJ

La réparation par le c-NHEJ démarre lorsque l'hétérodimère Ku70-Ku80 se lie aux extrémités d'ADN de la DSB afin de les protéger (**Figure 13**). Ce duo va permettre le recrutement et l'activation de l'activité kinase de DNA-PKcs (DNA-dependant Protein Kinase catalytic subunit) ([Dobbs et al., 2010](#); [Jette and Lees-Miller, 2015](#); [Jiang et al., 2015](#)). Si les extrémités sont compatibles pour la ligation, l'ADN ligase IV, aidée par le facteur XRCC4, va permettre la ligation des brins d'ADN ([Grawunder et al., 1997](#)). Dans le cas contraire, c'est la nucléase Artémis, recrutée par DNA-PKcs qui est chargée de modifier ces extrémités afin de les rendre compatibles. Elle possède une activité endonucléase capable de résoudre à la fois des extrémités sortantes 5' et 3', sur de l'ADN duplexé aux transitions ADN simple brin/double brin près des DSBs ([Chang and Lieber, 2016](#); [Franco et al., 2008](#)). Parmi les protéines importantes, on retrouve également les facteurs XRCC4-like factor (XLF, également appelé Cernunnos), Paralog of XRCC4 and XLF (PAXX) et APLF qui stimulent l'activité de ligation de l'ADN ligase IV et de XRCC4 ([Ahnesorg et al., 2006](#); [Callebaut et al., 2006](#); [Frit et al., 2019](#); [Graham et al., 2016](#)). Encore une fois, bien que la délétion ou la déficience de ces facteurs entraîne un défaut de CI, celui-ci n'est pas total ([Bjorkman et al., 2015](#); [Casellas et al., 1998](#); [Cook et al., 2003](#); [Franco et al., 2008](#); [Manis et al., 1998](#)). De plus, on observe alors un profil différent au niveau des jonctions de switch, celles-ci présentant une utilisation d'homologies de séquences plus longues et une perte totale de ligation directe amenant l'apparition de jonctions franches ([Pan-Hammarström et al., 2005](#); [Soulas-Sprauel et al., 2007](#); [Yan et al., 2007](#)). Cette voie est en réalité la voie alternative nommée A-EJ, que nous allons détailler dans la prochaine partie.

### 7.2.2. La réparation des DSBs par le A-EJ

Contrairement au c-NHEJ, la voie A-EJ démarre par une courte résection de l'ADN afin d'exposer une micro-homologie de 4 à 20 nucléotides d'un brin par rapport à un autre brin situé à l'autre extrémité de la cassure ([Pannunzio et al., 2018](#)) (**Figure 13**). Même si cette voie est encore mal connue, il semblerait que Parp1 soit tout d'abord impliquée en permettant la reconnaissance des extrémités de l'ADN des DSBs ([Audebert et al., 2004](#); [Robert et al., 2009](#); [Wang et al., 2006](#)). Parp1 permet alors de recruter Mre11 et CtIP pour réaliser la courte résection de ces extrémités ([Bennardo et al., 2008](#); [Dinkelmann et al., 2009b](#); [Lee-Theilen et al., 2011](#); [Rass et al., 2009](#); [Xie et al., 2009](#)). C'est la phosphorylation de CtIP qui permet à Mre11 de pouvoir réaliser cette résection ([Chang et al., 2017](#)). Après résection, les micro-homologies des 2 brins de part et d'autre de la cassure peuvent s'hybrider et alors être ligués par l'ADN ligase I ou III ([Lu et al., 2016](#); [Paul et al., 2013](#); [Simsek et al., 2011](#)). Le choix de l'une ou l'autre ligase est encore mal compris mais il semblerait qu'elles puissent se compenser entre elles ([Lu et al., 2016](#); [Masani et](#)

[al., 2016](#)). Parmi les autres facteurs impliqués, on retrouve HMCES (5-Hydroxymethylcytosine binding, ES-cell-specific) qui permet de protéger les extrémités sortantes d'ADN simple brin et qui semble être requis pour l'A-EJ, comme en témoigne la baisse de l'efficacité de CI et de micro-homologie observée dans les jonctions de switch dans les cellules déficientes pour ce facteur ([Mohni et al., 2019](#); [Shukla et al., 2020](#)). Enfin, Rad52, en inhibant la liaison de Ku70/Ku80 aux DSBs, semblerait faciliter l'utilisation de l'A-EJ ([Zan et al., 2017](#)). Il est ainsi à noter que l'A-EJ est utilisée pour réparer les DSBs induites lors de la CI même dans les cellules dans lesquelles le c-NHEJ est compétent ([Cortizas et al., 2013](#); [Lee-Theilen et al., 2011](#); [Robert et al., 2009](#)). De nombreuses études se penchent sur l'interaction et le choix de réparation entre c-NHEJ et A-EJ, et permettront de mieux comprendre le rôle et l'importance de cette voie et de ses facteurs associés pour le mécanisme de CI.

# RESULTATS

## Chapitre 1 :

### Vers l'identification de nouveaux partenaires d'AID

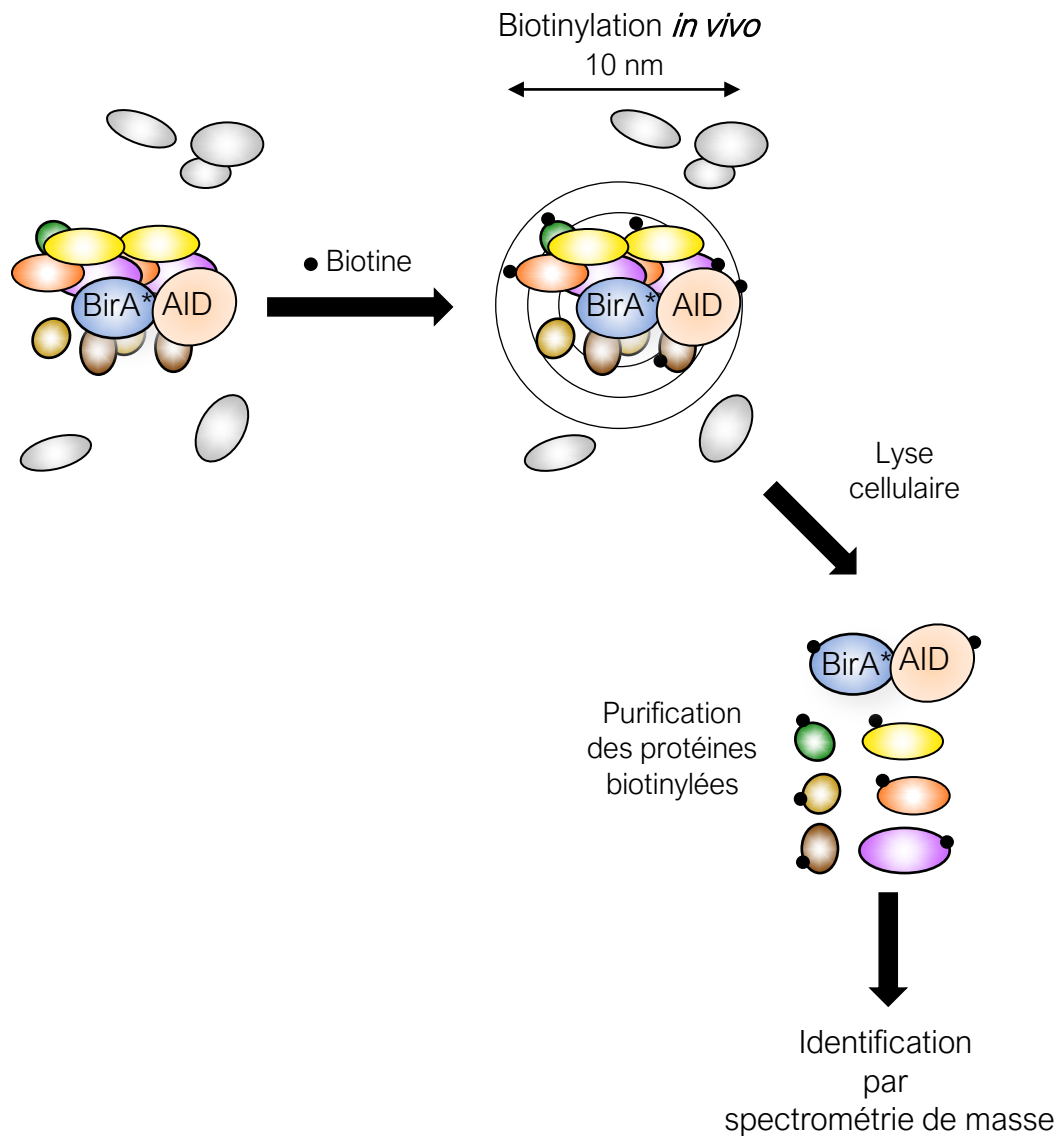
## Hypothèse de travail et objectif du projet

Au cours des réponses immunes, les mécanismes de CI et d'HMS permettent ensemble la génération d'une réponse humorale spécifique, adaptée et durable. De façon extraordinaire, elles sont toutes les deux strictement dépendantes de l'expression et de l'activité d'AID, qui a pour rôle de désaminer des résidus cytosines en uraciles dans l'ADN, générant ainsi des mésappariements qui sont à l'origine des mutations et des DSBs nécessaires à l'HMS et à la CI respectivement. L'HMS et la CI sont donc initiées par des dommages physiologiques dans l'ADN dont la génération et la réparation doivent être finement contrôlés afin de ne pas aboutir à l'apparition de mutations ou de DSBs en dehors des loci Ig et subséquentement aboutir à la transformation cancéreuse des lymphocytes B. Malgré des avancées récentes, les mécanismes protégeant les lymphocytes B de l'instabilité génomique causée par AID en promouvant la bonne régulation et le ciblage spécifique de son activité restent mal compris. De récentes avancées ont été faites pour mieux comprendre la façon dont AID est ciblé à certains gènes et pas à d'autres, mais il est possible qu'il existe encore d'autres mécanismes jouant un rôle dans le ciblage d'AID.

L'objectif de mon projet est ainsi d'identifier l'ensemble des protéines avec lesquelles AID interagit, ce qui devrait nous en apprendre plus sur la façon dont son action est ciblée et régulée dans les lymphocytes B.

Afin d'identifier de nouveaux co-facteurs d'AID, nous avons décidé d'utiliser la technologie de biotinylation de proximité BiID. Celle-ci est basée sur l'utilisation d'une biotine ligase bactérienne mutée (BirA<sup>R118G</sup>, abrégée en BirA\*), qui est fusionnée à une protéine cible (AID dans ce cas) (Figure 14). En présence de biotine, BirA\* est alors capable de biotinyler *in vivo* toutes les protéines qui l'entourent dans un rayon estimé de 10 nm (Kim et al., 2014), et donc les protéines interagissant avec AID. Après une lyse cellulaire, il est alors possible de purifier efficacement les protéines biotinylées à l'aide de billes magnétiques couplées à de la streptavidine qui reconnaît très fortement la biotine, puis de les identifier par spectrométrie de masse (MS ; de l'anglais mass spectrometry) (Figure 14).

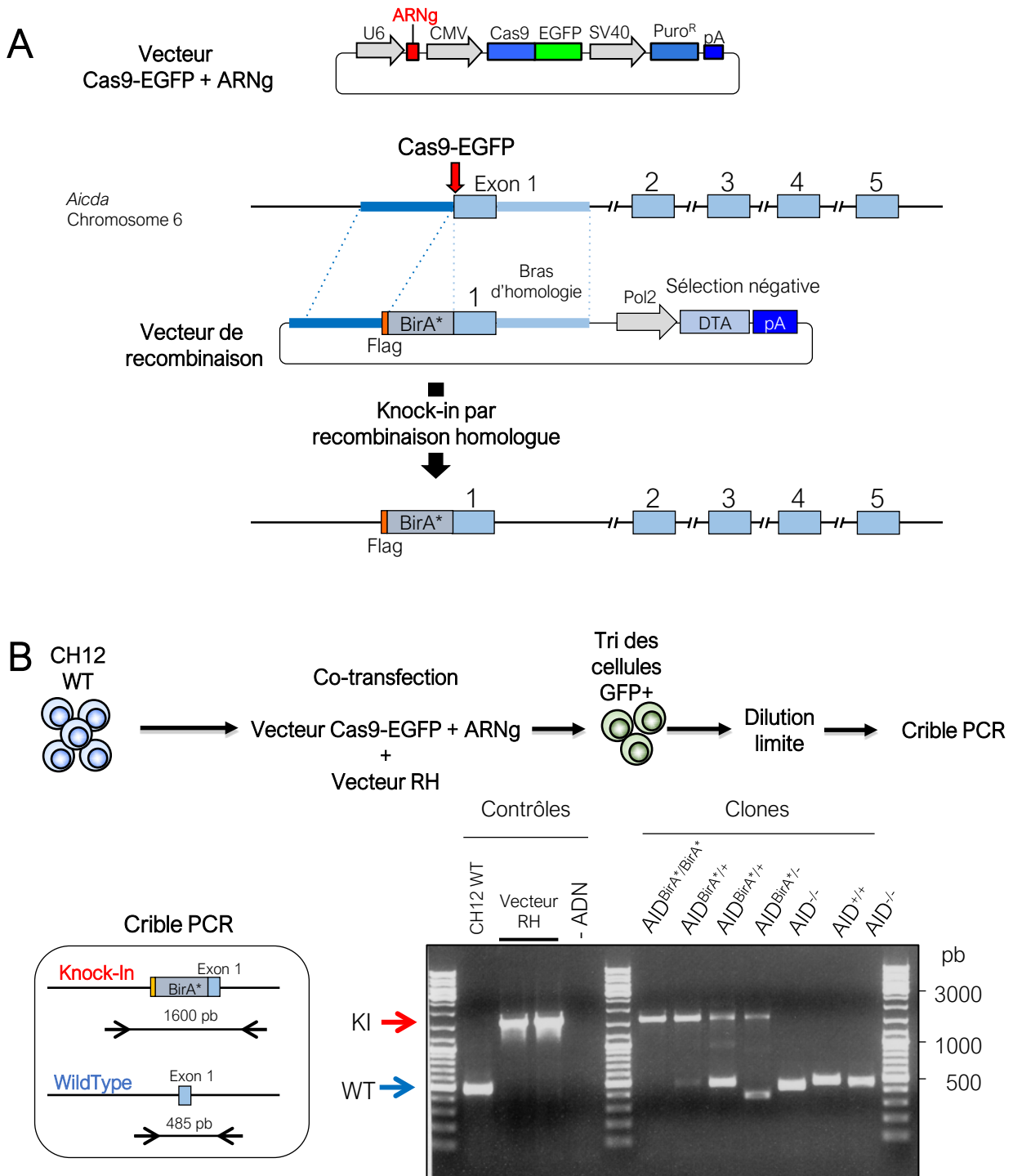




**Figure 14 : Purification et identification des partenaires protéiques d'AID à l'aide de la technologie BioID.** La technologie BioID repose sur l'utilisation d'une biotine ligase mutante (BirA\*), capable de biotinyler toutes les protéines l'entourant dans un faible rayon, en présence de biotine. Ainsi, l'expression d'une protéine de fusion entre AID et BirA\* permet à BirA\* la biotinylation *in vivo* des protéines interagissant avec AID. Après une lyse cellulaire, les protéines biotinyllées peuvent être purifiées à l'aide de billes magnétiques couplées à de la streptavidine, puis identifiées par spectrométrie de masse.

## I- Génération et caractérisation de knock-in de BirA\* dans *Aicda*

Le prérequis principal nécessaire à l'utilisation de la technologie BioID est l'expression d'une protéine de fusion entre BirA\* et la protéine cible. Ainsi, afin d'aboutir à l'expression de la protéine de fusion entre AID et BirA\* à des niveaux physiologiques et ainsi éviter tout biais lié à la surexpression d'une protéine, nous avons opté pour la génération d'un knock-in (KI). Celui-ci a été conçu dans la lignée cellulaire murine CH12, un modèle de choix pour l'étude de la CI dans lequel il est possible d'induire l'expression d'AID par stimulation des cellules avec du CD40L, de l'IL-4 et du TGF- $\beta$  (CIT) ([Nakamura et al., 1996](#)). L'utilisation de ce modèle devrait donc servir à l'identification des partenaires d'AID durant le processus de CI. Etant donné que le domaine C-ter est connu pour être requis pour la CI ([Barreto et al., 2003b](#)) et que nous craignons que l'ajout de BirA\* du côté C-ter d'AID ne perturbe ce domaine, nous avons tout d'abord choisi de réaliser le KI dans l'exon 1 d'*Aicda*. Ainsi, l'insertion de la séquence codante pour Flag-BirA\* juste après le codon d'initiation d'*Aicda* permet en théorie l'expression d'une protéine de fusion Flag-BirA\*-AID (**Figure 15A**). Afin de réaliser ce KI, nous avons fait appel à la technologie Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) - CRISPR associated protein 9 (Cas9) pour générer une DSB ciblée à l'exon 1 d'*Aicda*. Cette cassure peut éventuellement être réparée par recombinaison homologue (RH) en utilisant comme modèle de réparation un vecteur de recombinaison introduit dans la cellule. A ces fins, ce dernier contient la séquence à insérer dans le génome, flanquée de 2 bras d'homologie d'1 kb où se fera la recombinaison nécessaire à l'insertion de la séquence d'intérêt dans le génome de la cellule. Il contient également une cassette de sélection négative tuant les cellules qui l'auraient intégré aléatoirement et en entier par l'expression de la toxine diphtérique A (DTA) ([Yagi et al., 1993](#)). L'autre vecteur requis pour le KI code à la fois pour la Cas9 couplée à l'EGFP et pour son ARN guide (ARNg) ciblant donc son action à l'exon 1 d'*Aicda* (**Figure 15A**). Après la co-transfection de CH12 WT avec ces 2 vecteurs, les cellules GFP+ exprimant Cas9-EGFP ont été triées 24h plus tard puis diluées pour générer des lignées clonales. Ces lignées ont ensuite été génotypées à l'aide d'un crible PCR qui permet de distinguer facilement la présence d'un allèle KI de par la différence de taille de l'amplicon alors visible sur un gel d'agarose (**Figure 15B**). Une autre PCR utilisant des oligonucléotides extérieurs aux bras d'homologie est également réalisée sur les clones positifs pour discriminer ceux qui auraient intégré aléatoirement dans leur génome la séquence de BirA\* et des bras d'homologie malgré la cassette de sélection négative. Enfin, les clones sont séquencés pour valider l'insertion sans mutation et dans le bon cadre de lecture de la séquence codante pour BirA\*. Grâce à un western blot (WB) anti-AID réalisé avec des extraits totaux des clones KI stimulés pour induire l'expression de Flag-BirA\*-AID, on observe bien la présence d'une bande spécifique aux clones

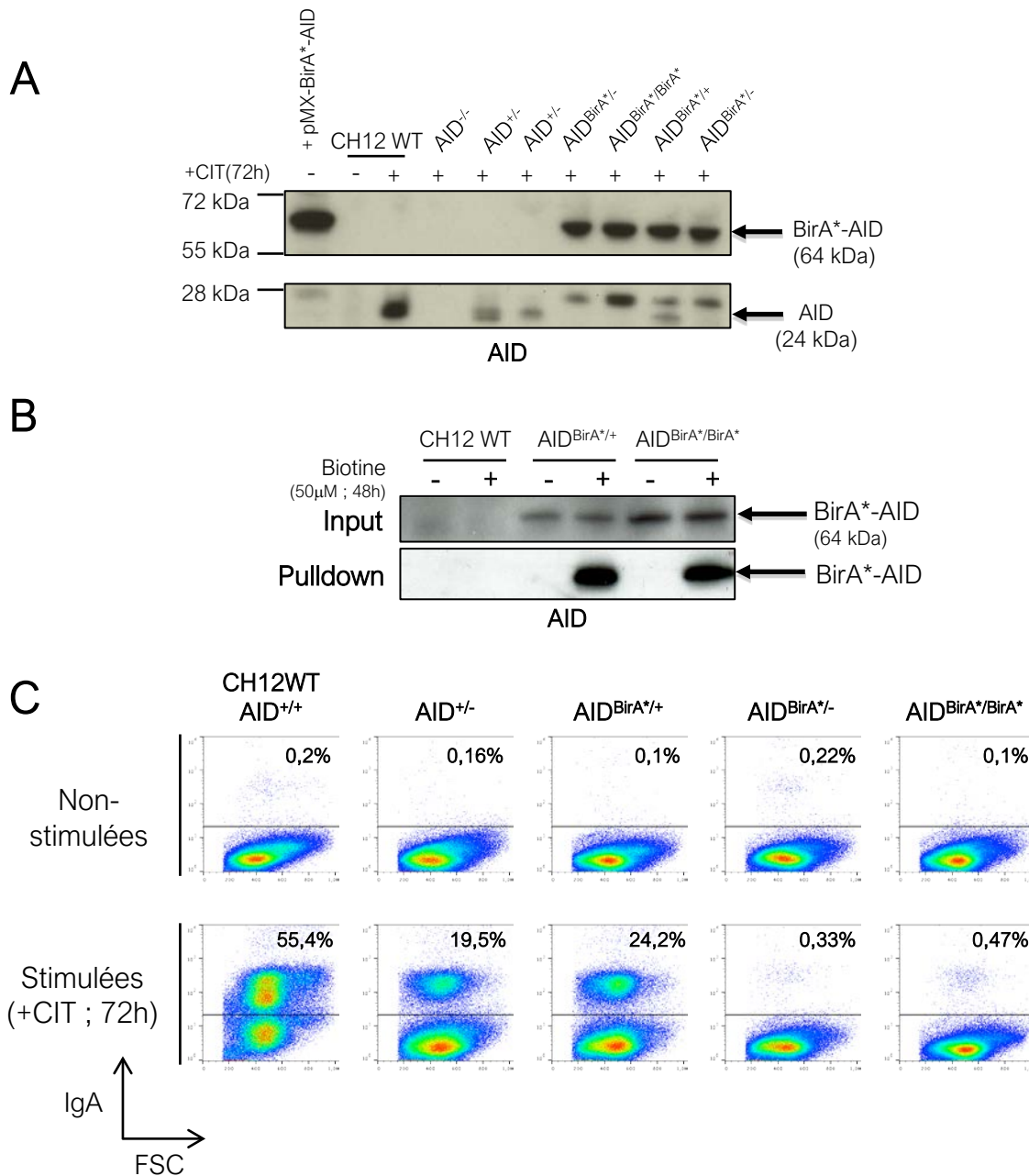


**Figure 15 : Génération et criblage de clones CH12 knock-in pour l'expression de Flag-BirA\*-AID. (A)** Deux vecteurs ont été générés pour réaliser un knock-in (KI) dans l'exon 1 d'*Aicda*. Le premier vecteur permet l'expression transitoire de Cas9-EGFP et d'un ARN guide (ARNg), ce qui permet de réaliser une cassure double-brin qui doit être réparée par recombinaison homologue (RH) en se servant du deuxième vecteur comme d'un modèle de réparation pour permettre l'insertion ciblée de la séquence codante pour BirA\*-Flag. **(B)** Pour réaliser le knock-in, des cellules CH12 wild-type (WT) sont co-transfectées avec le vecteur Cas9 et le vecteur de RH, puis les cellules GFP positives sont triées 24h plus tard avant d'être diluées pour générer des clones qui seront criblés par PCR 10 à 14 jours plus tard. Le crible PCR utilise deux oligonucléotides situés de part et d'autre de l'exon 1, qui permettent la génération d'un amplicon de 1600 paires de bases (pb) quand un allèle KI est amplifié, contre 485 pb pour un allèle sauvage. La taille des amplicons des différents clones est comparée sur un gel d'agarose 1% en utilisant le vecteur de RH et des CH12 WT pour la taille d'un allèle KI et WT respectivement.

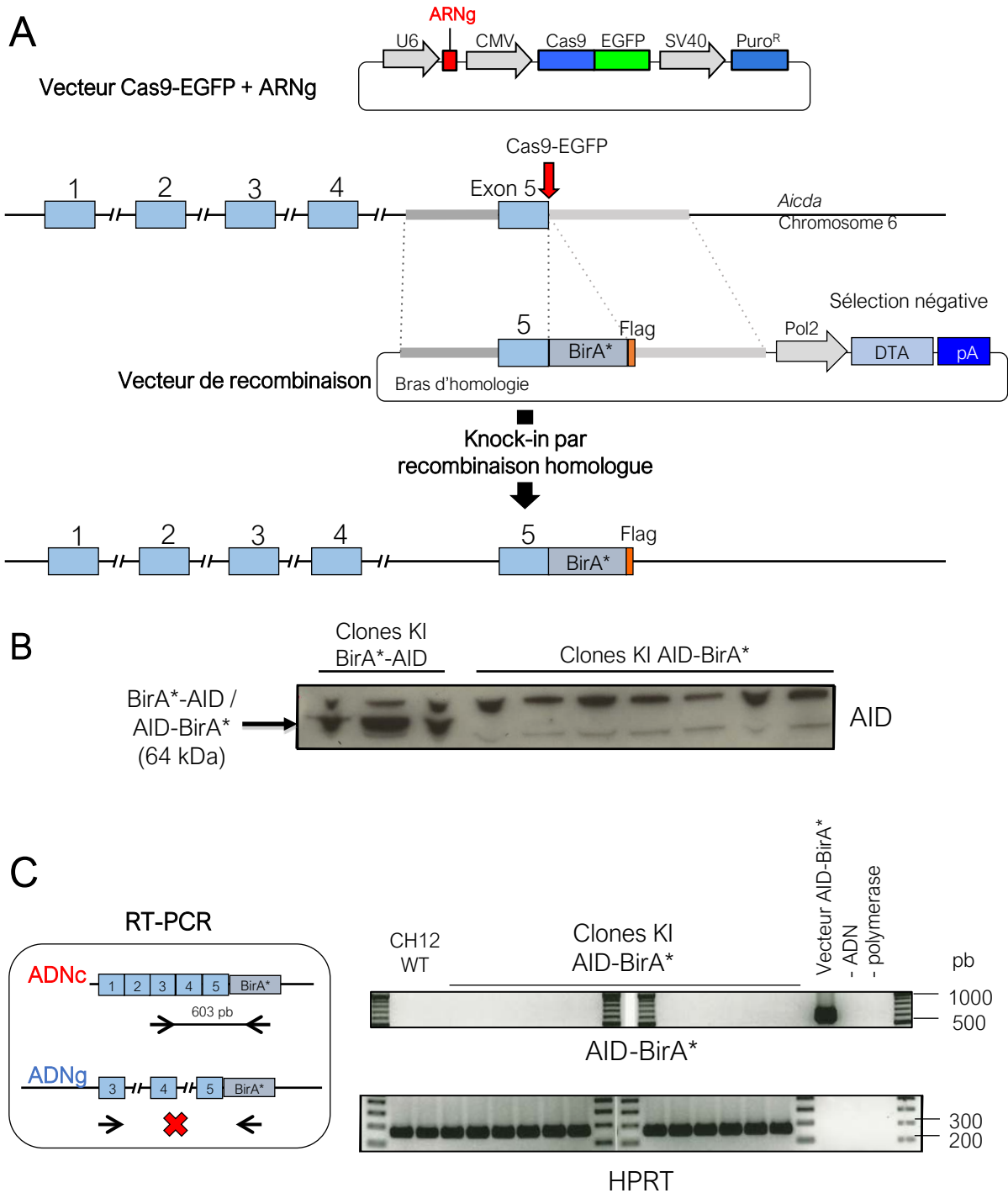
KI au poids attendu de Flag-BirA\*-AID de 64 kilodaltons (kDa). Ainsi, on la retrouve également dans le contrôle positif constitué par des CH12 infectées avec un vecteur rétroviral pMX-PIE contenant la séquence codante pour Flag-BirA\*-AID, mais pas dans les conditions ne possédant pas le KI (**Figure 16A**). Il manque néanmoins un contrôle de chargement (comme une révélation du niveau d'actine  $\beta$  par exemple) pour pouvoir réellement comparer le niveau d'expression entre AID des CH12 WT et Flag-BirA\*-AID des clones KI. Après avoir montré que BirA\*-AID est bien exprimée, j'ai cherché à savoir si BirA\* était bien capable d'entraîner une biotinylation. Ainsi, après avoir stimulé des CH12 WT ainsi qu'un clone KI hétérozygote et homozygote en présence ou en absence de biotine dans le milieu de culture, j'ai réalisé une précipitation des protéines biotinylées à l'aide de billes couplées à de la streptavidine. J'ai alors pu confirmer par WB qu'il était possible de purifier la protéine de fusion auto-biotinylée, mais uniquement lorsque de la biotine est ajoutée dans le milieu de culture, ce qui indique que BirA\* est bien active lorsqu'elle est fusionnée à AID, et que son activité est inductible par l'ajout de biotine dans le milieu de culture (**Figure 16B**). A la suite de leur stimulation, les clones KI ont également été analysés par cytométrie en flux pour leur capacité à réaliser la CI (**Figure 16C**). Malheureusement, tous les clones exprimant uniquement Flag-BirA\*-AID ( $AID^{BirA^*/BirA^*}$  ou  $AID^{BirA^*/-}$ ) sont incapables de réaliser la CI vers IgA. De plus, les clones hétérozygotes exprimant à la fois Flag-BirA\*-AID et AID ( $AID^{BirA^*/+}$ ) ont le même phénotype que des clones  $AID^{+/-}$ , à savoir une efficacité de CI réduite de moitié par rapport à des cellules CH12 WT, ce qui est conforme à l'incapacité de Flag-BirA\*-AID d'initier la CI et à l'haploinsuffisance d'AID (**Figure 16C**) ([Sernandez et al., 2008](#)).

Après une discussion avec J. Di Noia, un collaborateur ayant également utilisé la technologie BioID pour identifier les partenaires d'AID ([Methot et al., 2018](#)), celui-ci nous a indiqué que la fusion d'une protéine en N-ter d'AID l'empêche d'être importée efficacement dans le noyau. Cet effet est possiblement dû au masquage du NLS lui aussi situé en N-ter d'AID. De plus, étant donné que BirA\* pèse près de 34 kilodaltons (kDa), BirA\*-AID pesant 64 kDa est de fait trop grosse pour être importée de façon passive dans le noyau.

Ainsi, face à l'incapacité de BirA\*-AID à initier la CI, nous avons alors décidé de réaliser un KI dans l'exon 5 d'*Aicda* pour aboutir cette fois-ci à la fusion de BirA\* en C-ter d'AID. De façon similaire au KI dans l'exon 1, j'ai co-transfecté des CH12 WT avec un vecteur codant pour Cas9-EGFP et son ARNg ainsi que le vecteur de recombinaison (**Figure 17A**), puis trié les cellules GFP positives et réalisé le crible PCR permettant le géotypage de ces clones. J'ai ainsi obtenu 1 clone homozygote et plusieurs clones hétérozygotes pour le KI dans l'exon 5, qui ont été validés par séquençage. Cependant, le WB anti-AID réalisé avec des extraits totaux de clones KI stimulés ne montre aucune bande au poids de 64 kDa pour les clones KI AID-BirA\*, contrairement aux clones



**Figure 16 : Caractérisation fonctionnelle des clones CH12 knock-in BirA\*-AID. (A)** Analyse par western blot avec un anticorps anti-AID de l'expression de la protéine de fusion BirA\*-AID dans les clones KI. Les extraits utilisés proviennent de cellules stimulées ou non avec du CD40L, de l'IL-4 et du TGF- $\beta$  (+CIT) pendant 72h afin d'induire l'expression d'AID et/ou de BirA\*-AID. Des extraits totaux de cellules CH12 transduites avec un rétrovirus permettant l'expression de BirA\*-AID (pMX BirA\*-AID) ainsi que des extraits totaux de CH12 sauvages (WT) ont été ajoutés comme contrôle positif et négatif, respectivement. Le génotype des différentes cellules analysées est indiqué au-dessus de l'image. **(B)** Analyse par western blot avec un anticorps anti-AID de l'efficacité de biotinylation par BirA\* et de purification de la protéine de fusion BirA\*-AID, avant (Input) et après purification (Pulldown) des protéines biotinylées à partir d'extraits totaux de clones KI précédemment stimulés. **(C)** Comparaison de l'efficacité d'induction de commutation isotypique vers IgA par AID ou BirA\*-AID, dans des cellules CH12 stimulées (+CIT) ou non pendant 72h. Le nombre indiqué dans chaque graphique correspond au pourcentage de cellules IgA positives observé. Le génotype des différentes cellules analysées est indiqué au-dessus des graphiques de cytométrie.



**Figure 17 : Stratégie et caractérisation des clones CH12 knock-in AID-BirA\*.** (A) Schéma de la stratégie du knock-in (KI) de la séquence codante pour BirA\*-Flag dans l'exon 5 d'*Aicda*. De façon similaire au KI réalisé dans l'exon 1, 2 vecteurs sont nécessaires au KI de la séquence codante pour BirA\*-Flag dans l'exon 5 pour aboutir à l'expression d'AID-BirA\*-Flag : 1 vecteur Cas9-EGFP + ARNg pour générer une cassure double brin dans l'exon 5, qui peut être réparée à l'aide du vecteur de recombinaison pour aboutir au KI. (B) Analyse par western blot anti-AID de l'expression de Flag-BirA\*-AID ou d'AID-BirA\*-Flag dans des clones possédant le KI permettant l'expression de ces protéines de fusion. (C) Analyse par reverse transcriptase PCR (RT-PCR) de la quantité de transcrits matures permettant l'expression d'AID-BirA\* dans des clones KI. Seule une amplification à partir d'un ADNc génère un fragment de 603 pb qui est visible sur un gel d'agarose 1%. Des cellules CH12 WT et un plasmide contenant l'ADN complémentaire codant pour AID-BirA\* ont été ajoutés comme contrôle négatif et positif respectivement. Une amplification du transcrit mature codant pour l'hypoxanthine-guanine phosphoribosyltransférase (HPRT) a également été ajoutée comme contrôle du bon déroulement de la RT.

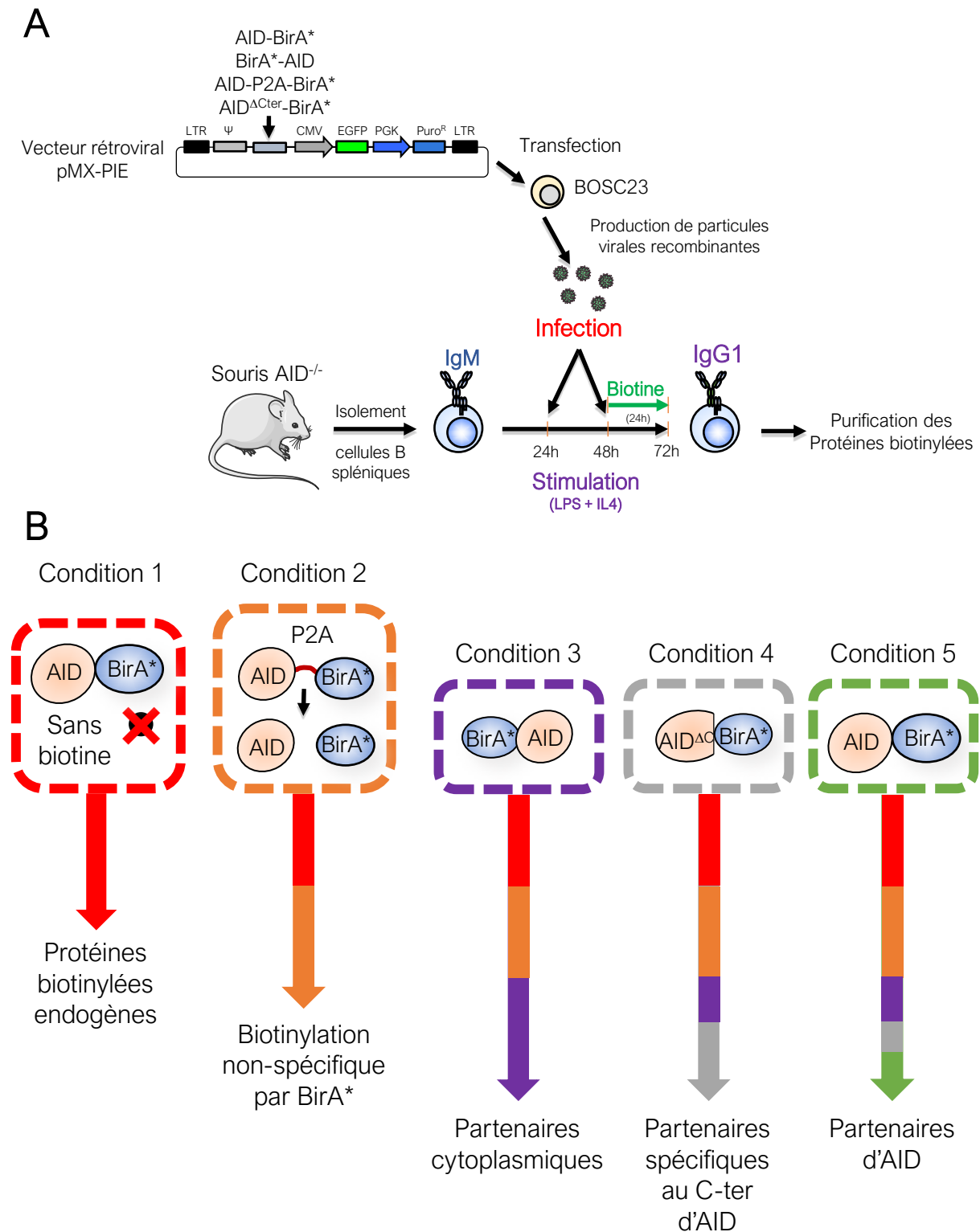
KI BirA\*-AID, alors que les 2 protéines de fusion pèsent le même poids et devraient donc apparaître au même niveau (Figure 17B). Une reverse transcription suivie d'une PCR (RT-PCR) réalisée à partir de l'ARN des clones KI a de plus révélé qu'il n'y a pas de transcrits matures codant pour AID-BirA\*, ce qui indique que le KI de BirA\*-Flag dans l'exon 5 d'AID entraîne probablement un défaut de transcription d'*Aicda*, ou éventuellement un défaut d'épissage du transcrit en ARNm mature (Figure 17C). Nous ne connaissons donc pas la raison pour laquelle ce KI n'est pas exprimé, d'autant plus que cette région du gène n'est pas connue pour être nécessaire à l'expression d'AID, contrairement à la région intronique.

## II- Mise en place du système BioID dans des cellules B primaires murines

Etant donné que la protéine de fusion issue du KI dans l'exon 1 n'est pas fonctionnelle et que le KI dans l'exon 5 perturbe l'expression d'AID, nous avons dû opter pour une nouvelle stratégie et avons alors envisagé d'induire l'expression de la protéine de fusion par transduction rétrovirale (Figure 18A). Nous avons alors décidé de changer de lignée cellulaire et d'utiliser ainsi des cellules B primaires provenant de souris AID<sup>-/-</sup>, pour être dans des conditions les plus physiologiques possibles malgré la surexpression induite par la transduction rétrovirale. De plus, le fait que ces cellules n'expriment pas AID empêche toute compétition entre AID endogène et la protéine de fusion. De plus, cela signifie que ces cellules sont normalement incapables de réaliser la CI en l'état. Cependant, si elles sont infectées avec un rétrovirus leur permettant d'exprimer une version d'AID capable d'initier la CI, il est alors possible de pouvoir observer l'apparition de cellules exprimant des IgG1 de surface, après une stimulation de ces cellules avec du LPS et de l'IL-4 (Figure 18A). Ce système d'infection rétroviral présentant l'avantage de pouvoir utiliser facilement plusieurs conditions en même temps, j'ai alors réalisé le clonage de vecteurs rétroviraux (pMX-PIE) permettant l'expression de BirA\*-AID, d'AID<sup>ΔCter</sup>-BirA\*-Flag (AID<sup>ΔCter</sup>-BirA\*), d'AID-BirA\*-Flag (AID-BirA\*) et d'AID-P2A-BirA\*-Flag (AID-P2A-BirA\*).

D'après les résultats obtenus dans les CH12, la protéine BirA\*-AID devrait rester dans le cytoplasme et ainsi nous permettre d'identifier et de discriminer les partenaires cytoplasmiques d'AID, pour potentiellement identifier plus facilement les partenaires nucléaires d'AID après l'analyse des résultats de MS. La fusion en C-ter d'AID est la condition clé de cette expérience. Si elle est fonctionnelle, c'est elle qui permettra l'identification des protéines interagissant avec AID lors de la CI. AID<sup>ΔCter</sup> est une version mutante d'AID à qui il manque les 10 derniers acides aminés formant le domaine C-ter d'AID. Ce domaine a été identifié comme étant requis pour initier la CI



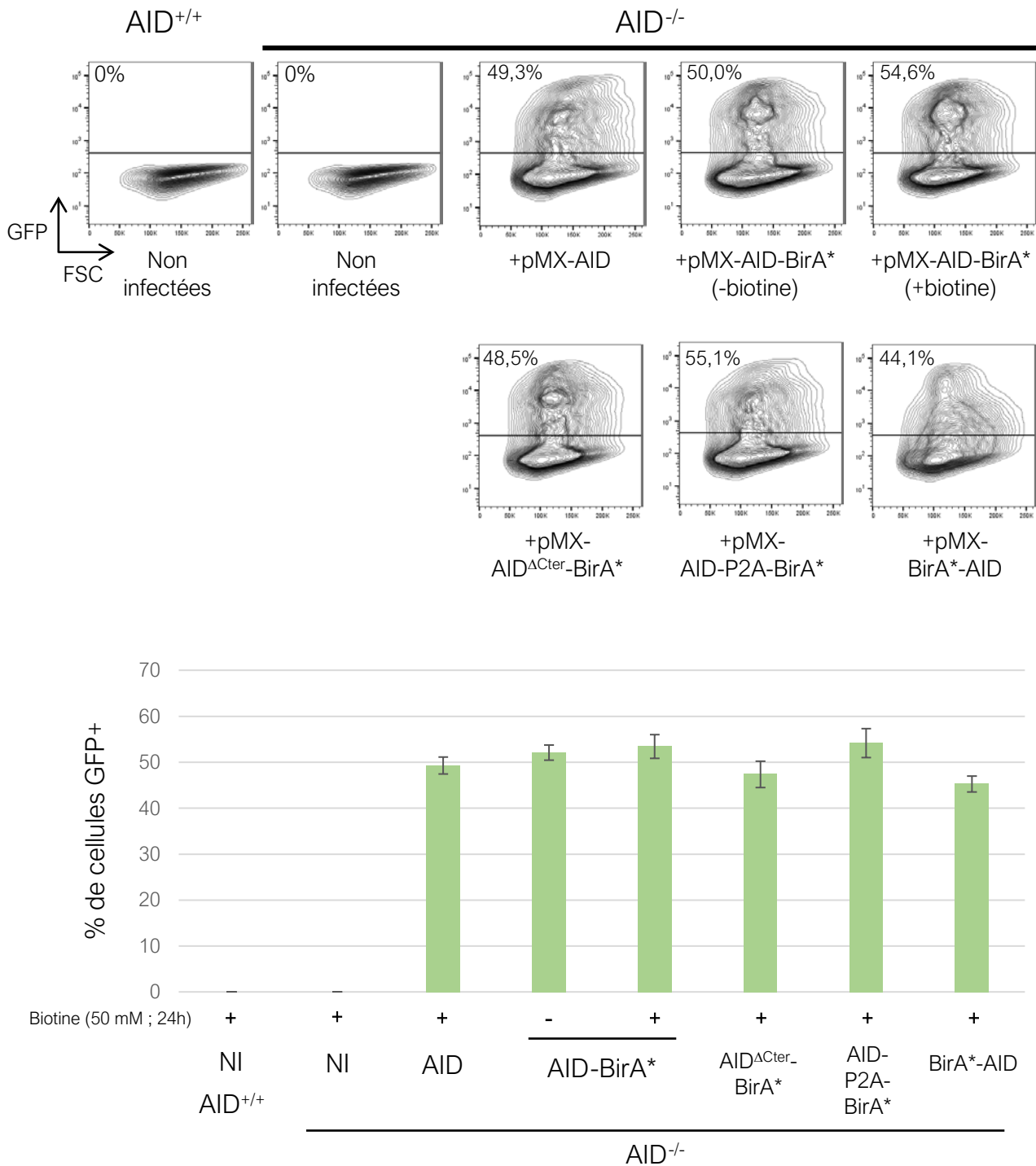


**Figure 18 : La technologie BioID dans des cellules B primaires de souris. (A)** Schéma expérimental pour l'utilisation de BioID dans . Afin de faire exprimer les protéines nécessaires à l'utilisation de BioID dans des cellules B primaires de souris AID<sup>-/-</sup>, des cellules BOSC23 sont transfectées avec les vecteurs rétroviraux (pMX-PIE) contenant les séquences codantes pour AID-BirA\*, BirA\*-AID, AID-P2A-BirA\* ou AID<sup>ΔCter</sup>-BirA\*. Les BOSC23 vont alors produire des particules virales qui serviront à infecter les cellules B primaires afin qu'elles puissent exprimer la protéine encodée par le rétrovirus, qui, si elle est compétente, permettra d'initier la CI vers IgG1. Les cellules sont ainsi stimulées pendant 72h à l'aide de LPS et d'IL-4 et de la biotine est également ajoutée au milieu de culture durant les 24 dernières heures de stimulation afin d'activer BirA\* et de pouvoir purifier les protéines biotinylées. **(B)** 5 conditions expérimentales sont utilisées dans les B primaires : la condition 1 où la biotine n'est pas ajoutée au milieu, ce qui permet, en l'absence d'activation de BirA\*, de purifier les protéines biotinylées de façon endogène dans les cellules. La condition 2, par l'expression d'AID séparée de BirA\* par un peptide « auto-clivant » P2A, va permettre de dissocier AID et BirA\* et ainsi purifier les protéines interagissant uniquement avec BirA\*. Enfin, Les conditions 3,4 et 5 permettront d'identifier et de discriminer les partenaires cytoplasmiques d'AID, les partenaires spécifiques du domaine C-terminal d'AID et tous les partenaires d'AID, par l'expression de BirA\*-AID, AID<sup>ΔCter</sup>-BirA\* et AID-BirA\*, respectivement.

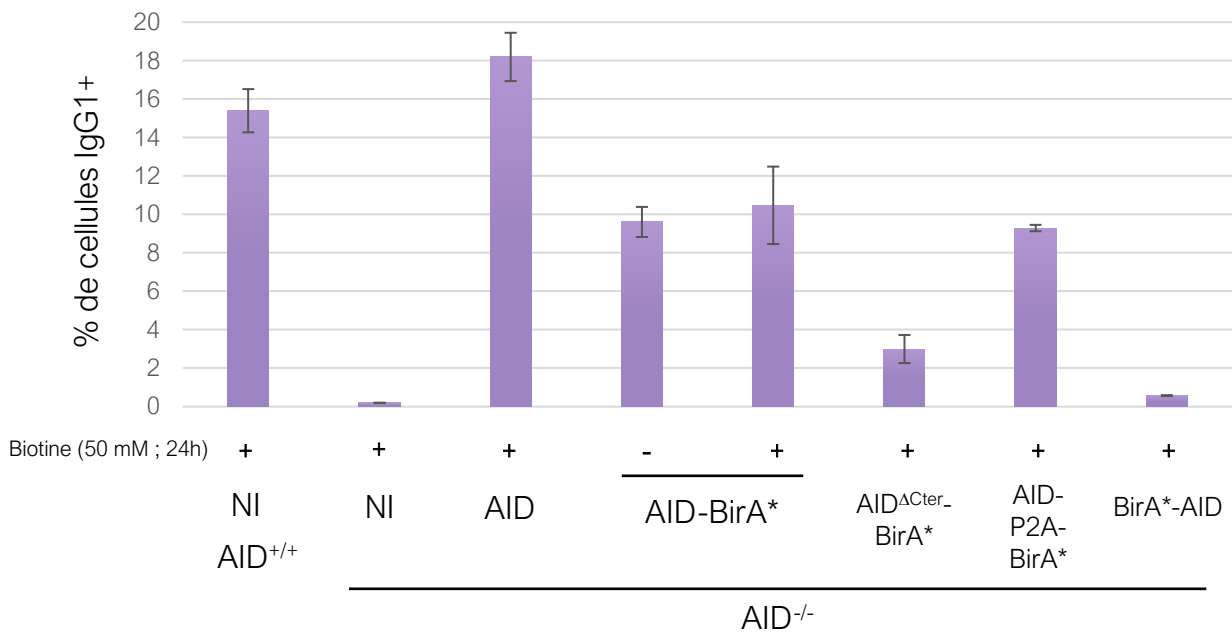
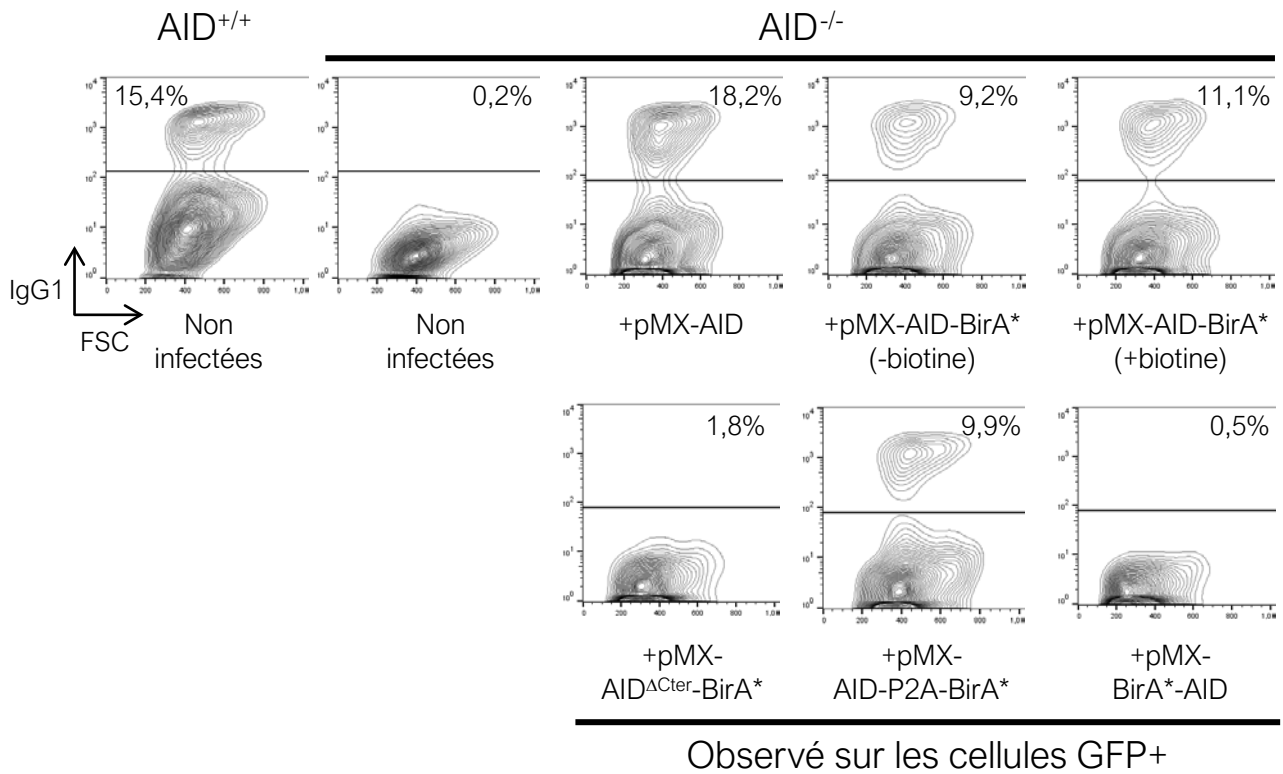


mais pas pour l'HMS, bien qu'il puisse être ciblé à S $\mu$  et y désaminer des cytosines, ce qui permet même la génération de DSBs sans toutefois permettre la CI ([Barreto et al., 2003b](#); [Doi et al., 2009](#); [Ta et al., 2003b](#)). Pour expliquer cette particularité, une hypothèse a donc été émise selon laquelle AID<sup>ΔCter</sup> est incapable de recruter des co-facteurs nécessaires à la réparation des DSBs induites lors de la CI, sans que ces facteurs n'aient cependant été identifiés. Ainsi, nous espérons pouvoir identifier ces facteurs encore inconnus en fusionnant BirA\* à AID<sup>ΔCter</sup>. Enfin, la condition AID-P2A-BirA\* permet de dissocier AID et BirA\* par la présence d'un peptide P2A. Le P2A, pour porcine teschovirus-1 2A, provient donc d'un teschovirus porcine et fait partie de la famille de peptide 2A dits « auto-clivants » ([Szymczak and Vignali, 2005](#)). La dissociation du peptide en 2 parties n'est pas due à un clivage protéolytique, mais résulte d'un phénomène appelé « saut de ribosome » (ribosome skipping en anglais) dans lequel ce dernier, lors de la traduction de la séquence codante du peptide, va passer de la lecture d'un codon à un autre sans toutefois ajouter de liaison peptidique entre une glycine et une proline située vers l'extrémité 3' du peptide ([Donnelly et al., 2001](#)). L'avantage de ce système est de permettre une expression stœchiométrique de plusieurs protéines à partir d'un ARN polycistronique, grâce à un peptide relativement court (22 acides aminés pour le P2A). L'efficacité du « clivage » n'est cependant pas de 100%, mais nous avons choisi le P2A car c'est un des peptides donnant les meilleurs résultats à cet égard ([Kim et al., 2011](#)). Cela implique donc qu'une petite partie des ARNm aboutira à une expression faible de la protéine de fusion AID-P2A-BirA\* non « clivée », mais également à AID-P2A sans BirA\* dans le cas où le ribosome stoppe la traduction après son « saut ». Cette condition permet donc, dans la plupart des cas, de dissocier BirA\* d'AID et d'identifier et de discriminer ainsi les protéines qui auraient tendance à s'associer à BirA\* et qui constitueraient ainsi des faux-positifs présents dans les autres conditions. Enfin, il est primordial d'avoir une condition expérimentale où BirA\* n'est pas présente ou n'est pas activée, afin d'identifier les protéines biotinylées de façon indépendante de BirA\* qui existent dans toute cellule vivante ([Tytgat et al., 2015](#)) (**Figure 18B**).

Du fait de la durée de vie limitée des cellules B primaires *ex vivo* (3-5 jours maximum), il est malheureusement impossible d'enrichir la proportion de cellules ayant intégré le génome viral en utilisant un antibiotique comme moyen de sélection avant de stimuler les cellules, bien qu'une cassette de résistance à la puromycine soit contenue dans le vecteur rétroviral. Ainsi, les cellules B primaires sont infectées à 2 reprises, 24h et 48h après leur mise en culture, afin de maximiser l'efficacité d'infection et faire en sorte qu'un maximum de cellules puisse exprimer la protéine de fusion. Enfin, de la biotine est ajoutée dans le milieu de culture pour activer BirA\* à partir de 48h et pendant 24h (**Figure 18A**).



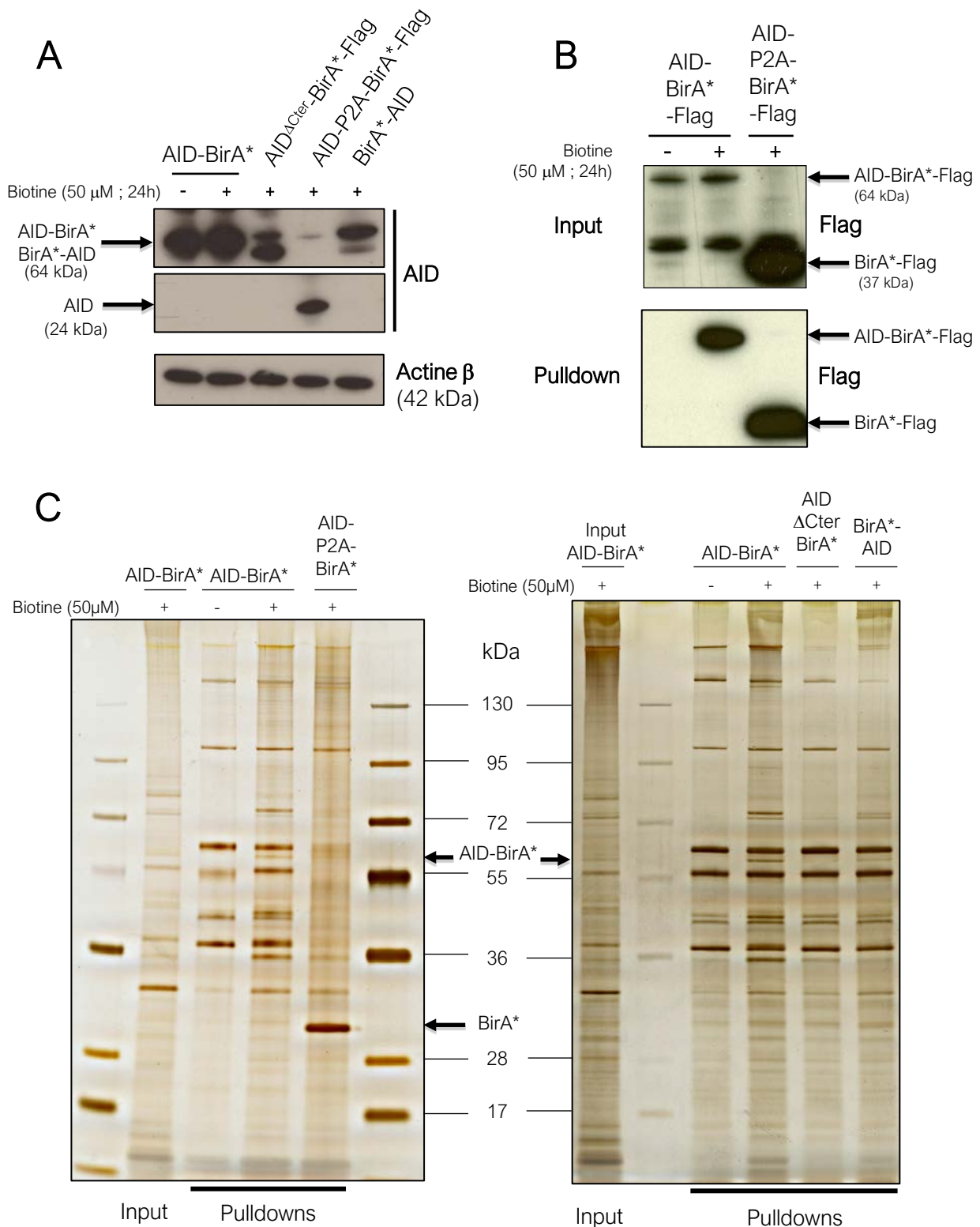
**Figure 19 : Les vecteurs pMX-PIE permettent une transduction rétrovirale efficace des cellules B primaires.** L'efficacité de la transduction des cellules B primaires AID<sup>-/-</sup> a été évaluée en observant le pourcentage de cellules exprimant la GFP encodée par les vecteurs rétroviraux par cytométrie en flux. Des cellules B primaires AID<sup>+/+</sup> et AID<sup>-/-</sup> ont été ajoutées comme contrôle négatif d'expression de la GFP. Les nombres indiqués correspondent au pourcentage de cellules GFP positives. Les données issues de ces analyses sont représentées dans l'histogramme et sont issues de 4 expériences indépendantes. Les conditions NI correspondent aux cellules non infectées, et les différentes constructions encodées par les pMX-PIE ayant servi à infecter les cellules B primaires sont annotées sous les barres du diagramme en bâtons. Les barres d'erreurs du diagramme en bâtons correspondent à l'écart-type standard des données issues de ces 4 expériences.



**Figure 20 : La déficience de CI des B primaires AID<sup>-/-</sup> peut être sauvée par l'expression d'AID-BirA\*.** L'efficacité de CI des cellules B primaires AID<sup>-/-</sup> transduites ou non avec des pMX-PIE (pMX) a été évaluée en observant le pourcentage de cellules exprimant des IgG1 de surface par cytométrie en flux. Pour les cellules transduites, ce pourcentage est observé sur la population GFP positive. Les nombres indiqués correspondent au pourcentage de cellules IgG1 positives. Les données issues de ces analyses sont représentées dans l'histogramme et sont issues de 4 expériences indépendantes. Les conditions NI correspondent aux cellules non infectées, et les différentes constructions encodées par les pMX-PIE ayant servis à infecter les cellules B primaires sont annotées sous les barres du diagramme en bâtons. Les barres d'erreurs du diagramme en bâtons correspondent à l'écart-type standard des données issues de ces 4 expériences.

L'analyse de l'efficacité de la transduction rétrovirale ainsi que de l'efficacité de CI a été réalisée par cytométrie en flux après 72h de culture. On retrouve ainsi une efficacité d'infection d'environ 50% pour tous les pMX-PIE, et la présence ou l'absence de biotine n'a pas l'air d'avoir d'effet sur cette efficacité (Figure 19). On peut néanmoins noter une petite tendance pour les vecteurs codant pour AID<sup>ΔCter</sup>-BirA\* et BirA\*-AID à être légèrement moins bien infectés. Néanmoins, nous avons considéré que cette différence était marginale et permet bien de comparer les conditions entre elles. Une forte différence d'efficacité d'infection aurait logiquement entraîné une différence globale de la quantité de protéines de fusion exprimées dans chaque condition, ce qui rend impossible l'idée d'une analyse semi-quantitative et comparative entre les conditions. Ce problème aurait néanmoins pu être contourné en triant les cellules GFP+ et en égalisant le nombre de celles-ci entre chaque condition pour réaliser la précipitation des protéines biotinylées.

En ce qui concerne l'efficacité de CI, les résultats sont conformes à ce que nous attendions : alors que l'expression de BirA\*-AID ou d'AID<sup>ΔCter</sup>-BirA\* ne permet pas ou très peu de sauver la déficience en CI, celle d'AID-BirA\* et d'AID-P2A-BirA\* le peut, au moins partiellement (Figure 20). Ceci indique donc que la fusion de BirA\* en C-ter d'AID est fonctionnelle du point de vue de la CI. On observe néanmoins une baisse de l'efficacité de CI en comparant l'efficacité d'AID-BirA\* avec celle de cellules B AID<sup>+/+</sup> ou de celles infectées avec le pMX-PIE permettant l'expression d'AID seule. Sachant que l'efficacité de CI est corrélée à l'expression d'AID, ce défaut pourrait provenir d'une différence du niveau d'expression entre les différentes protéines, mais pourrait également venir d'une baisse de la capacité d'AID à initier la CI lorsqu'elle est fusionnée avec BirA\*. AID-BirA\* semble néanmoins être présente en plus grande quantité qu'AID issue de la condition AID-P2A-BirA\* (Figure 21A), bien que l'on observe à peu près la même efficacité de CI dans ces 2 conditions (11,1% vs 9,9%). Ceci laisse donc penser qu'AID fusionnée à BirA\* est moins capable d'initier la CI, et ce, probablement à cause d'un défaut de localisation, de ciblage aux régions S, ou de son activité catalytique. Malheureusement, l'absence d'extraits protéiques provenant de cellules B AID<sup>-/-</sup> infectées avec le pMX-AID ou de cellules B AID<sup>+/+</sup> sur ce WB ne permet pas de confirmer ou d'infirmer cette hypothèse. Ce WB permet également d'observer qu'AID<sup>ΔCter</sup>-BirA\* et BirA\*-AID (bande du haut) sont moins bien exprimés qu'AID-BirA\*, et que la majorité d'AID-P2A-BirA\* est correctement clivée (Figure 21A). Après une purification des protéines biotinylées à petite échelle sur 10 millions de cellules, j'ai également pu confirmer que la biotinylation par BirA\* restait bien activable, comme le montre le fait que la protéine AID-BirA\* n'est précipitée et visible que lorsque de la biotine est ajoutée au milieu de culture (Figure 21B). De façon surprenante, BirA\* a l'air d'être présente en bien plus grande quantité dans la condition AID-P2A-BirA\* que dans la protéine de fusion AID-BirA\*, alors même que l'inverse était observé pour AID dans le WB



**Figure 21 : Les protéines de fusion sont bien exprimées et sont capables d'être purifiées avec leurs partenaires. (A)** Western blot anti-AID évaluant la bonne expression des différentes protéines de fusion encodées par les pMX-PIE. Les extraits totaux de cellules B primaires infectées avec des rétrovirus codant pour les constructions indiquées au-dessus de l'image ont été utilisés pour vérifier l'expression des différentes protéines de fusion en présence ou en absence d'ajout de biotinylation dans le milieu. Le blot d'actine  $\beta$  est utilisé comme contrôle de charge. **(B)** Analyse de l'efficacité de la purification (pulldown) d'AID-BirA\* et d'AID-P2A-BirA\* en présence ou en absence de biotinylation grâce à un western blot anti-Flag. **(C)** Marquages à l'argent d'extraits protéiques de cellules B primaires AID<sup>-/-</sup> transduites avec différents pMX-PIE et cultivées en présence ou en absence de biotinylation, avant (Input) ou après purification (pulldowns) des protéines biotinyliées à l'aide de streptavidine.

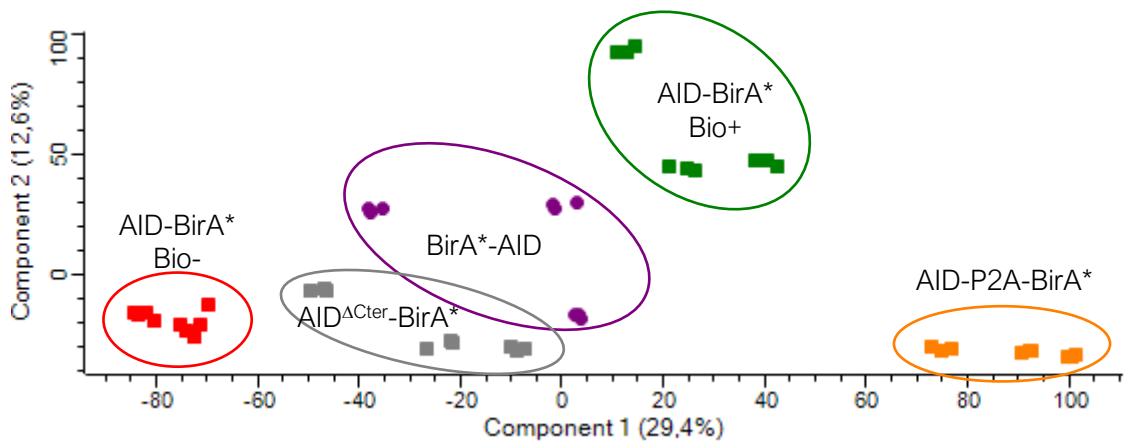
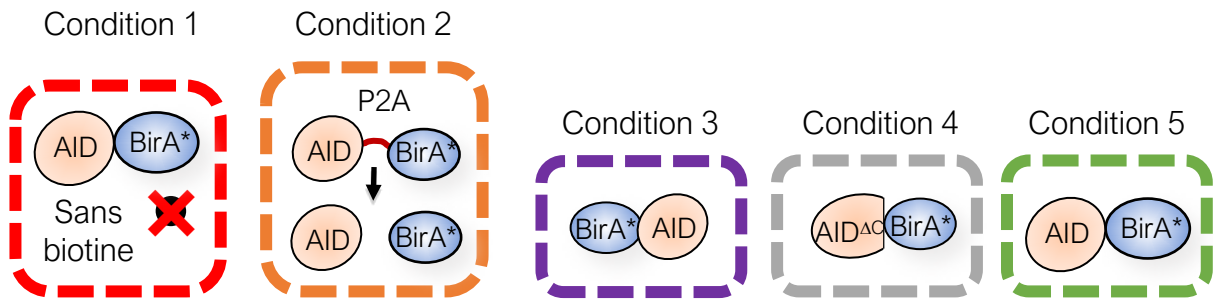
précédent (**Figure 21A**). Le P2A permettant en théorie une expression stœchiométrique d'AID et de BirA\*, la seule hypothèse plausible est que la présence du peptide P2A clivé en C-ter d'AID entraîne d'une manière ou d'une autre sa dégradation, ce qui fait que BirA\* est plus présente qu'AID dans cette condition. Quoiqu'il en soit, on peut également observer que BirA\* est purifiée en plus grande quantité qu'AID-BirA\* (**Figure 21B**). De manière générale, les différences d'expression entre les protéines de fusion contenant BirA\* posent d'importants problèmes pour l'analyse des résultats qui seront mentionnés plus tard dans les résultats et dans la discussion. Enfin, ces différences d'efficacité de purification peuvent également s'observer grâce aux marquages à l'argent des protéines purifiées dans toutes les conditions (**Figure 21C**). En effet, il est bien possible d'observer la présence d'une bande spécifique pour la condition AID-BirA\* +biotine (Bio+) entre les marqueurs 55 et 72 kDa correspondant donc probablement à AID-BirA\* (64 kDa), mais cette bande est bien moins marquée dans les conditions AID<sup>ACTer</sup>-BirA\* ou BirA\*-AID. Les profils des protéines purifiées restent néanmoins similaires dans ces 3 conditions, et se démarquent notamment de la condition sans biotine, de l'input, ou de la condition AID-P2A-BirA\*. Cela est conforme aux résultats attendus et montre que l'expérience a globalement fonctionné.

### III- Analyse des résultats de spectrométrie de masse

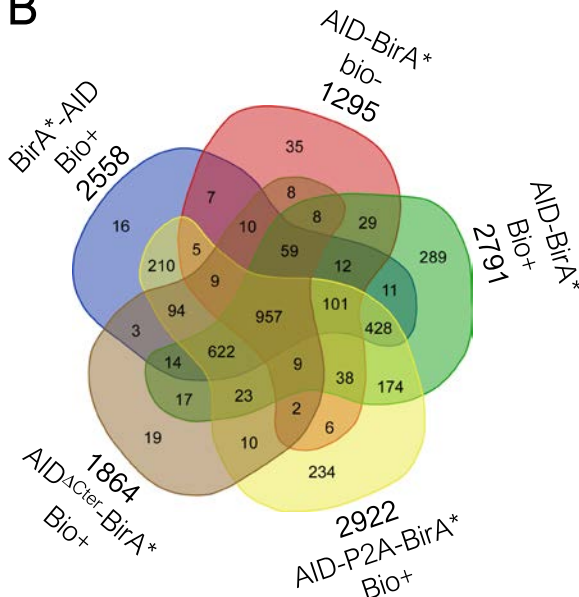
J'ai alors réalisé cette même expérience à large échelle en récoltant 50 millions de cellules par condition avant de les lyser et de purifier les protéines biotinyllées pour leur identification par MS. Trois réplicats biologiques ont été réalisés pour chaque condition. Les données de l'efficacité d'infection et de la CI des différents réplicats biologiques sont incluses dans les figures 20 et 21 respectivement. J'ai utilisé le logiciel Perseus ([Tyanova and Cox, 2018](#); [Tyanova et al., 2016b](#)) pour l'analyse des données obtenues, qui permet de réaliser les tests statistiques standards pour étudier des données de MS. Chaque réplicat biologique de chaque condition est injecté en 3 fois dans le spectromètre de masse, donnant lieu à 3 réplicats techniques. J'ai ainsi éliminé de l'analyse les protéines n'étant identifiées que dans un seul réplicat technique sur 3, afin d'éviter l'analyse de résultats faux-positifs. En réalisant une analyse par composante principale (PCA ; de l'anglais principal component analysis) où chaque point représente les données issues d'un réplicat technique, on peut observer que les réplicats de chaque condition sont assez bien regroupés, à l'exception d'un réplicat biologique de la condition BirA\*-AID qui semble un peu différent des 2 autres (**Figure 22A**). Ceci indique que les résultats obtenus sont assez reproductibles et qu'il est possible de dissocier statistiquement les différentes conditions entre elles. En revanche, si l'on part de l'hypothèse qu'il n'existe que des différences subtiles entre



**A**



**B**



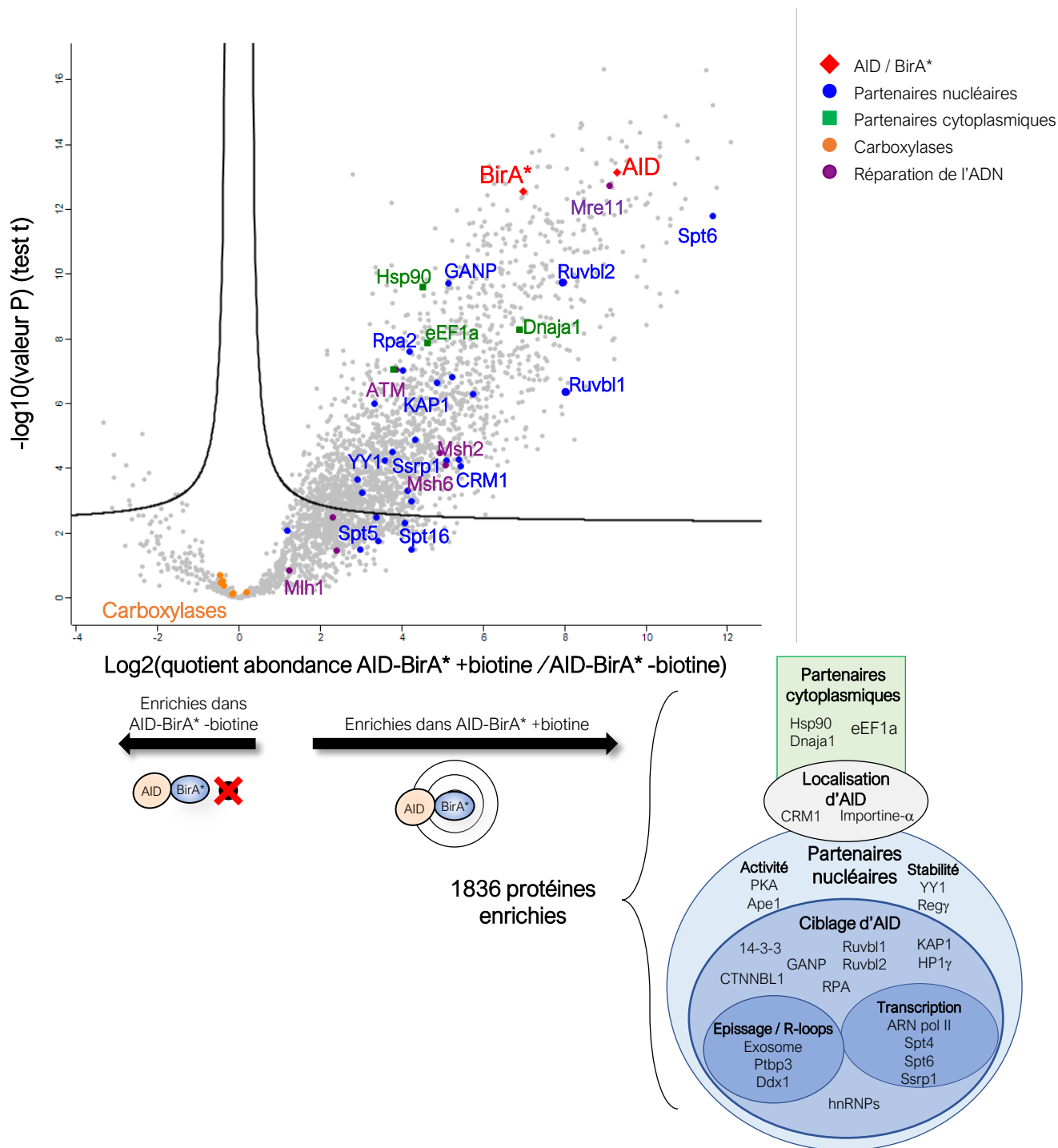
Identifiées dans  $\geq 1$  réplikat biologique sur 3

**Figure 22 : Comparaison des résultats obtenus grâce aux différentes protéines de fusion de BioID. (A)** Analyse par composante principale (PCA) des données d'abondance associées aux protéines identifiées dans les conditions AID-BirA\* - biotine (Bio- ; en rouge), AID-P2A-BirA\* (en orange), BirA\*-AID (en violet), AID-BirA\* +biotine (Bio+ ; en vert) et AID $\Delta$ Cter-BirA\* (en gris). Chaque symbole représente les données issues d'un réplikat technique. **(B)** Diagramme de Venn montrant le nombre de protéines identifiées dans chaque condition (nombre annoté sous le nom de la condition) et leur présence ou absence au sein des autres conditions.

AID<sup>ΔCTer</sup>-BirA\* et AID-BirA\*, on aurait dû s'attendre à les voir beaucoup plus proches l'un de l'autre, ce qui n'est pas le cas ici (**Figure 22A**). Ceci peut probablement s'expliquer par le nombre de protéines identifiées dans chaque condition (**Figure 22B**). On retrouve ainsi nettement moins de protéines dans la condition AID<sup>ΔCTer</sup>-BirA\* que dans les autres conditions où BirA\* est activée, ce qui s'explique par la différence d'expression et de purification observée auparavant (**Figure 21A**). Cette condition est donc à mi-chemin entre la condition sans biotine et les autres conditions avec biotine, conformément à ce qui est observé dans la PCA (**Figure 22A**).

En recoupant les protéines identifiées dans chaque expérience, on observe que la plupart des protéines sont identifiées dans au moins 2 conditions différentes, à l'exception d'environ 200 à 250 protéines exclusives à AID-BirA\* Bio+ et à AID-P2A-BirA\* (**Figure 22B**). Pour AID-P2A-BirA\*, ces résultats peuvent s'expliquer par le fait que BirA\* dissociée d'AID est alors capable de se déplacer plus librement dans la cellule et de biotinyler alors d'autres protéines. Les protéines exclusives à AID-P2A-BirA\* sont donc probablement celles interagissant avec elle qui ne le pouvaient pas quand BirA\* est associée à AID ou sont juste le fait d'un bruit de fond plus élevé généré par BirA\* seule. Enfin, les protéines exclusives à AID-BirA\* Bio+ sont possiblement les partenaires d'AID attendus qui ne s'associent ni à BirA\*-AID ni à AID<sup>ΔCTer</sup>-BirA\*, probablement mêlées à des faux-positifs générés par BirA\* plus exprimé dans cette condition (**Figure 22B**). Ce diagramme de Venn représente ainsi les données qui ont été utilisées pour réaliser des analyses statistiques d'enrichissement des protéines. Ces analyses statistiques sont représentées sous la forme de « volcano plots » dans lesquels l'abondance des protéines identifiées dans 2 conditions sont comparées. La première comparaison entre AID-BirA\* -biotine et +biotine permet donc en théorie de discriminer les protéines biotinylées de façon endogène dans la cellule ainsi que les potentiels contaminants s'accrochant de façon non spécifique aux billes magnétiques couplées à la streptavidine utilisées lors de la purification des protéines (**Figure 23**). Dans cette comparaison, plus de 1800 protéines sont enrichies dans la condition AID-BirA\* Bio+ avec une FDR (False Discovery Rate) inférieure à 5%. Parmi ces 1836 protéines, on retrouve la majorité des partenaires d'AID déjà connus, tels que les partenaires cytoplasmiques Hsp90, Hsp40 (Dnaja1) et eEf1a ([Hasler et al., 2011](#); [Methot et al., 2015](#); [Orthwein et al., 2010](#); [Orthwein et al., 2012](#)), mais aussi beaucoup de partenaires nucléaires impliqués dans l'activité, la stabilisation et le ciblage d'AID, comme des sous-unités de l'ARN pol II ([Nambu et al., 2003](#)), RPA ([Yamane et al., 2011b](#)), des sous-unités de l'exosome à ARN ([Basu et al., 2011](#)) et Ptbp3, un facteur récemment identifié ([Chen et al., 2018](#)) (**Figure 23**). En revanche et de façon surprenante, Spt5, un facteur clé du ciblage d'AID à l'ARN pol II en pause ([Pavri et al., 2010](#)), ne dépasse pas le seuil de significativité. Globalement, même si beaucoup de protéines sont enrichies et une bonne proportion d'entre elles





**Figure 23 : Analyse statistique de l'enrichissement de protéines entre les conditions AID-BirA\* +biotine et AID-BirA\* -biotine.** La comparaison statistique de l'abondance des protéines identifiées dans les conditions AID-BirA\* +biotine et AID-BirA\* -biotine est réalisée à l'aide d'un volcano plot. Les valeurs d'abondance de chaque condition pour chaque protéine font l'objet d'un test t dont la valeur P est corrigée par la méthode de Benjamini-Hochberg. L'inverse du log en base 10 de cette valeur est représentée en ordonnée, alors que le log en base 2 du quotient des valeurs d'abondance d'AID-BirA\* +biotine sur celles d'AID-BirA\* -biotine est représenté sur l'axe des abscisses. Chaque point représente une protéine différente. Le seuil de significativité correspond à une FDR (False Discovery Rate) de 5%. Les différents partenaires déjà connus d'AID enrichis dans la condition AID-BirA\* +biotine sont indiqués dans le schéma en bas à droite.

sont issues du bruit de fond généré par AID, le fait qu'on retrouve la plupart des partenaires d'AID indique que la technologie BioID est bien fonctionnelle et peut permettre l'identification de nouveaux partenaires d'AID. Afin de réduire cette importante liste de protéines enrichies, j'ai également comparé la condition AID-BirA\* Bio+ aux autres conditions (Figure 23). En la comparant avec AID-P2A-BirA\*, le nombre de protéines enrichies baisse de façon drastique à seulement 108 protéines enrichies, et l'on ne retrouve pratiquement plus aucun partenaire connu d'AID, hormis Spt6 et les 2 sous unités d'Hsp-90, qui sont parmi les protéines les plus abondantes dans la condition AID-BirA\* (Figure 24A). Cette baisse importante du nombre de protéines enrichies et la perte des partenaires connus d'AID est très probablement due au fait que BirA\* est plus exprimée dans la condition AID-P2A-BirA\* et génère ainsi un bruit de fond trop important pour que le signal de partenaires moins courants d'AID et donc moins abondamment biotinylés soient statistiquement enrichis dans la condition AID-BirA\*. Ainsi, plus BirA\* est exprimée, plus les protéines qu'elle rencontre sont biotinylées et donc abondamment purifiées puis identifiées par MS. Pour tenter de palier à ce problème, j'ai réalisé une normalisation des données en soustrayant la valeur médiane d'abondance de chaque réplicat technique (Figure 24B). Ainsi, plus les valeurs d'abondance sont élevées au sein d'un échantillon, plus la soustraction sera importante, ce qui permet de rééquilibrer artificiellement les valeurs d'abondance pour chaque protéine. Il est important de noter que cette normalisation ne s'applique pas pour comparer AID-BirA\* +biotine et -biotine, puisque dans cette dernière condition on s'attend à ce que l'abondance des protéines identifiées reste faible (sauf pour certaines carboxylases qui sont biotinylées de façon naturelle). Ainsi, malgré le fait que cette normalisation permette d'identifier 446 protéines comme étant enrichies dans la condition AID-BirA\*, les partenaires nucléaires d'AID n'en font toujours pas partie. Dans la même logique, j'ai réalisé les volcano plots afin d'identifier les protéines enrichies dans la condition AID-BirA\* par rapport à BirA\*-AID (Figures 24C et D). Ici encore, les résultats ne sont malheureusement pas ceux attendus. Le but de la comparaison entre ces 2 conditions était d'identifier spécifiquement les partenaires nucléaires d'AID discriminant les partenaires cytoplasmiques communs aux 2 conditions. Or, en utilisant les données non normalisées, il s'avère que parmi les 584 protéines enrichies, aucune ou presque n'est un partenaire nucléaire déjà connu d'AID, alors même que 2 partenaires cytoplasmiques sont enrichis alors qu'ils ne devraient logiquement pas l'être (Figure 24C). De plus, AID et BirA\* sont aussi statistiquement enrichies alors qu'elles devraient être présentes au même niveau dans les 2 conditions, ce qui corrobore l'hypothèse selon laquelle AID-BirA\* est exprimée à des niveaux plus élevés que BirA\*-AID, et que cela fausse les résultats obtenus. La normalisation des données réduit de fait le nombre de protéines enrichies dans AID-BirA\* en le faisant passer à 260 protéines, mais AID et BirA\* restent toujours enrichies, ce qui démontre que cette méthode de normalisation a aussi ses

Figure 24 : Analyse statistique de l'enrichissement de protéines identifiées dans les conditions AID-BirA\* +biotine, AID-P2A-BirA\*, BirA\*-AID et AID<sup>ΔCter</sup>-BirA\*.

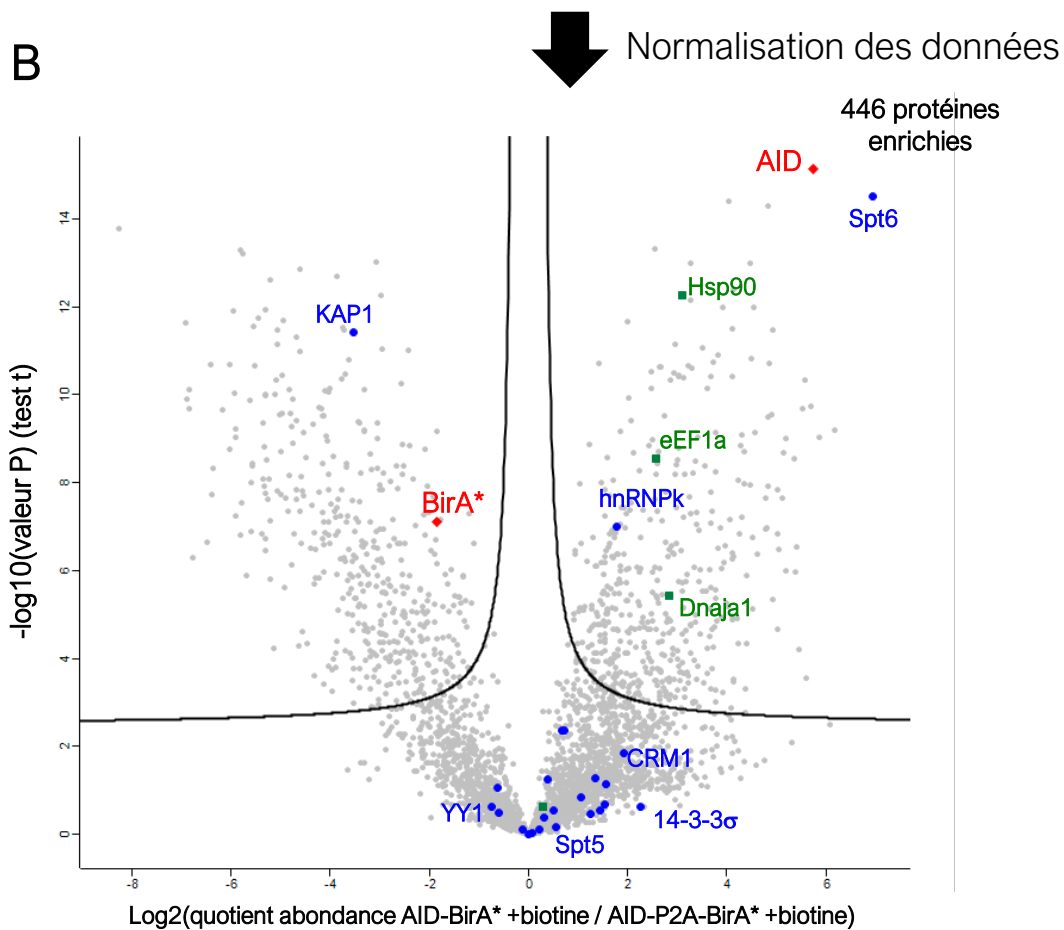
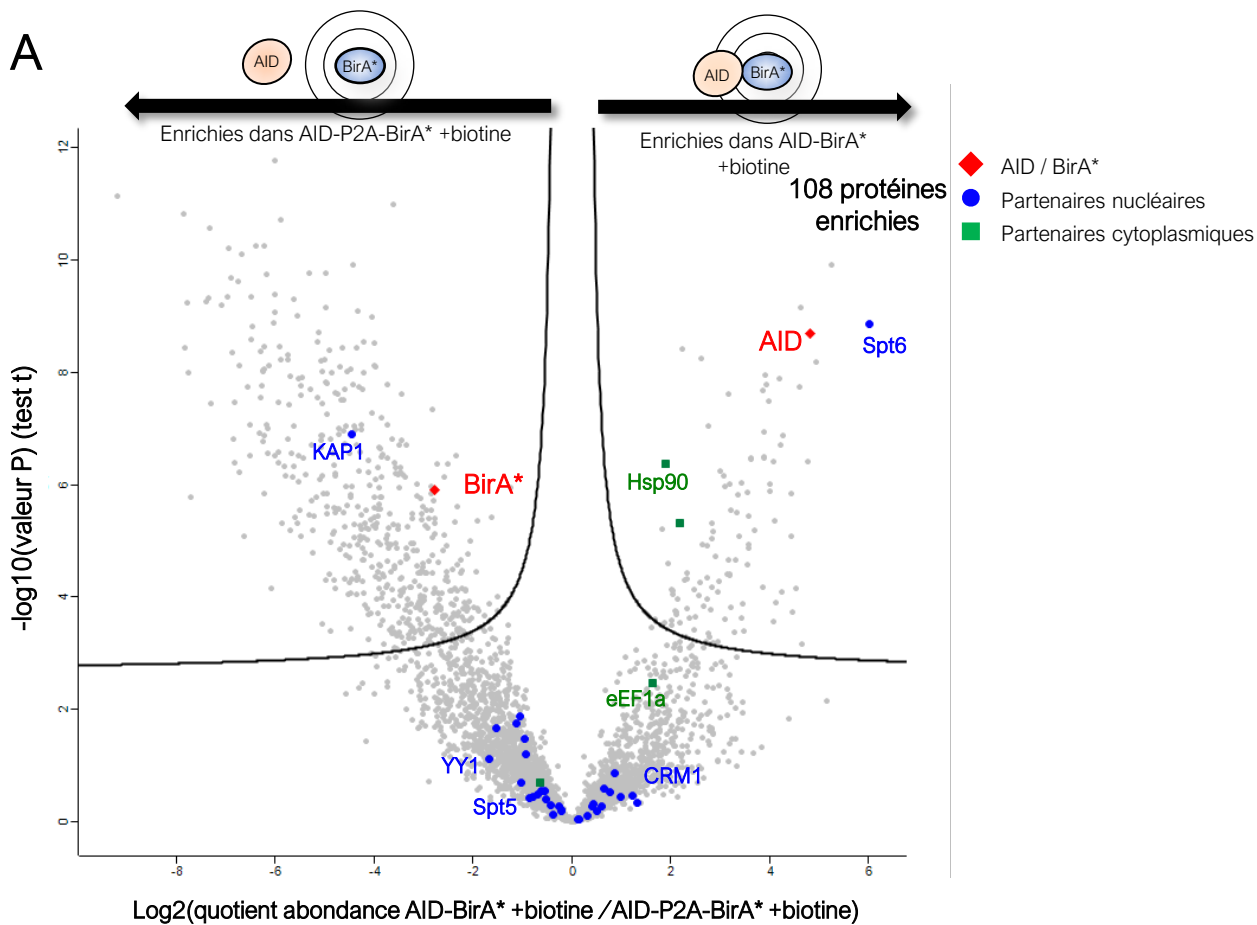
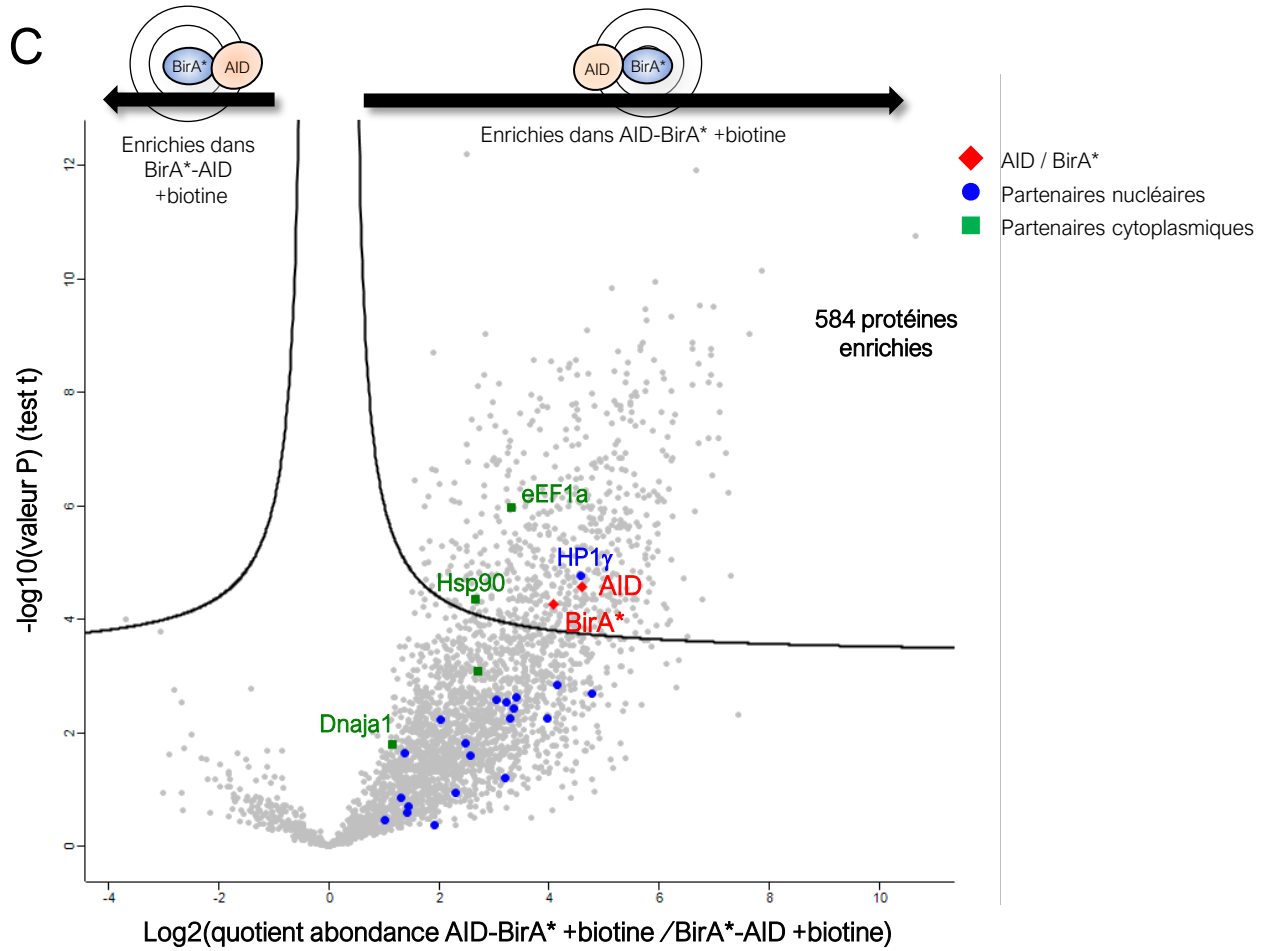


Figure 24 : Analyse statistique de l'enrichissement de protéines identifiées dans les conditions AID-BirA\* +biotine, AID-P2A-BirA\*, BirA\*-AID et AID<sup>ΔCter</sup>-BirA\*.



**D**

Normalisation des données  
(soustraction de la médiane)

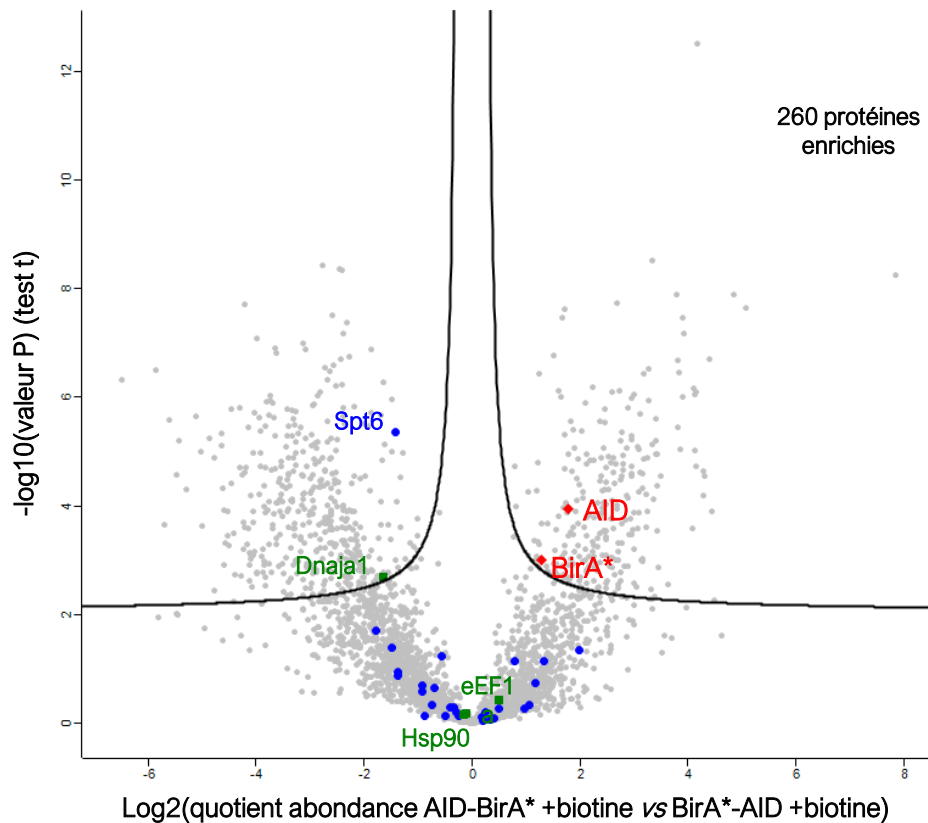
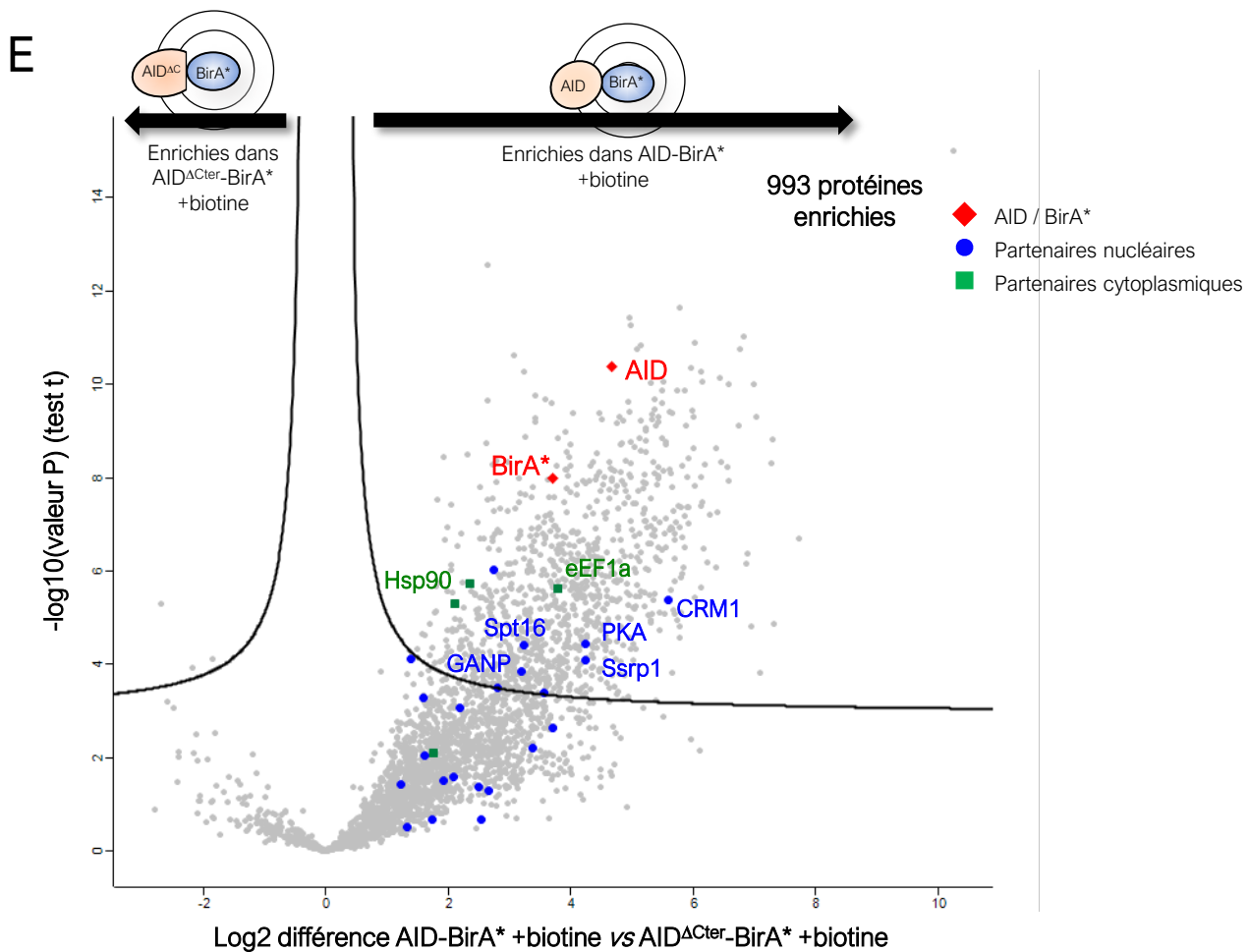
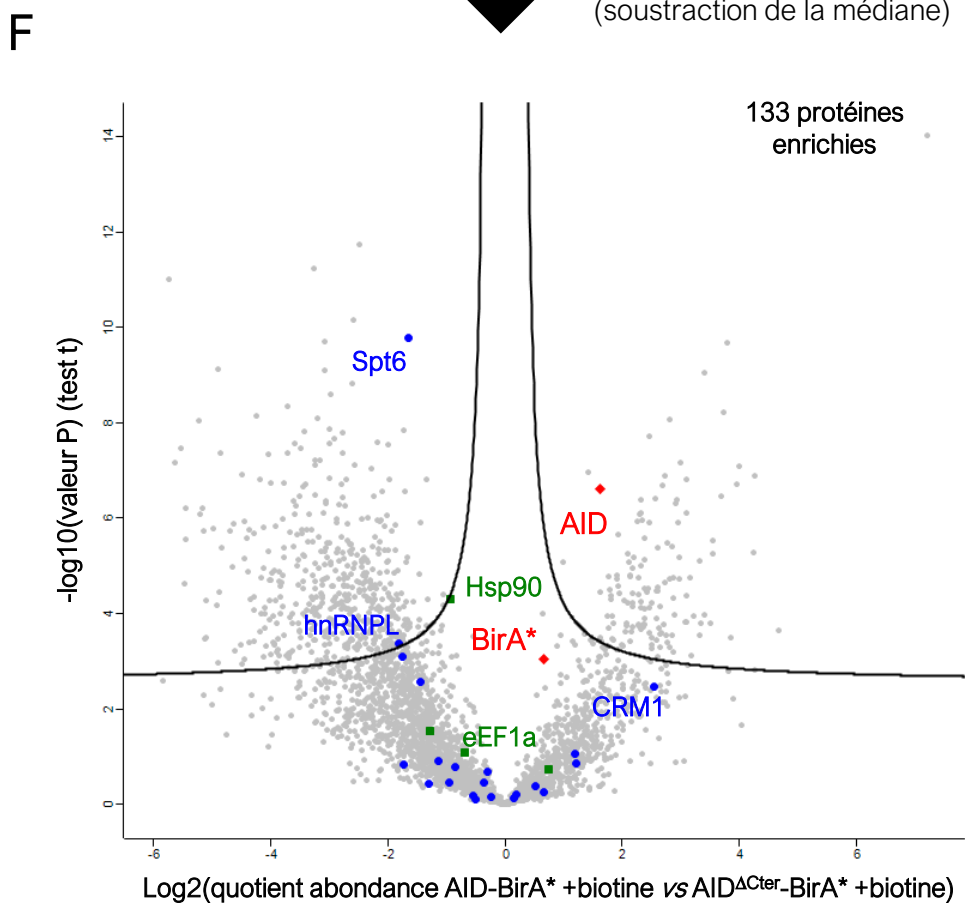


Figure 24 : Analyse statistique de l'enrichissement de protéines identifiées dans les conditions AID-BirA\* +biotine, AID-P2A-BirA\*, BirA\*-AID et AID<sup>ΔCter</sup>-BirA\*.



Normalisation des données  
(soustraction de la médiane)



**Figure 24 : Analyse statistique de l'enrichissement de protéines identifiées dans les conditions AID-BirA\* +biotine, AID-P2A-BirA\*, BirA\*-AID et AID<sup>ΔCter</sup>-BirA\*.** (A) La comparaison statistique de l'abondance des protéines identifiées dans les conditions AID-BirA\* +biotine et AID-P2A-BirA\* +biotine est réalisée à l'aide d'un volcano plot. Les valeurs d'abondance de chaque condition pour chaque protéine font l'objet d'un test t dont la valeur P est corrigée par la méthode de Benjamini-Hochberg. L'inverse du log en base 10 de cette valeur est représentée en ordonnée, alors que le log en base 2 du quotient des valeurs d'abondance d'AID-BirA\* +biotine sur celles d'AID-BirA\* -biotine est représenté sur l'axe des abscisses. Chaque point représente une protéine différente. Le seuil de significativité correspond à une FDR (False Discovery Rate) de 5%. (B) Volcano plot issu de la comparaison de l'abondance des protéines normalisée par soustraction de la médiane de chaque réplicat technique dans les conditions AID-BirA\* +biotine et AID-P2A-BirA\* +biotine. (C,D) Volcanos plot comparant l'abondance normalisée ou non des protéines identifiées dans AID-BirA\* +biotine et BirA\*-AID +biotine. (E,F) Volcanos plot comparant l'abondance normalisée ou non des protéines identifiées dans AID-BirA\* +biotine et AID<sup>ΔCter</sup>-BirA\* +biotine.

limites (Figure 24D). Enfin, l'analyse de l'enrichissement des protéines entre les conditions AID-BirA\* et AID<sup>ΔCter</sup>-BirA\* souffre des mêmes travers que ceux mentionnés auparavant : du fait de la différence de leur niveau d'expression, on retrouve sans normalisation 993 protéines enrichies dans la condition AID-BirA\* (Figure 24E). Il est improbable que le domaine C-ter d'AID permette de recruter autant de protéines, et il est ainsi à nouveau raisonnable de penser qu'au moins une grande partie de ces protéines constitue une partie du bruit de fond aspécifique généré par BirA\*. Ce nombre chute à 133 protéines enrichies après normalisation, mais le fait qu'AID soit toujours enrichie permet de douter encore une fois de la pertinence de ces résultats (Figure 24F). En recoupant les listes de protéines enrichies issues des différents volcano plots, on peut observer sans surprise que la majorité des protéines enrichies par la comparaison des conditions avec biotine le sont également dans la comparaison avec AID-BirA\* -biotine.

Grâce aux différentes listes issues des volcano plots générés avec les données normalisées, j'ai également pu réaliser une analyse de l'enrichissement des termes d'ontologie génique (GO ; de l'anglais gene ontology) présentant un rapport avec l'activité d'AID. La liste de protéines enrichies versus AID-BirA\* -biotine montre la présence importante de protéines associés à l'épissage d'ARNm, à l'ADN simple brin ou à la réparation de l'ADN et localisés au nucléole ou à la chromatine (Figure 25). En revanche, un très grand nombre de ces protéines est perdu dans les autres comparaisons avec les conditions AID-P2A-BirA\* et BirA\*-AID. Enfin, l'analyse des protéines enrichies dans AID-BirA\* +biotine versus AID<sup>ΔCter</sup>-BirA\* montre une perte totale des protéines incluses dans ces termes GO, hormis pour 3 protéines nucléolaires (résultats non montrés). Ceci confirme donc que la comparaison entre AID-BirA\* +biotine et -biotine paraît prometteuse, mais que les autres comparaisons n'ont malheureusement pas l'air d'enrichir le type de protéines auquel nous nous attendions.

A la suite de l'analyse de ces résultats, nous avons donc préféré nous focaliser sur la liste obtenue par la comparaison d'AID-BirA\* +biotine et -biotine, étant donné qu'elle contient beaucoup de partenaires déjà connus d'AID, ainsi qu'un nombre de protéines ayant une fonction associée aux mécanismes de ciblage déjà connus d'AID. Les 50 protéines les plus enrichies de cette comparaison sont présentés dans la table 1, et le classement des partenaires déjà connus d'AID est montré dans la table 2.

#### IV- Choix de gènes candidats à étudier

Un crible CRISPR/Cas9 visant à identifier les facteurs requis pour une CI efficace a été réalisé par un autre membre du laboratoire. Ce crible est mentionné plus en détail dans l'ébauche du manuscrit du prochain chapitre. Celui-ci a permis d'établir une liste de plus de 3000 gènes

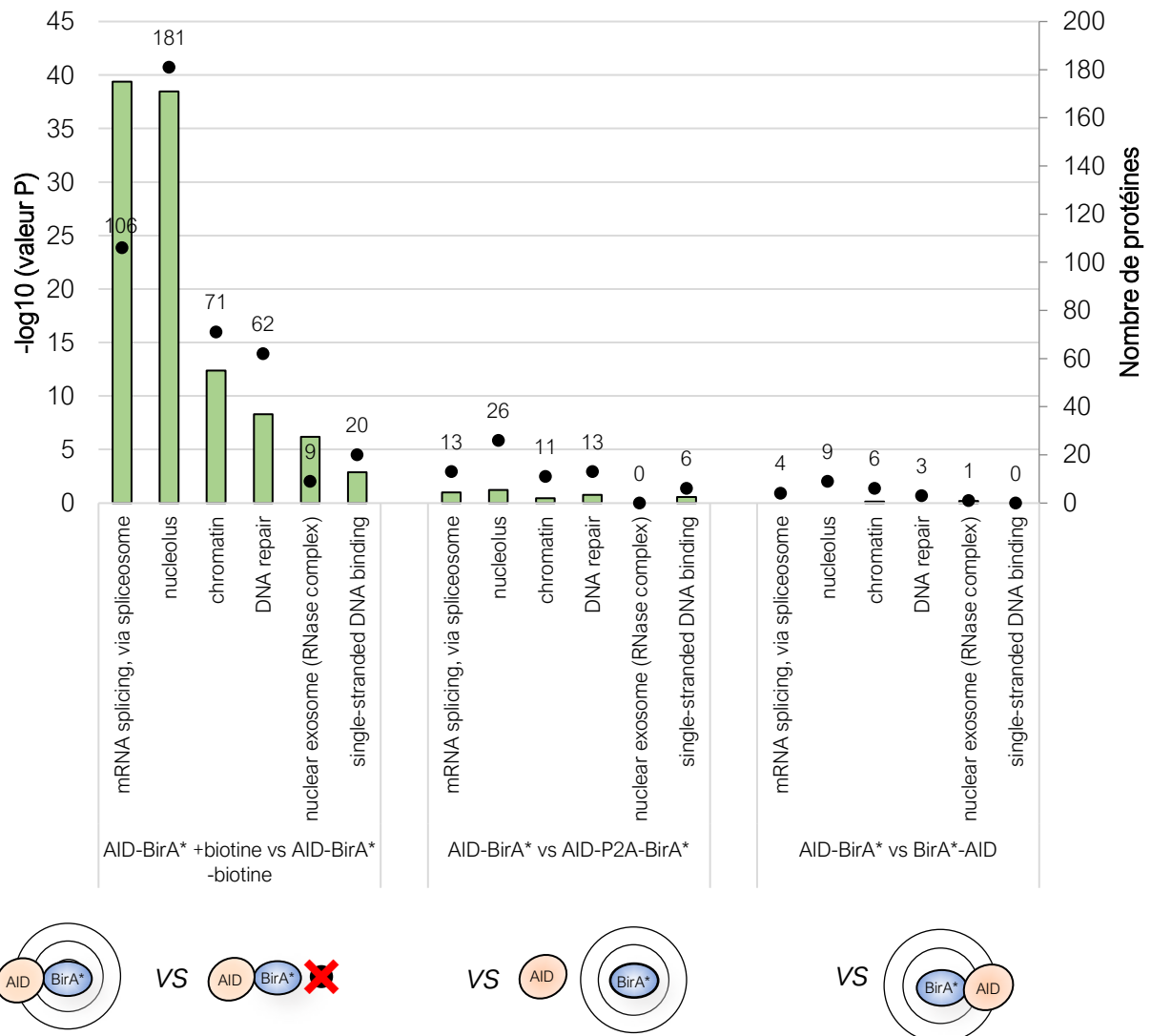


Figure 25 : Enrichissement statistique de processus biologiques liés à AID des protéines enrichies dans AID-BirA\* par rapport aux autres conditions. Analyse de termes d'ontologie génique (GO) issus des listes de protéines enrichies par comparaison de la condition AID-BirA\* +biotine avec AID-BirA\* -biotine (gauche), avec AID-P2A-BirA\* (milieu) ou avec BirA\*-AID (droite). Le nombre de protéines incluses dans chaque terme GO est représenté par un point noir alors que son enrichissement statistique est représenté sous forme de diagramme en bâtons.



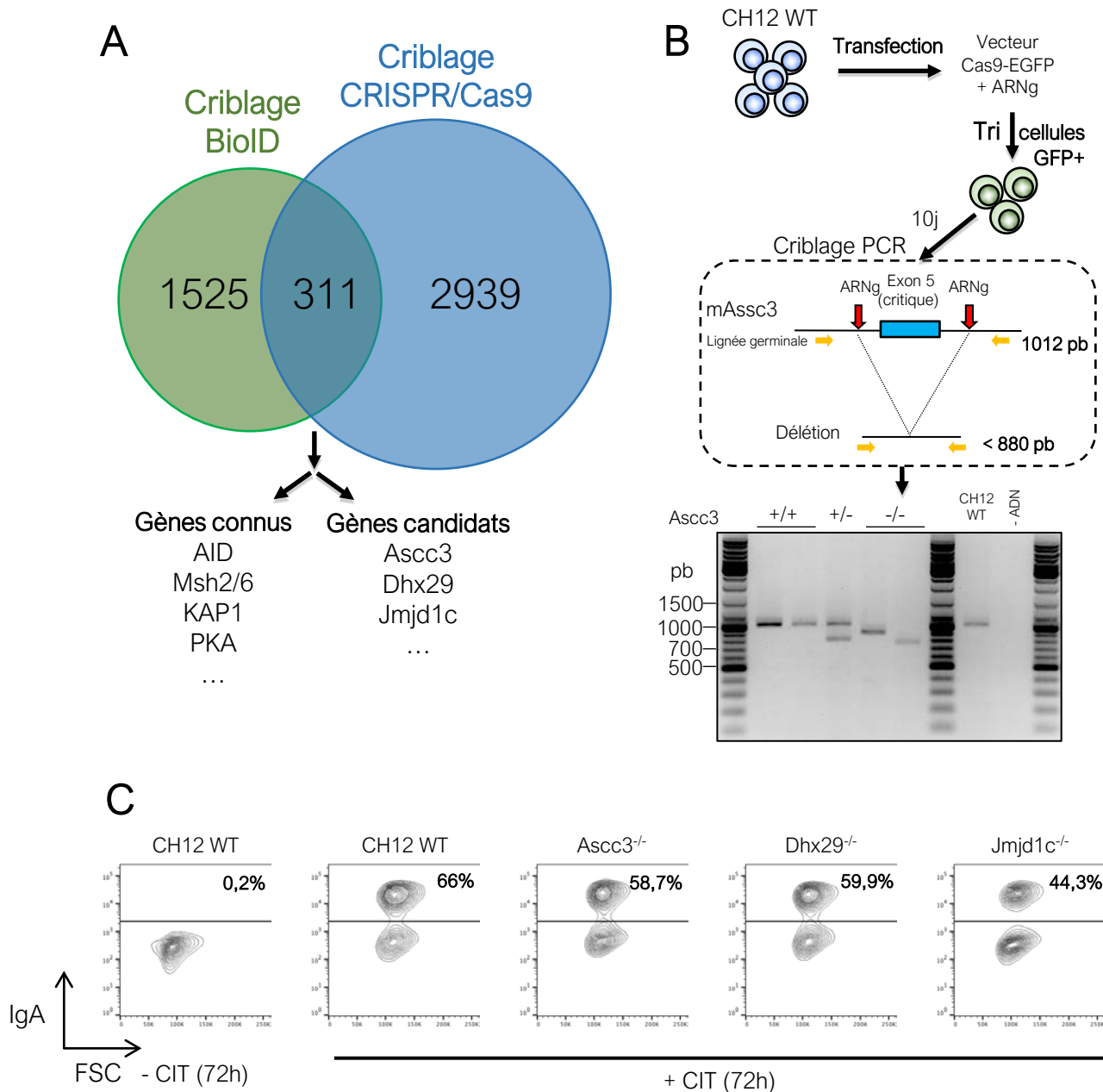
potentiellement impliqués dans le mécanisme global de la CI. Une partie de ces gènes est notamment impliquée dans la stimulation des cellules B, dans la transcription du locus ou la réparation de l'ADN, mais également dans la régulation d'AID. Ainsi, on retrouve également dans cette liste une partie des partenaires d'AID déjà connus tels que KAP1, PKA, le complexe FACT ou des sous-unités du complexe de l'exosome par exemple.

J'ai ainsi recoupé cette liste avec les données issues de la comparaison entre AID-BirA\* +biotine et AID-BirA\* -biotine afin de réduire la grande quantité de potentiels partenaires d'AID ([Figure 26A](#)). On obtient alors une nouvelle liste de 311 protéines qui interagissent potentiellement avec AID et qui ont un impact sur la CI. Parmi ces 311 candidats, j'ai notamment choisi 3 gènes candidats afin d'étudier leur rôle dans la régulation d'AID lors de la CI : Activating signal cointegrator (ASC)1 complex subunit 3 (Ascc3), DExH-box hélicase 29 (Dhx29) et Jumonji Domain Containing 1C (Jmjd1c). J'ai choisi d'étudier ces protéines car leur fonction connue a un rapport avec les mécanismes en jeu lors de la CI.

Ascc3 fait partie du complexe ASC1 et agit en tant qu'hélicase à ADN de 3' vers 5', et est importante pour réparer les dommages d'alkylation dans l'ADN. Il expose ainsi l'ADN sous forme simple brin pour qu'ils puissent être réparés par Alkbh3 ([Dango et al., 2011](#)). Il est ainsi possible que ce facteur soit impliqué dans le ciblage d'AID et permette de lui fournir de l'ADN simple brin comme substrat.

L'hélicase à ARN ATP-dépendante Dhx29 est impliquée dans l'initiation de la traduction afin de rendre plus accessible l'extrémité 5' structurée de certains ARNm ([Dhote et al., 2012](#); [Parsyan et al., 2009](#)). De par son interaction avec Melanomadifferentiation-associated gene-5 (MDA5) et retinoic acid-inducible gene I (RIG-I), elle sert également de co-senseur à l'ARN viral, et est de fait requise pour la production d'interféron de type I dans les cellules épithéliales respiratoires humaines ([Oshiumi et al., 2016](#); [Sugimoto et al., 2014](#); [Zhu et al., 2018](#)). Etant une hélicase à ARN, une hypothèse est qu'elle est peut-être impliquée dans la formation de R-loops ou dans le cadre du passage des structures de G4 à ARN à des R-loops, de façon similaire à Ddx1 ([Ribeiro de Almeida et al., 2018b](#)).

Le protéine Jmjd1c déméthyle la lysine 9 de l'histone H3 donc régule ainsi l'expression de plusieurs gènes dont certains sont impliqués dans des leucémies myéloïdes aiguës ([Chen et al., 2015](#); [Xu et al., 2020](#)). De manière intéressante, elle déméthyle également MDC1 ce qui permet d'améliorer son interaction avec RNF8 et son recrutement aux DSBs ([Watanabe et al., 2013](#)). RNF8 permet à son tour le recrutement et l'accumulation de 53BP1 aux DSBs ([Li et al., 2010b](#); [Ramachandran et al., 2010](#); [Santos et al., 2010](#)) pour favoriser leur réparation et la recombinaison



**Figure 26 : Délétion et caractérisation de gènes candidats issus des cribles BiLD et CRISPR/Cas9. (A)** Diagramme de Venn montrant le recoupement entre les résultats issus de l'utilisation de BiLD et ceux issus d'un crible CRISPR/Cas9 visant à identifier les facteurs requis pour la CI. Des 311 facteurs communs aux 2 listes, certains sont déjà connus pour être requis ou importants pour la CI, mais de nombreux autres n'ont pas de rôle connu dans la CI, tels qu'Ascc3, Dhx29 ou Jmjd1c par exemple. **(B)** Schéma illustrant la stratégie utilisée pour générer et cribler des cellules CH12 knock-out pour un gène donné. Des CH12 WT sont transfectées avec un vecteur permettant l'expression de Cas9-EGFP et de 2 ARNg s'hybridant de part et d'autre d'un exon critique du gène ciblé. La stratégie utilisée pour générer le knock-out d'Ascc3 a été montrée en exemple. La bonne délétion de l'exon critique est confirmée par PCR qui génère un produit de 1012 pb lors de l'amplification d'un allèle WT et un produit de 880 pb ou moins lorsque l'allèle est délété. La taille des amplicons est ensuite vérifiée en les déposant sur un gel d'agarose. **(C)** Analyse par cytométrie en flux de l'effet de la délétion de gènes candidats sur l'effet de la CI. Des cellules CH12 WT ainsi que des lignées clonales dans lesquelles les gènes candidats Ascc3, Dhx29 ou Jmjd1c sont délétés ont été stimulées ou non (-CIT) pendant 72h à l'aide de CD40L, d'IL-4 et de TGF- $\beta$  (+CIT). L'efficacité de CI résultant de cette stimulation a par la suite été déterminée par l'expression d'IgA de surface par cytométrie en flux.

entre 2 DSBs distales lors de la CI ([Dong et al., 2015](#); [Manis et al., 2004a](#); [Reina-San-Martin et al., 2007](#)).

Pour caractériser le rôle de ces 3 candidats, j'ai réalisé un KO des gènes qui les codent par CRISPR/Cas9 dans la lignée CH12 (**Figure 26B**). Pour cela, j'ai généré un vecteur par gène codant pour la Cas9 et pour 2 ARNg ciblant les régions introniques juste en amont et en aval d'un exon critique, afin d'aboutir à sa délétion complète et au décalage du cadre de lecture permettant le KO. De façon similaire à la génération de KI, des cellules CH12 WT sont alors transfectées avec ce vecteur, puis les cellules GFP positives exprimant la Cas9 sont triées pour générer des clones uniques. Ces lignées clonales sont génotypes 10 jours plus tard par un criblage PCR permettant de confirmer la délétion de l'exon grâce à 2 oligonucléotides s'hybridant en amont et en aval des sites de coupure 5' et 3' respectivement (**Figure 26B**). Les clones sont également séquencés afin de valider le KO. J'ai ainsi obtenu plusieurs clones KO pour *Ascc3*, *Dhx29* ou *Jmjd1c*. Après stimulation de ces clones, il est possible d'observer que la délétion de *Jmjd1c* semble entraîner une légère baisse de l'efficacité de la CI par rapport aux CH12 WT, alors que la délétion d'*Ascc3* ou de *Dhx29* n'en entraîne pas (**Figure 26C**).

## V- Discussion

Au cours de ce projet, j'ai réalisé la mise au point de la technologie BioID afin d'identifier les partenaires protéiques d'AID. J'ai tout d'abord montré qu'aucun des 2 KI qui permet la fusion de BirA\* en N-ter ou en C-ter d'AID n'est fonctionnel dans les CH12. J'ai alors mis au point les conditions pour utiliser BioID par infection rétrovirale de cellules B primaires murines, en utilisant plusieurs conditions pour faciliter l'identification de protéines candidates. Malheureusement, ces conditions supplémentaires n'ont pas donné de résultats probants, mais la comparaison des conditions basales dans lesquelles BirA\* fusionnée à AID est activée ou non nous a tout de même permis d'obtenir une liste de protéines intéressantes et interagissant potentiellement avec AID. En parallèle, un crible CRISPR/Cas9 visant à identifier les gènes requis pour la CI a été réalisé dans le laboratoire. Le recoupement des données de ces 2 cribles a mené à une réduction de la liste de candidats, me permettant de choisir 3 gènes dont j'ai réalisé le KO dans les CH12 afin, dans un premier temps, d'évaluer leur impact sur l'efficacité de CI. L'absence de Jmjd1c semble ainsi entraîner une légère mais constante baisse de CI dans plusieurs clones.

### **BioID, une technique relativement nouvelle avec ses avantages et ses limites**

Afin d'identifier les partenaires d'une protéine cible, une approche couramment utilisée consiste à réaliser une précipitation de cette protéine cible interagissant avec ses partenaires dans des conditions dites natives pour tenter de préserver les interactions protéiques ([LaCava et al., 2016](#); [Mehta and Trinkle-Mulcahy, 2016](#)). Cette technique de co-immunoprécipitation (co-IP) a été réalisée dans notre laboratoire avant mon arrivée, et a notamment permis l'identification de KAP1 ([Jeevan-Raj et al., 2011](#)), des cohésines ([Thomas-Claudepierre et al., 2013](#)) et des sous-unités du complexe médiateur ([Thomas-Claudepierre et al., 2016](#)) comme partenaires d'AID. Cette expérience, malgré son succès, présentait quelques défauts : d'une part, AID taggée en N-ter avait dû être surexprimée dans un grand nombre de cellules CH12 et d'autre part, ces cellules n'étaient pas stimulées pour réaliser la CI. Par conséquent, c'est avec l'idée de palier à ces défauts ainsi qu'aux limitations techniques de la co-IP (qui ne permet de détecter que des partenaires courants ou sous forme d'un complexe), que nous avons entrepris d'utiliser la technologie BioID.

Au regard des points évoqués auparavant, l'avantage principal de BioID réside dans le fait que le marquage des protéines partenaires de la cible se fait dans des cellules vivantes, ce qui permet de ne pas avoir à conserver ces interactions après la lyse cellulaire. En outre, cette technologie permet aussi, en théorie, d'identifier des interactions plus transitoires qu'avec une co-IP classique. Pour cela, BirA\* catalyse la biotinylation des protéines l'entourant par une réaction en 2 étapes : elle transforme tout d'abord la biotine en biotinyl-5'-AMP, produit pour lequel elle a une affinité

réduite due à sa mutation (R118G) ([Roux et al., 2012](#)). Elle va ainsi relâcher le biotiny-5'-AMP de son site catalytique et celui-ci va s'accumuler sous forme d'un « nuage » autour de BirA\*. Il est capable de réagir avec les lysines exposées des protéines environnantes, et va ainsi permettre leur biotinylation covalente ([Roux et al., 2012](#)). Ceci implique donc que BirA\* biotinye également des protéines qui n'interagissent ni directement avec elle, ni directement avec la protéine à laquelle elle est fusionnée, ce qui entraîne l'apparition d'un bruit de fond de faux-positifs.

Le fait que de la biotine soit liée aux protéines proches constitue un avantage : en effet, la streptavidine, utilisée pour purifier les protéines biotinylées, possède une affinité extrêmement forte pour la biotine ( $K_d = 10^{-14}$  M). Il est ainsi possible de procéder à des lavages stricts capables d'éliminer la plupart des contaminants et des protéines non-spécifiques sans risquer de décrocher les protéines biotinylées. Cependant, cela implique qu'il faut alors recourir à l'utilisation de détergents ou d'urée pour éluer les protéines biotinylées, mais celle-ci décroche également les molécules de streptavidine couplées aux billes magnétiques. Cette streptavidine monomérique de petite taille (14 kDa) peut être considérée comme un contaminant présent en très grande abondance pouvant masquer alors la présence d'autres petites protéines faiblement abondantes.

Enfin, l'inconvénient majeur pour l'utilisation de BioID est la nécessité de générer une protéine de fusion entre BirA\* et la protéine cible. Si, tout comme dans notre cas, un KI n'est pas envisageable, il faut alors que celle-ci soit exprimée de manière ectopique, à un niveau d'expression qui peut être différent de la protéine endogène. De plus, cette fusion peut entraîner un changement de la localisation ou encore de l'activité catalytique de la protéine cible.

### **Fonctionnalité des protéines de fusion**

Il aurait été bon de vérifier si la répartition intra-cellulaire de chaque protéine de fusion est conforme à ce que l'on attend. Cela aurait pu être réalisé à l'aide d'expériences d'immunofluorescence (IF), cependant nous n'avons pas de bon anticorps anti-AID ou anti-Flag pour ce type d'expérience. Une alternative aurait été de réaliser un fractionnement des compartiments cellulaires et d'observer par WB si les protéines de fusion sont présentes ou non dans le noyau, et en quelle quantité. Quoiqu'il en soit, la fusion en C-ter ne semble pas ou peu changer la localisation d'AID, puisqu'AID-BirA\* est capable de sauver la déficience en CI des cellules AID<sup>-/-</sup>. Une étude utilisant des mutants d'AID ne possédant plus leur NES et fusionnés à la GFP en C-ter sont bien quasi-exclusivement nucléaires ([Patenaude et al., 2009](#)), ce qui montre que ce type de fusion n'altère pas l'import d'AID. Enfin, il aurait pu être envisagé de comparer le ciblage d'AID et d'AID-BirA\* aux régions S par CHIP. Un défaut partiel de recrutement d'AID-BirA\* pourrait alors expliquer que cette fusion semble moins capable qu'AID seule d'initier la CI. De

même, l'activité catalytique de chaque fusion d'AID aurait pu être testée à l'aide d'un test de réversion pour la résistance à la rifampicine dans des bactéries ([Methot et al., 2018](#)). Dans ce test, une cassette de résistance à la rifampicine peut devenir à nouveau fonctionnelle à l'aide des mutations générées par AID, si elle est active. La proportion de bactéries ayant réalisé une réversion détermine alors l'efficacité de l'activité catalytique des différentes protéines de fusion. Cela permettrait notamment de déterminer si AID-BirA\* est aussi active qu'AID seule.

### **Analyse des différentes listes de protéines enrichies**

Afin de faciliter l'analyse des protéines identifiées à l'aide de la condition d'intérêt AID-BirA\*, nous avons souhaité utiliser d'autres conditions comme BirA\*-AID, AID-P2A-BirA\* et AID<sup>ACTer</sup>-BirA\* et AID-BirA\* en absence de biotine. Je vais ainsi discuter des résultats obtenus à l'aide de chacune de ces conditions.

#### **AID-BirA\* vs. BirA\*-AID :**

La condition BirA\*-AID devait permettre d'identifier principalement les partenaires cytoplasmiques d'AID. Même si nous n'avons pas validé la localisation de cette protéine, son incapacité presque totale à initier la CI a conforté l'hypothèse selon laquelle elle est exclusivement cytoplasmique ([Figure 20](#)). Cependant, au vu des résultats obtenus avec les données de MS, on peut se demander si c'est réellement le cas. En effet, on retrouve une bonne partie des partenaires nucléaires d'AID identifiés dans cette condition, ce qui fait qu'ils ne sont pas considérés comme statistiquement enrichis dans la comparaison AID-BirA\* contre BirA\*-AID. Une expérience possible pour savoir si BirA\*-AID rentre ou non dans le noyau serait de traiter les cellules exprimant BirA\*-AID avec de la leptomycine B, un inhibiteur de l'exportine-1 qui exporte AID en dehors du noyau ([Kudo et al., 1999](#); [McBride et al., 2004](#); [Patenaude et al., 2009](#)). Ainsi, par immunofluorescence ou en réalisant un fractionnement pour récupérer des extraits nucléaires, il serait possible d'observer si BirA\*-AID s'accumule dans le noyau en présence de leptomycine B. Si c'est le cas, cela veut dire qu'elle réalise toujours des aller-retours entre le noyau et le cytoplasme, et ceci pourrait être la raison pour laquelle une bonne partie des partenaires nucléaires d'AID sont biotinylés dans cette condition. Il serait alors également intéressant de comprendre pourquoi cette protéine de fusion n'est pas du tout capable d'initier la CI, à des niveaux observés encore plus faibles que ceux d'AID<sup>ACTer</sup>-BirA\* ([Figure 20](#)). De plus, il serait intéressant de voir si cette condition pourrait être remplacée par une autre où le NLS d'AID serait supprimé, l'empêchant ainsi d'entrer dans le noyau.

### AID-BirA\* vs. AID-P2A-BirA\* :

L'utilisation de la condition AID-P2A-BirA\* devait permettre d'identifier les protéines interagissant préférentiellement avec BirA\* pour pouvoir les éliminer de l'analyse. De la même façon que pour BirA\*-AID, cette condition n'a pas apporté les résultats attendus. En effet, on peut voir que c'est la condition dans laquelle le plus de protéines ont été identifiées (presque 3000) ([Figure 22B](#)). Etant donné le fait que la biotinylation des protéines ne représente pas forcément une interaction directe, ce nombre important de protéines identifiées reflète probablement le fait que BirA\* dissociée d'AID est alors libre de se déplacer dans toute la cellule, et génère ainsi beaucoup de bruit de fond. De plus, elle a l'air bien mieux exprimée qu'AID-BirA\* ([Figure 21B](#)), expliquant ainsi qu'elle génère plus de bruit de fond, et biotinylyle abondamment des protéines qui n'interagissent pas forcément avec elle, dont certains partenaires nucléaires d'AID. Ainsi, ce pourrait être la raison pour laquelle ces derniers sont « perdus » de la liste de protéines enrichies lorsqu'AID-BirA\* est comparée à AID-P2A-BirA\*. Par conséquent, à la place de ce contrôle, il aurait été préférable d'utiliser une protéine semblable à AID mais pas impliquée dans les mêmes processus cellulaires, limitant spatialement l'activité de BirA\*. Il aurait ainsi été plus avantageux d'utiliser APOBEC2, comme cela a été fait pour l'étude publiée qui utilise également BioID pour identifier des partenaires d'AID ([Methot et al., 2018](#)).

### AID-BirA\* vs. AID<sup>ACTer</sup>-BirA\* :

La protéine AID<sup>ACTer</sup>-BirA\* devait permettre, par soustraction de liste, aboutir à l'identification de facteurs potentiellement requis pour la CI, mais pas pour l'HMS, en accord avec le phénotype observé d'AID<sup>ACTer</sup> ([Barreto et al., 2003b](#)). La comparaison sans normalisation d'AID-BirA\* et d'AID<sup>ACTer</sup>-BirA\* montre une grosse différence d'abondance moyenne entre ces 2 conditions, ce qui fait que plus de 900 protéines sont considérées comme étant enrichies dans AID-BirA\* ([Figure 24E](#)), alors que l'on s'attendait plutôt à l'enrichissement d'un nombre limité de protéines. Nous pensons que cela est dû au fait qu'AID<sup>ACTer</sup> est majoritairement nucléaire car elle ne possède pas de NES. Or, il a été montré que la demi-vie d'AID est grandement réduite dans le noyau ([Aoufouchi et al., 2008](#)), ce qui pourrait signifier que ces protéines de fusion sont trop rapidement dégradées pour pouvoir générer un signal convenable et abondant. Cette demi-vie raccourcie est aussi probablement la raison pour laquelle AID<sup>ACTer</sup>-BirA\* semble moins bien exprimée que les autres constructions ([Figure 21A](#)). Quoiqu'il en soit, la liste de 133 protéines enrichies après normalisation des données a également été recoupée avec la liste issue du crible CRISPR/Cas9, pouvant donner lieu à l'étude de potentiels partenaires interagissant avec le domaine-Cter d'AID et requis pour la CI (résultats non montrés).

## AID-BirA\* +biotine vs. AID-BirA\* -biotine :

La fusion de BirA\* en C-ter d'AID s'est avérée être la condition qui pourrait nous permettre d'identifier de nouveaux partenaires d'AID impliqués dans sa régulation lors de la CI. En effet, c'est la seule fusion capable de restaurer (au moins partiellement) la déficience de CI dans des cellules B primaires AID<sup>-/-</sup>. De façon intéressante, une étude publiée après l'obtention de mes résultats de MS a également utilisé cette protéine de fusion ([Methot et al., 2018](#)). Toutes les protéines qu'ils identifient comme enrichies dans leur expérience sont également retrouvées dans ma liste de protéines enrichies, révélant que mes résultats sont reproductibles et robustes. Cette liste de 1836 protéines contient sûrement un certain nombre de faux-positifs, mais contient également des partenaires moins fréquents et plus transitoires d'AID, identifiés moins abondamment. Du fait de la localisation majoritairement cytoplasmique d'AID, nous nous attendions à ce que les partenaires cytoplasmiques soient les plus enrichis ([Patenaude and Di Noia, 2010](#); [Patenaude et al., 2009](#); [Rada et al., 2002a](#)). Or, ça n'a pas l'air d'être le cas, puisqu'on trouve qu'une partie des partenaires nucléaires sont plus enrichis, comme pontin (Ruvb1) et reptin (Ruvb2) du complexe INO80 ([Kracker et al., 2015](#)) ou encore des sous-unités de l'ARN pol II ([Nambu et al., 2003](#)) par exemple. Toutefois, d'autres facteurs requis pour le ciblage d'AID tels que Spt5 ne sont pas enrichis. Cela pourrait être dû au fait que Spt5 (et d'autres facteurs) interagissent avec le domaine N-terminal d'AID, et que BirA\* fusionnée en C-ter d'AID est alors placée trop loin de ces partenaires pour les biotinyler efficacement.

## Conclusions sur le système BioID

Historiquement, l'utilisation de BioID s'est faite pour identifier des protéines contenues au sein de complexes stables comme celui des pores nucléaires ([Kim et al., 2014](#)) ou des éléments structurels comme la lamina nucléaire ([Chojnowski et al., 2015](#); [Roux et al., 2012](#)). Dans ce contexte, la technologie BioID est efficace et a permis l'identification de nombreux partenaires protéiques ([Varnaite and MacNeill, 2016](#)). Dans notre cas, la faible proportion d'AID nucléaire, la nature transitoire de ses interactions ainsi que sa localisation dynamique faites pour limiter son action mutagène sont autant de difficultés rencontrées pour l'identification de partenaires impliqués dans la régulation d'AID. De plus, l'accumulation relativement lente du signal de biotinylation par BirA\*, qui se compte en heures, peut faire que certaines protéines dont la cinétique d'interaction est rapide ne seront pas ou très peu biotinylées. Parmi celles-ci, on peut notamment penser aux protéines impliquées dans la réparation de l'ADN, dont la cinétique d'action se compte plutôt en secondes ou en minutes ([Haince et al., 2008](#); [Kochan et al., 2017](#)).



De plus, le manque de contrôle de l'activité de BirA\* fait qu'on observe en réalité l'ensemble des interactions ayant lieu au cours de la « vie » d'une protéine, qu'elles soient reliées au processus biologique d'intérêt ou non. Les résultats que nous avons obtenus montrent néanmoins que BioID reste une bonne approche pour étudier les mécanismes de régulation d'AID dans leur ensemble et identifier de manière globale les protéines interagissant avec AID, même si le bruit de fond généré est important, et qu'il est alors difficile de distinguer les vrais co-facteurs peu enrichis des faux-positifs se trouvant à proximité de BirA\*. Cela rend également le choix de candidats à étudier difficile à faire. Celui-ci aurait dû être facilité par la comparaison d'AID-BirA\* avec les autres protéines de fusion, mais cela n'a pas donné les résultats escomptés.

Afin de tenter d'améliorer et d'affiner les résultats que nous avons obtenus grâce à l'utilisation de BioID dans des cellules B primaires, nous envisageons également l'utilisation de BioID dans la lignée CH12. Cela présenterait plusieurs avantages : il deviendrait ainsi possible de réaliser une sélection avec un antibiotique comme la puromycine pour enrichir la proportion de cellules transduites. Il n'était pas possible de le faire avec des cellules B primaires car, une fois mises en culture *ex vivo*, celles-ci meurent au bout de quelques jours. Ainsi, l'efficacité d'infection ne serait plus un facteur limitant et le signal de biotinylation total serait plus élevé pour le même nombre de cellules. L'infection rétrovirale et la stimulation des cellules pouvant alors être découplée, il est alors également envisageable de rajouter la biotine dans le milieu aux alentours de 24h post-stimulation, là où, dans un contexte normal d'induction de l'expression d'AID, sa quantité est maximale ([Cortizas et al., 2016a](#)). Il est ainsi possible d'imaginer que certains facteurs régulant l'action d'AID puissent être également induits rapidement après la stimulation des cellules, et que ceux-ci n'ont pas pu être identifiés au vu de la configuration expérimentale utilisée dans les cellules B primaires. Enfin, l'utilisation de la lignée CH12 facilite la culture d'un grand nombre de cellules, et nous envisageons ainsi également de réaliser du fractionnement d'extraits cellulaires pour enrichir les partenaires nucléaires d'AID.

### **Gènes candidats et perspectives**

La réalisation d'un crible CRISPR/Cas9 pour identifier les facteurs requis pour la CI nous a permis de faire baisser drastiquement le nombre de candidats. Ce crible, basé sur l'importance de la fonction des protéines lors la CI offre une approche non biaisée complémentaire au crible BioID, et permet de choisir facilement les candidats à étudier, qui jouent donc potentiellement un rôle sur la régulation d'AID lors de la CI. Cependant, la nature du crible CRISPR/Cas9 à générer en théorie un seul KO par cellule exclut donc les facteurs avec des redondances fonctionnelles de ce recoupement. De plus, ce crible visant à identifier les facteurs entraînant une perte de

l'efficacité de CI, les facteurs qui régulent négativement l'action d'AID et dont l'absence entraîne potentiellement une hausse de l'efficacité de CI sont également exclus du recoupement de listes.

J'ai ainsi choisi 3 gènes candidats issus de ce recoupement et des résultats préliminaires issus de la stimulation de clones KO montrent que parmi eux, Jmjd1c pourrait effectivement jouer un rôle dans la régulation d'AID lors de la CI. Par conséquent, du fait que Jmjd1c déméthyle MDC1 et permette le recrutement de RNF8 et 53BP1 aux DSBs ([Li et al., 2010b](#); [Ramachandran et al., 2010](#); [Santos et al., 2010](#); [Watanabe et al., 2013](#)), il se pourrait que son interaction avec AID régule et promeuve la bonne réparation des DSBs lors de la CI. Ces résultats prometteurs, offrent ainsi la possibilité qu'AID participe elle-même à la réparation des DSBs dont elle est la cause, et ce mécanisme permettrait d'ajouter un niveau de régulation et de protection supplémentaire vis-à-vis des dommages induits à AID dans l'ADN. Ainsi, ce facteur mérite de plus amples investigations.

Pour conclure, l'utilisation de la technologie BioID dans des lymphocytes B primaires m'a permis d'identifier plus de 1800 protéines comme interagissant potentiellement avec AID. Parmi celles-ci, on retrouve la majorité des partenaires déjà connus d'AID, mais également beaucoup de protéines non spécifiques probablement enrichies par le bruit de fond de biotinylation important généré par BirA\*. Ainsi, pour pouvoir tirer avantage de cette liste, il y a nécessité de la croiser avec une autre liste basée sur un crible fonctionnel, ce qui permet de réduire nettement le nombre de candidats et de faire ressortir ceux qui ont un impact sur l'efficacité de CI. La caractérisation fonctionnelle de ces facteurs devrait nous permettre de mieux comprendre comment l'action d'AID est régulée lors de la CI.

**Table 1. Top 50 des protéines les plus enrichies entre la condition AID-BirA\* +biotine et AID-BirA\* -biotine**

Nom du gène	Rang +Biotine vs -biotine	Log2 abondance +biotine / -biotine	Description
Hcls1	1	12,1877	hematopoietic cell specific Lyn substrate 1
Supt6h	2	11,9959	Transcription elongation factor SPT6
Swap70	3	11,6016	SWA-70 protein
Coro1b	4	11,5278	coronin, actin binding protein 1B
Ppp6r1	5	11,4583	protein phosphatase 6, regulatory subunit 1
Atg2b	6	11,2241	autophagy related 2B
Eif4g1	7	11,2212	eukaryotic translation initiation factor 4, gamma 1
Prcc2c	8	11,0511	proline-rich coiled-coil 2C
Pacsin2	9	10,8832	protein kinase C and casein kinase substrate in neurons 2
Ankrd44	10	10,8795	ankyrin repeat domain 44
Larp1	11	10,8614	La ribonucleoprotein domain family, member 1
Vcpip1	12	10,6788	valosin containing protein (p97)/p47 complex interacting protein 1
Eif5	13	10,5523	eukaryotic translation initiation factor 5
Ankrd17	14	10,5514	ankyrin repeat domain 17
Ubp2l	15	10,5436	ubiquitin-associated protein 2-like
Csde1	16	10,5028	cold shock domain containing E1, RNA binding
Nap1l4	17	10,4839	nucleosome assembly protein 1-like 4
Ranbp2	18	10,4778	RAN binding protein 2
Tln1	19	10,3704	talin 1
Ktn1	20	10,3547	kinectin 1
Nudc	21	10,3391	nuclear distribution gene C homolog (Aspergillus)
Hdlbp	22	10,3160	high density lipoprotein (HDL) binding protein
Zbp1	23	10,2690	Z-DNA binding protein 1
Ppp6r3	24	10,2371	protein phosphatase 6, regulatory subunit 3
Samsn1	25	10,1934	SAM domain, SH3 domain and nuclear localization signals, 1
Serbp1	26	10,1776	serpine1 mRNA binding protein 1
Lpxn	27	10,1449	leupaxin
Ppp6c	28	10,0013	protein phosphatase 6, catalytic subunit
Wwox	29	9,9359	WW domain-containing oxidoreductase
Strap	30	9,9061	serine/threonine kinase receptor associated protein
Inpp5d	31	9,9055	inositol polyphosphate-5-phosphatase D
Ankrd28	32	9,8858	ankyrin repeat domain 28
Sec16a	33	9,8224	SEC16 homolog A (S. cerevisiae)
Fkbp8	34	9,8221	FK506 binding protein 8
Mtdh	35	9,7295	metadherin
Mre11	36	9,6313	Double-strand break repair protein MRE11
Snx9	37	9,5783	sorting nexin 9
Nap1l1	38	9,4893	nucleosome assembly protein 1-like 1
Coro1c	39	9,4437	coronin, actin binding protein 1C
G3bp1	40	9,3733	GTPase activating protein (SH3 domain) binding protein 1
Septin7	41	9,3571	Septin 7
Eif4enif1	42	9,3296	eukaryotic translation initiation factor 4E nuclear import factor 1
Tagln2	43	9,3041	transgelin 2
Aicda	44	9,3012	activation-induced cytidine deaminase
Lcp1	45	9,2848	lymphocyte cytosolic protein 1
Xrn1	46	9,2683	5'-3' exoribonuclease 1
Ncf4	47	9,2590	neutrophil cytosolic factor 4
Rangap1	48	9,2032	RAN GTPase activating protein 1
Tgtp1	49	9,1607	T cell specific GTPase 1
Flna	50	9,1461	filamin, alpha

Table montrant les 50 protéines les plus enrichies dans la condition AID-BirA\* +biotine comparée à AID-BirA\* -biotine, leur facteur d'enrichissement (log2 abondance +biotine/-biotine) ainsi que leur classement en fonction de ce facteur d'enrichissement.

**Table 2. Enrichissement des protéines partenaires d'AID ou reliées à la réparation de l'ADN dans la condition AID-BirA\* +biotine vs. -biotine**

Nom du gène	Rang +Biotine vs -biotine	Log2 abondance +biotine / -biotine	Description
Supt6h	2	11,9959	Transcription elongation factor SPT6
Mre11	36	9,6313	Double-strand break repair protein MRE11
Aicda	44	9,3012	activation-induced cytidine deaminase
Psmc3	88	8,2928	proteasome (prosome, macropain) activator subunit 3 (PA28 gamma, Ki)
Ruvbl1	91	8,2289	RuvB-like protein 1
Dnaja1	94	8,1850	DnaJ (Hsp40) homolog, subfamily A, member 1
Jmjd1c	104	8,1031	jumonji domain containing 1C
Ascc3	105	8,1027	activating signal cointegrator 1 complex subunit 3
Ruvbl2	120	7,9785	RuvB-like protein 2
BirA*-Flag	220	6,9821	Biotin ligase BirA*(R118G)
Dhx29	253	6,718074375	DEAH (Asp-Glu-Ala-His) box polypeptide 29
Polr2a	421	5,8384	polymerase (RNA) II (DNA directed) polypeptide A
Syncrip	440	5,7391	synaptotagmin binding, cytoplasmic RNA interacting protein
Exosc9	452	5,6952	exosome component 9
Xpo1	462	5,6428	exportin 1, CRM1 homolog (yeast)
Polr2b	494	5,5231	polymerase (RNA) II (DNA directed) polypeptide B
Msh2	566	5,2160	mutS homolog 2 (E. coli)
Msh6	597	5,0752	mutS homolog 6 (E. coli)
Polh	605	5,0599	polymerase (DNA directed), eta (RAD 30 related)
Rad50	657	4,8885	RAD50 homolog (S. cerevisiae)
Ssrp1	660	4,8838	structure specific recognition protein 1
Exosc7	664	4,8643	exosome component 7
Mcm3ap	685	4,7732	minichromosome maintenance deficient 3 (S. cerevisiae) associated protein
Eef1a1	724	4,6338	eukaryotic translation elongation factor 1 alpha 1
Trim28	735	4,5977	tripartite motif-containing 28
Cbx3	737	4,5969	chromobox 3
Hsp90ab1	759	4,5254	heat shock protein 90 alpha (cytosolic), class B member 1
Exosc8	765	4,5044	exosome component 8
Rad51	811	4,3843	RAD51 homolog
Polr2m	856	4,2871	polymerase (RNA) II (DNA directed) polypeptide M
Supt16h	871	4,2377	FACT complex subunit SPT16
Prkaca	872	4,2339	protein kinase, cAMP dependent, catalytic, alpha
Ddx1	885	4,2072	DEAD (Asp-Glu-Ala-Asp) box polypeptide 1
Rpa2	890	4,1979	replication protein A2
Ywhaq	912	4,1377	14-3-3 protein theta
Ywhah	944	4,0723	14-3-3 protein eta
Exosc2	1057	3,8461	exosome component 2
Atm	1083	3,8003	ataxia telangiectasia mutated homolog (human)
Hsp90aa1	1088	3,7953	heat shock protein 90, alpha (cytosolic), class A member 1
Hnrnpf	1096	3,7688	heterogeneous nuclear ribonucleoprotein F
Polr2e	1148	3,6618	polymerase (RNA) II (DNA directed) polypeptide E
Prkacb	1167	3,6150	protein kinase, cAMP dependent, catalytic, beta
Exosc4	1233	3,4751	exosome component 4
Mcm6	1302	3,3670	minichromosome maintenance deficient 6 (MIS5 homolog, S. pombe) (S. cerevisiae)
Hnrnpk	1325	3,3309	heterogeneous nuclear ribonucleoprotein K
Mcm3	1328	3,3298	minichromosome maintenance deficient 3 (S. cerevisiae)
Yy1	1356	3,2731	YY1 transcription factor
Exosc6	1371	3,2463	Exosome complex component MTR3
Supt5h	1384	3,2224	Transcription elongation factor SPT5
Mcm4	1390	3,2163	minichromosome maintenance deficient 4 homolog (S. cerevisiae)
Mcm5	1402	3,1902	minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)
Ptbp3	1460	3,0875	polypyrimidine tract binding protein 3
Mcm7	1511	3,0103	minichromosome maintenance deficient 7 (S. cerevisiae)
Mcm2	1570	2,8942	minichromosome maintenance deficient 2 mitotin (S. cerevisiae)
Ywhag	1704	2,6818	14-3-3 protein gamma

Table montrant différentes protéines d'intérêt leur facteur d'enrichissement (log2 abondance +biotine/-biotine) ainsi que leur classement en fonction de ce facteur d'enrichissement dans la condition AID-BirA\* +biotine comparée à AID-BirA\* -biotine. Sont indiqués, en rouge : AID et BirA\*, en violet : facteurs de réparation de l'ADN, en bleu : partenaires nucléaires d'AID, en vert : partenaires cytoplasmiques d'AID, et en gris : les gènes candidats.

Chapitre 2 :

Caractérisation du rôle de Fam72a dans la diversification des immunoglobulines

La réalisation d'un crible fonctionnel CRISPR/Cas9 visant à identifier les acteurs moléculaires impliqués lors de la CI au sein du laboratoire a permis d'établir une liste de gènes candidats. Parmi eux, un des plus intéressants est Fam72a, un gène codant pour une protéine mal caractérisée et n'ayant pas de rôle connu dans la CI. Dans d'autres études, Fam72a a été identifiée comme étant impliquée dans la neurodégénérescence ([Nehar et al., 2009](#)), mais est également surexprimée dans de nombreux cancers courants tels que ceux du colon, du sein, du poumon, de l'utérus et de l'ovaire ([Guo et al., 2008](#); [Heese, 2013](#); [Pramanik et al., 2015](#)). Son expression a également été associée à des gènes régulant la mitose et la formation de centrosomes ([Rahane et al., 2019](#)).

De plus, une étude clé a permis de montrer que Fam72a s'associe au domaine N-terminal d'Ung2, et que cette interaction est dépendante du tryptophane 125 (W125) de Fam72a. En effet, l'utilisation d'un mutant Fam72a<sup>W125R</sup> ne permet pas de co-précipiter Ung2, contrairement à la version sauvage de la protéine ([Guo et al., 2008](#)). Cette étude montre également que la rupture de l'interaction entre Fam72a et Ung2 n'altère ni la localisation cellulaire d'Ung2 *in vivo*, ni son activité enzymatique *in vitro*. De façon intéressante, le génome humain possède 4 gènes paralogues pour FAM72 (nommés FAM72A à FAM72D), contrairement aux génomes des autres mammifères qui ne possèdent qu'une seule copie. Or, 2 de ces paralogues codent pour un W125 (FAM72A et FAM72B) alors que les 2 autres codent pour un R125 (FAM72C et FAM72D), ce qui empêche théoriquement ces isoformes d'interagir avec UNG2 ([Kutzner et al., 2015](#)).

Au vu de l'importance primordiale d'Ung2 dans les mécanismes de diversification des anticorps, nous avons cherché à caractériser le rôle de Fam72a et de son interaction avec Ung2 lors de la CI et de l'HMS. Les résultats que nous avons obtenus sont présentés dans le manuscrit qui suit.

## I- Manuscrit en cours de rédaction –

**Fam72a contrôle la balance entre la réparation de l'ADN mutagène et fidèle durant la diversification des anticorps.**

## **Fam72a controls the balance between error-prone and error-free DNA repair during antibody diversification.**

Mélanie Rogier<sup>1,2,3,4,\*</sup>, Jacques Moritz<sup>1,2,3,4,\*</sup>, Isabelle Robert<sup>1,2,3,4</sup>, Chloé Lescale<sup>5</sup>, Vincent Heyer<sup>1,2,3,4</sup>, Anne-Sophie Thomas-Claudepierre<sup>1,2,3,4</sup>, Arthur Abello<sup>1,2,3,4</sup>, Ludovic Deriano<sup>5</sup> and Bernardo Reina-San-Martin<sup>1,2,3,4,#</sup>.

<sup>1</sup> *Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Illkirch, France.*

<sup>2</sup> *Institut National de la Santé et de la Recherche Médicale (INSERM), U1258, Illkirch, France.*

<sup>3</sup> *Centre National de la Recherche Scientifique (CNRS), UMR7104, Illkirch, France.*

<sup>4</sup> *Université de Strasbourg, Illkirch, France.*

<sup>5</sup> *Genome Integrity, Immunity and Cancer Unit, Equipe Labellisée Ligue Contre Le Cancer, Institut Pasteur, 75015 Paris, France.*

\* These authors contributed equally to this work

# Corresponding author: Bernardo Reina-San-Martin: [reinab@igbmc.fr](mailto:reinab@igbmc.fr)

### **Abstract**

Efficient humoral responses rely on DNA damage, mutagenesis and error-prone DNA repair. B cell receptor diversification through somatic hypermutation (SHM) and class switch recombination (CSR) are initiated by cytidine deamination in DNA mediated by activation induced cytidine deaminase (AID)<sup>1</sup> and by the subsequent excision of the resulting uracils by Uracil DNA glycosylase (UNG) and by mismatch repair (MMR) proteins<sup>2-4</sup>. Although uracils arising in DNA are faithfully repaired<sup>2-7</sup>, it is not known how these pathways are co-opted to generate mutations and double stranded DNA breaks (DSBs) in the context of SHM and CSR<sup>2,4,8</sup>. Here we have performed a genome-wide CRISPR/Cas9 knockout screen for genes involved in CSR. The screen identified FAM72A, a protein that interacts with the nuclear isoform of UNG (UNG2)<sup>9</sup> and that is overexpressed in several cancers<sup>9</sup>. We show that the FAM72A-UNG2 interaction controls the protein levels of UNG2 and that CSR is defective in *Fam72a*<sup>-/-</sup> B cells due to the specific upregulation of UNG2. Moreover, we show that in *Fam72a*<sup>-/-</sup> B cells SHM is reduced by 5-fold and that upregulation of UNG2 results in a skewed mutation pattern. Our results are consistent with a model in which FAM72A interacts with UNG2 to control its physiological level by triggering its degradation. Consequently, deficiency in *Fam72a* leads to supraphysiological levels of UNG2 and enhanced uracil excision, shifting the balance from error-prone to error-free DNA repair. Our findings have potential implications for tumorigenesis, as *Fam72a* overexpression would lead to reduced UNG2 levels, shifting the balance toward mutagenic DNA repair and rendering cells more prone to acquire mutations.

During immune responses, mammalian B cells undergo somatic hypermutation (SHM) and class switch recombination (CSR) to diversify the B cell receptor repertoire and mount efficient and adapted humoral responses<sup>2,3</sup>. SHM modifies antibody affinity for the triggering antigen by introducing point mutations at the variable region of both heavy (IgH) and light (IgL) chains, thereby allowing affinity maturation and the subsequent generation of high-affinity antibodies<sup>2,3</sup>. CSR modulates antibody effector functions by replacing the isotype expressed from IgM/IgD to IgG, IgA or IgE<sup>10</sup>. This is achieved by a deletional recombination event that takes place at the IgH locus between a donor ( $S\mu$ ) and an acceptor switch region ( $Sx$ )<sup>10</sup> that brings into proximity the V region and the exons encoding for a new constant region, thus allowing the expression of an antibody with the same specificity but with a different isotype<sup>10</sup>. Mechanistically, both SHM and CSR are triggered by activation induced cytidine deaminase (AID), an enzyme which deaminates cytosines into uracils in DNA<sup>11</sup>. AID-generated U:G mismatches are processed mainly by the base-excision repair (BER) enzyme uracil-DNA-glycosylase (UNG)<sup>4,12</sup> but also by mismatch repair (MMR)<sup>13</sup> proteins to introduce mutations or double-stranded DNA breaks (DSBs) during SHM and CSR, respectively<sup>2-4</sup>. During SHM, DNA replication over U:G mismatches generates transition mutations at G:C base pairs. Uracil excision by UNG, followed by replication over abasic sites introduces both transition and transversion mutations at C:G base pairs. In addition, single stranded DNA surrounding the U:G mismatch or abasic site can be excised by both MMR and BER to introduce transition and transversion mutations at A:T base pairs. During CSR, cytidine deamination and DNA cleavage on opposite strands generate DSBs<sup>10</sup>, which are a necessary intermediates for recombination<sup>3</sup>. While uracils in DNA are usually faithfully repaired<sup>2-7</sup>, the mechanisms by which AID-induced uracils are instead processed in an error-prone way in the context of antibody diversification through SHM and CSR is largely unknown<sup>2-4</sup>.

To get insight into the molecular mechanisms controlling CSR, we conducted a genome-wide CRISPR/Cas9 knockout screen (**Fig. 1A**). The screen was performed with the second-generation mouse Brie (mBrie)<sup>14</sup> gRNA library and using the IgM<sup>+</sup> murine B cell line CH12, which can be induced to express endogenous levels of AID and undergo CSR from IgM to IgA very efficiently *in vitro* when cultured with TGF- $\beta$ , IL-4 and an anti-CD40 antibody<sup>15</sup>. To conduct the screen, we established a CH12 cell line stably expressing Cas9 (CH12<sup>Cas9</sup>; **Fig. S1A**) and sub-cloned the gRNA library into a retroviral vector (pMX-mBrie; **Fig. S1B**), which efficiently transduces CH12 cells. CH12<sup>Cas9</sup> cells were transduced with the pMX-mBrie gRNA library and selected with puromycin for two weeks to allow



for the Cas9-mediated generation of knockouts. Cells were then stimulated to undergo CSR for 72h. Cells which failed to undergo CSR (which remained IgM+) or that succeed (which became IgA+) were sorted using magnetic beads (**Fig. 1A and S1C**). Genes, whose loss-of-function affect the efficiency of CSR, were identified by looking at gRNA enrichment in the IgM+ versus IgA+ populations through deep sequencing<sup>16</sup> (**Fig. 1A**). The screen identified 654 candidate genes with a significant ( $p < 0.05$ ) gRNA enrichment and having more than 2 effective gRNAs (**Fig. 1B**). The screen was successful as it identified genes known to be required for CSR (*Aicda*, *Ung*, *Trp53bp1*) together with transcription factors driving AID expression (*Batf*, *Pten*, *Irf4*), Non-homologous end joining factors (*Xrcc5*, *Fam35a*), mismatch repair proteins (*Msh2*, *Msh6*, *Pms1*, *Pms2*, *Mlh1*), members of the TGF- $\beta$ , IL-4 and CD40 signaling pathways (*Tgfbr1*, *Tgfbr2*, *Il4ra*, *Il2rg*, *CD40*) and other genes known to be implicated in CSR (*Mediator*, *Trim28*, *Zmynd8*, etc.). Consistent with this, gene ontology analysis revealed a significant enrichment in pathways that are relevant for CSR and SHM (**Fig. 1C**). Importantly, the screen revealed a significant number of genes with no known function in CSR, including *Fam72a* (**Fig. 1B**).

We focused our analysis on *Fam72a* because of its high-position in the gene-ranking (**Fig. 1B**) and since similarly to AID, its expression is enhanced in wild-type splenic primary B cells undergoing CSR (**Fig. S1D**). *Fam72a* encodes a poorly characterized protein, which has been shown to bind to the nuclear isoform of Uracil DNA Glycosylase (UNG2) in a tryptophan 125-dependent manner<sup>9</sup> and which is overexpressed in multiple cancers<sup>9</sup>. To validate the screen and determine whether *Fam72a* is required for CSR, we generated *Fam72a*<sup>-/-</sup> CH12 cells using the high-fidelity Cas9 (Cas9-HF1)<sup>17</sup> (**Fig. S2A and S2B**). Interestingly, we found that the efficiency of CSR was reduced by 50-60 % in four independent *Fam72a*<sup>-/-</sup> CH12 clones when compared to *Fam72a*<sup>+/+</sup> controls (**Fig. 1D**). The CSR defect observed was not due to impaired AID expression (**Fig. S2C**) or defects in proliferation (**Fig. S2D**) or switch region transcription (**Fig. S2E**). We conclude that *Fam72a* is required for efficient CSR in CH12 cells.

To investigate the physiological role of *Fam72a* in CSR and SHM *in vivo*, we obtained a *Fam72a*<sup>-/-</sup> mouse model (**Fig. S3A**) generated by the knockout mouse program (KOMP)<sup>18</sup>. *Fam72a*<sup>-/-</sup> mice gave offspring at Mendelian ratios and showed no obvious deleterious phenotype. *Fam72a*<sup>-/-</sup> mice did not display any specific defects in B cell development, and all B cell populations were found to be represented in normal proportions and numbers in the bone marrow and the spleen (**Fig. 2A, 2B,**

**S3B and S3C**). This suggests that FAM72A is dispensable for RAG1/2- and non-homologous end joining (NHEJ)-dependent V(D)J recombination.

To determine whether *Fam72a*<sup>-/-</sup> B cells display a defect in CSR, we purified resting splenic B cells, labeled them with CFSE to track proliferation and cultured them under conditions that induce CSR to different isotypes. We found that *Fam72a*<sup>-/-</sup> B cells display a 40-70% reduction in the efficiency of CSR to all isotypes tested (**Fig. 2C and 2D**), which was not due defects in proliferation (**Fig. 2C**), AID expression (**Fig. 2E**), or switch region transcription (**Fig. 2F**). We conclude that deficiency in *Fam72a* results in a B cell intrinsic CSR defect in mouse primary B cells.

To demonstrate that the CSR defect observed is due to deficiency in *Fam72a* and investigate the functional relevance of the FAM72A-UNG2 interaction, we re-expressed FAM72A and the FAM72A<sup>W125R</sup> or FAM72A<sup>W125A</sup> UNG2-binding defective mutants<sup>9</sup> in *Fam72a*<sup>-/-</sup> CH12 B cells. We found that the CSR defect observed was indeed due to the absence of FAM72A, as re-expression of FAM72A (**Fig. S2B**) rescued the CSR defect (**Fig. 3A and 3C**), ruling out unlikely Cas9-HF1 off-target effects. Interestingly, we found that contrary to wildtype FAM72A, the FAM72A<sup>W125R</sup> or FAM72A<sup>W125A</sup> mutants were unable to rescue the CSR defect (**Fig. 3A and 3C**). Furthermore, we found that the residual CSR activity observed in *Fam72a*<sup>-/-</sup> B cells is dependent on the catalytic activity of UNG, as expression of an UNG-inhibitor (*Ugi*)<sup>19</sup> abolishes CSR (**Fig. 3B and 3C**). We conclude that the CSR defect observed is due to the lack of FAM72A, that the FAM72A-UNG2 interaction is of functional relevance for CSR and that the residual CSR activity is dependent on the catalytic activity of UNG.

Alternative promoter usage and splicing of the *Ung* gene generates a mitochondrial (UNG1) and a nuclear (UNG2) isoform<sup>20</sup> (**Fig. S2F**), both of which have been shown to be able to sustain CSR<sup>21</sup>. To examine the relative contribution of both these isoforms relative to FAM72A, we generated *Ung1*<sup>-/-</sup>, *Ung2*<sup>-/-</sup>, *Ung*<sup>-/-</sup>, *Fam72a*<sup>-/-</sup> *Ung1*<sup>-/-</sup> and *Fam72a*<sup>-/-</sup> *Ung2*<sup>-/-</sup> CH12 cell clones using CRISPR/Cas9-HF1 (**Fig. S2F**) and tested their ability to undergo CSR. As shown before<sup>12,21</sup>, while *Ung1*<sup>-/-</sup> and *Ung2*<sup>-/-</sup> cells underwent CSR at wildtype levels (**Fig. 3D and S2G**), CSR in *Ung*<sup>-/-</sup> CH12 cells was completely abolished (**Fig. 3D and S2G**). Surprisingly, while *Fam72a*<sup>-/-</sup> *Ung2*<sup>-/-</sup> rescued the CSR defect observed in *Fam72a*<sup>-/-</sup> CH12 cells, *Fam72a*<sup>-/-</sup> *Ung1*<sup>-/-</sup> did not (**Fig. 3D and S2G**). This result prompted us to further explore the significance of the FAM72A-UNG2 interaction. For this, we analyzed the expression level of UNG1 and UNG2 in the different single and double knockout CH12 cell lines by

Western blot. Surprisingly, we found that loss-of-function of *Fam72a* results in a significant and specific up-regulation of the UNG2 isoform (**Fig. 3E**), which accumulated on chromatin (**Fig. 3F**). This suggests that FAM72A controls the protein level of UNG2 and its access to chromatin. This phenotype, which was not at the mRNA level (**Fig. S2H**), was also observed in *Fam72a*<sup>-/-</sup> primary B cells (**Fig. 3G**). Furthermore, when we overexpressed FAM72A in *Fam72a*<sup>-/-</sup> CH12 cells (**Fig. S2B**), the up-regulation of UNG2 was almost abolished (**Fig. 3E**). Significantly, overexpression of the FAM72A<sup>W125A</sup> mutant (**Fig. S2B**), which does not interact with UNG2<sup>9</sup>, did not down-regulate UNG2 (**Fig. 3E**). We conclude that FAM72A specifically regulates the protein level of the UNG2 isoform by controlling its degradation and that the CSR defect observed in *Fam72a*<sup>-/-</sup> and *Fam72a*<sup>-/-</sup> *Ung1*<sup>-/-</sup> cells is due to increased levels of the UNG2 isoform. Consistent with this, we found that overexpression of UNG2 in wildtype CH12 cells was sufficient to suppress CSR (**Fig. 3H**). Therefore, it appears that the maintenance of the physiological levels of UNG2 is of critical importance for the efficiency of CSR and explains why mutants, with defective catalytic activity of UNG2, reconstitute CSR better than wildtype UNG2 in *Ung*<sup>-/-</sup> B cells<sup>22-24</sup>. Our results suggest that FAM72A, by controlling the level of UNG2, influences the usage of error-prone *versus* error-free DNA repair in response to AID-induced uracils in DNA.

Activated B cells expressing AID display mutations at switch regions during CSR<sup>25</sup> and at a region spanning immunoglobulin variable regions during SHM<sup>3</sup>. To determine whether deficiency in *Fam72a* and the concomitant up-regulation of UNG2 have an influence in the frequency and pattern of AID-induced mutations, we sequenced the 5' end of the S $\mu$  switch regions in *Fam72a*<sup>-/-</sup> and control CH12 cells that were stimulated to undergo CSR for three days. Interestingly, we found that the mutation frequency was reduced by 50% in *Fam72a*<sup>-/-</sup> CH12 cells when compared to controls (**Fig. S4A**). To confirm whether deficiency in *Fam72a* results in lower AID-induced mutation frequency, we analyzed the J<sub>H</sub>4 intron (J<sub>H</sub>4i), a sequence which gets heavily mutated by AID during SHM *in vivo*<sup>26</sup>, in germinal center B cells isolated from the Peyer's patches of unimmunized *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> mice (**Fig. 4A**). Interestingly, we found that the mutation frequency at the J<sub>H</sub>4i sequence was drastically reduced (five-fold) in *Fam72a*<sup>-/-</sup> B cells, when compared to controls (**Fig. 4B, 4C and S4B**). Consistent with UNG2 up-regulation, we found that the percentage of transition mutations at C:G base pairs, which are generated by replication over U:G mismatches were reduced from 57 % in *Fam72a*<sup>+/+</sup> to 16 % in *Fam72a*<sup>-/-</sup> B cells (**Fig. 4C and S4B**). Conversely, transversion mutations at C:G base pairs, which are generated through replication over UNG-generated abasic sites, increased from 43 % in *Fam72a*<sup>+/+</sup> to 84 % in *Fam72a*<sup>-/-</sup> B cells (**Fig. 4C and**

**S4B).** This phenotype is diametrically opposed to *Ung*<sup>-/-</sup> B cells<sup>12</sup>, in which transition mutations at C:G base pairs reaches 95 % and transversion mutations at C:G base pairs are suppressed to 5 %<sup>12</sup>. These results support the hypothesis that in *Fam72a*-deficient B cells, where UNG2 is upregulated, AID-induced uracils are more efficiently excised from DNA by UNG2 and that error-free repair is enforced. Furthermore, enhanced uracil excision would result in fewer U:G mismatches, rendering MMR less effective in generating mutations at A:T base pairs surrounding the deaminated cytosines. Therefore, the level of UNG2 needs to be tightly controlled to trigger error-prone DNA repair, which is essential for antibody diversification through SHM and CSR.

CSR is sensitive to the expression level of AID and concomitantly to the amount of AID-induced DSBs, which are obligatory intermediates<sup>3,10</sup>. Indeed, AID is haplo-insufficient and CSR is reduced in AID<sup>+/-</sup> B cells<sup>27</sup>. The CSR defect observed in *Fam72a*<sup>-/-</sup> B cells could be explained by enforced error-free repair due to the up-regulation of UNG2, which in turn would lead to the generation of fewer DSBs at switch regions. In wildtype B cells, a small fraction of AID-induced DSBs can be detected in the form of chromosome breaks or translocations<sup>28</sup>. To determine whether up-regulation of UNG2 in the absence of FAM72A might influence the level of *IgH* DSBs, we conducted fluorescence in situ hybridization experiments using two *IgH*-specific probes and chromosome 12 paint in activated AID<sup>-/-</sup>, *Ung*<sup>-/-</sup>, *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> cells reconstituted (or not) with FAM72A or FAM72A<sup>W125R</sup> and quantified aberrant metaphases in these cells (**Fig. 4D, 4E and Table S1**). We found lower levels of aberrant metaphases in *Fam72a*<sup>-/-</sup> cells or *Fam72a*<sup>-/-</sup> cells reconstituted with FAM72A<sup>W125R</sup> as compared to wild type and *Fam72a*<sup>-/-</sup> cells reconstituted with wild type FAM72A, suggesting that FAM72A-mediated UNG2 upregulation controls the level of AID-induced DSBs during CSR. We conclude that the CSR defect observed in the absence of *Fam72a* is due to the inefficient generation of AID-induced DSBs.

Our results are consistent with a model in which FAM72A interacts with UNG2 to control its physiological level by triggering its degradation. Consequently, deficiency in *Fam72a* leads to the specific upregulation of the UNG2 isoform and its accumulation on chromatin. This is sufficient to enforce uracil excision, resulting in a reduction in the efficiency of SHM and CSR. It is possible that supraphysiological levels of UNG2 in the absence of FAM72A tilt the balance toward error-free BER by favoring the recruitment of polymerase  $\beta$ <sup>5,6,29,30</sup>. A shift in the balance from error-free to error-prone DNA repair in response to uracils in DNA has significant implications in the onset of cancer. As *Fam72a* is overexpressed in several types of tumors and transformed cell lines<sup>9</sup> and as *Ung*<sup>-/-</sup>

mice develop B cell lymphomas<sup>31-33</sup>, it is possible that *Fam72a* overexpression would suppress the levels of UNG2 leading to inefficient uracil excision, enforcing error-prone DNA repair and making these cells more susceptible to accumulate mutations and hence more prone to tumorigenesis.

### **Acknowledgements**

B. Kavli for the anti-UNG1/2 antibody; L. Brino for help with gRNA library sub-cloning and CRISPR/Cas9 screen analysis; B. Jost, C. Thibault-Carpentier, D. Plassard and D. Dembele for gRNA high-throughput sequencing; B. Heller for cell culture; C. Ebel and M. Philipps for cell-sorting; M. Selloum, D. Ali-Hadji and I. Gonçalves Da Cruz for animal care. J. Moritz was supported by the Ministère de l'Enseignement Supérieur et de la Recherche, France and the Fondation ARC. M. Rogier was supported by La Ligue Contre le Cancer. This study was supported by grants from the Fondation Recherche Médicale (Equipe FRM EQU201903007818), Institut National du Cancer (INCa\_13852), Fondation ARC (ARCPJA32020060002061) and by the grant ANR-10-LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the program Investissements d'Avenir labelled ANR-10-IDEX-0002-02.

### **Author Contributions**

B.R.S.M. conceived the project. M.R., J.M., I.R., C.L., A-S.T.C-P., A.A. and V.H. performed experiments and analyzed the data. M.R., J.M. and B.R.S.M. wrote the manuscript. L.D. and B.R.S.M directed overall research and edited the manuscript.

Supplementary Information is available for this paper.

Correspondence and requests for materials should be addressed to B.R.S.M.

## Literature Cited

- 1 Neuberger, M. S., Harris, R. S., Di Noia, J. & Petersen-Mahrt, S. K. Immunity through DNA deamination. *Trends Biochem Sci* **28**, 305-312, doi:S0968000403001117 [pii] (2003).
- 2 Feng, Y., Seija, N., Di Noia, J. M. & Martin, A. AID in Antibody Diversification: There and Back Again. *Trends in Immunology* **41**, 586-600 (2020).
- 3 Methot, S. P. & Di Noia, J. M. Molecular Mechanisms of Somatic Hypermutation and Class Switch Recombination. *Adv Immunol* **133**, 37-87, doi:10.1016/bs.ai.2016.11.002 (2017).
- 4 Stratigopoulou, M., van Dam, T. P. & Guikema, J. E. J. Base Excision Repair in the Immune System: Small DNA Lesions With Big Consequences. *Front Immunol* **11**, doi: 10.3389/fimmu.2020.01084 (2020).
- 5 Krokan, H. E. & Bjoras, M. Base Excision Repair. *Csh Perspect Biol* **5**, doi: 10.1101/cshperspect.a012583 (2013).
- 6 Krokan, H. E., Drablos, F. & Slupphaug, G. Uracil in DNA - occurrence, consequences and repair. *Oncogene* **21**, 8935-8948, doi:10.1038/sj.onc.1205996 (2002).
- 7 Krokan, H. E. *et al.* Error-free versus mutagenic processing of genomic uracil-Relevance to cancer. *DNA Repair* **19**, 38-47, doi:10.1016/j.dnarep.2014.03.028 (2014).
- 8 Methot, S. P. *et al.* Consecutive interactions with HSP90 and eEF1A underlie a functional maturation and storage pathway of AID in the cytoplasm. *J Exp Med* **212**, 581-596, doi:10.1084/jem.20141157 (2015).
- 9 Guo, C. *et al.* Ugene, a newly identified protein that is commonly overexpressed in cancer and binds uracil DNA glycosylase. *Cancer Res* **68**, 6118-6126, doi:10.1158/0008-5472.CAN-08-1259 (2008).
- 10 Stavnezer, J. & Schrader, C. E. IgH chain class switch recombination: mechanism and regulation. *J Immunol* **193**, 5370-5378, doi:10.4049/jimmunol.1401849 (2014).
- 11 Petersen-Mahrt, S. K., Harris, R. S. & Neuberger, M. S. AID mutates E. coli suggesting a DNA deamination mechanism for antibody diversification. *Nature* **418**, 99-103, doi:10.1038/nature00862 (2002).
- 12 Rada, C. *et al.* Immunoglobulin Isotype Switching Is Inhibited and Somatic Hypermutation Perturbed in UNG-Deficient Mice. *Curr Biol* **12**, 1748-1755 (2002).
- 13 Zanotti, K. J. & Gearhart, P. J. Antibody diversification caused by disrupted mismatch repair and promiscuous DNA polymerases. *DNA Repair* **38**, 110-116, doi:10.1016/j.dnarep.2015.11.011 (2016).
- 14 Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* **34**, 184-191, doi:10.1038/nbt.3437 (2016).
- 15 Nakamura, M. *et al.* High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. *Int Immunol* **8**, 193-201 (1996).
- 16 Joung, J. *et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nature protocols* **12**, 828-863, doi:10.1038/nprot.2017.016 (2017).
- 17 Kleinstiver, B. P. *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490-495, doi:10.1038/nature16526 (2016).
- 18 KOMP. <[https://www.mmrrc.org/catalog/sds.php?mmrrc\\_id=47653](https://www.mmrrc.org/catalog/sds.php?mmrrc_id=47653)> (
- 19 Di Noia, J. & Neuberger, M. S. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature* **419**, 43-48 (2002).
- 20 Nilsen, H. *et al.* Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the UNG gene. *Nucleic acids research* **25**, 750-755, doi:DOI 10.1093/nar/25.4.750 (1997).
- 21 Sarno, A. *et al.* Uracil-DNA glycosylase UNG1 isoform variant supports class switch recombination and repairs nuclear genomic uracil. *Nucleic acids research* **47**, 4569-4585, doi:10.1093/nar/gkz145 (2019).
- 22 Begum, N. A. *et al.* Uracil DNA glycosylase activity is dispensable for immunoglobulin class switch. *Science* **305**, 1160-1163 (2004).
- 23 Di Noia, J. M. *et al.* Dependence of antibody gene diversification on uracil excision. *J Exp Med* **204**, 3209-3219 (2007).



- 24 Begum, N. A. *et al.* Requirement of non-canonical activity of uracil DNA glycosylase for class switch recombination. *J Biol Chem* **282**, 731-742 (2007).
- 25 Petersen, S. *et al.* AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. *Nature* **414**, 660-665, doi: 10.1038/414660a (2001).
- 26 Jolly, C. J., Klix, N. & Neuberger, M. S. Rapid methods for the analysis of immunoglobulin gene hypermutation: application to transgenic and gene targeted mice. *Nucleic acids research* **25**, 1913-1919 (1997).
- 27 Sernandez, I. V., de Yebenes, V. G., Dorsett, Y. & Ramiro, A. R. Haploinsufficiency of activation-induced deaminase for antibody diversification and chromosome translocations both in vitro and in vivo. *PLoS one* **3**, e3927, doi:10.1371/journal.pone.0003927 (2008).
- 28 Dev, H. *et al.* Shieldin complex promotes DNA end-joining and counters homologous recombination in BRCA1-null cells. *Nat Cell Biol* **20**, 954-965, doi:10.1038/s41556-018-0140-1 (2018).
- 29 Wu, X. & Stavnezer, J. DNA polymerase beta is able to repair breaks in switch regions and plays an inhibitory role during immunoglobulin class switch recombination. *J Exp Med* **204**, 1677-1689 (2007).
- 30 Schrader, C. E., Guikema, J. E., Wu, X. & Stavnezer, J. The roles of APE1, APE2, DNA polymerase beta and mismatch repair in creating S region DNA breaks during antibody class switch. *Philos Trans R Soc Lond B Biol Sci* **364**, 645-652, doi:10.1098/rstb.2008.0200 (2009).
- 31 Nilsen, H. *et al.* Gene-targeted mice lacking the Ung uracil-DNA glycosylase develop B-cell lymphomas. *Oncogene* **22**, 5381-5386, doi:10.1038/sj.onc.1206860 (2003).
- 32 Andersen, S. *et al.* Monoclonal B-cell hyperplasia and leukocyte imbalance precede development of B-cell malignancies in uracil-DNA glycosylase deficient mice. *DNA Repair* **4**, 1432-1441, doi:10.1016/j.dnarep.2005.08.004 (2005).
- 33 Nilsen, H., An, Q. & Lindahl, T. Mutation frequencies and AID activation state in B-cell lymphomas from Ung-deficient mice. *Oncogene* **24**, 3063-3066, doi:10.1038/sj.onc.1208480 (2005).
- 34 Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome biology* **15**, 554, doi:10.1186/s13059-014-0554-4 (2014).
- 35 Miyazaki, K. MEGAWHOP cloning: a method of creating random mutagenesis libraries via megaprimer PCR of whole plasmids. *Methods Enzymol* **498**, 399-406, doi:10.1016/B978-0-12-385120-8.00017-6 (2011).
- 36 Maccarthy, T., Roa, S., Scharff, M. D. & Bergman, A. SHMTool: a webserver for comparative analysis of somatic hypermutation datasets. *DNA Repair (Amst)* **8**, 137-141, doi:10.1016/j.dnarep.2008.09.006 (2009).
- 37 Engler, C., Gruetzner, R., Kandzia, R. & Marillonnet, S. Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS one* **4**, e5553, doi:10.1371/journal.pone.0005553 (2009).
- 38 Motohashi, K. A simple and efficient seamless DNA cloning method using SLICE from Escherichia coli laboratory strains and its application to SLiP site-directed mutagenesis. *Bmc Biotechnol* **15**, doi: 10.1186/s12896-015-0162-8 (2015).
- 39 Lescale, C. *et al.* RAG2 and XLF/Cernunnos interplay reveals a novel role for the RAG complex in DNA repair. *Nature communications* **7**, 10529, doi:10.1038/ncomms10529 (2016).
- 40 Jeevan-Raj, B. P. *et al.* Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. *J Exp Med* **208(8)**, 1649-1660, doi:10.1084/jem.20110118 (2011).
- 41 Holden, P. & Horton, W. A. Crude subcellular fractionation of cultured mammalian cell lines. *BMC Research Notes* **2**, doi:doi:10.1186/1756-0500-2-243 (2009).

## Methods

### ***CH12<sup>Cas9</sup> cell generation.***

CH12 cells were transduced with retroviral supernatants obtained by transfecting Bosc23 cells with a retrovirus (pMX-26) expressing a Cas9-mCherry-P2A-Hygromycin cassette (**Fig. S1A**). After 72h of hygromycin selection (Sigma; 300 µg/mL), mCherry+ cells were sorted into 96-well plates using a cell sorter FACS ARIA Fusion (Becton Dickinson) and cultured for 10 days with hygromycin (Sigma; 300 µg/mL). Individual clones were then tested for their ability to express Cas9 by western-blot (**Fig. S1A**), and their ability to undergo efficient CSR after stimulation (**Fig. S1A**). To test the functionality of Cas9, clones were transfected with gRNAs known to be effective (see Table S2 for gRNA sequences). Three days later, PCR was performed to identify clones with Cas9-mediated deletions (**Fig. S1A**). Clone #11, which had the highest Cas9 activity and which displayed a CSR efficiency similar to the parental CH12 cell line was selected to conduct the screen.

### ***Sub-cloning of the mBrie gRNA library.***

The mouse Brie CRISPR knockout pooled library (lentiCRISPRv2 backbone; a gift from David Root and John Doench (Addgene #73633) was sub-cloned into the pMX-28 retroviral vector (**Fig. S1B**) as described<sup>16</sup> to generate the pMX-mBrie gRNA library. Briefly, the U6p-gRNA-scaffold cassette was amplified by PCR (see Table S2 for primers), digested with BamHI and a NotI restriction enzymes and ligated into pMX-28 (**Fig. S1B**). Stbl4 electrocompetent cells (ThermoFisher) were electroporated using Eporator Eppendorf (Program1-1700V), plated on ampicillin agar plates and incubated at 37°C for 15 hours. Plasmid DNA was extracted and the U6-gRNA-scaffold cassette was amplified by PCR to add sequencing adaptors and barcodes (see Table S2 for primers) and then analyzed by deep sequencing to determine gRNA representation and library uniformity as described<sup>16</sup>.

### ***CRISPR/Cas9 knockout screen.***

CH12<sup>Cas9</sup> cells were transduced with the pMX-mBrie gRNA library in quadruplicate at a multiplicity of infection (MOI) of 0.3 and with a 300X coverage, as previously described<sup>16</sup>. Transduced cells were selected with puromycin (Sigma; 0.5 µg/mL) and hygromycin (Sigma; 300 µg/mL) for 15 days. Selected cells were then induced to undergo CSR with TGF-β (1 ng/ml; R&D Systems Europe), IL-4 (5 ng/ml; Peprotech) and an anti-CD40 antibody (200 ng/ml; eBioscience) for 72h. IgM+ and IgA+ cells were sorted using anti-IgM and anti-IgA coupled magnetic beads (Mitentyi). Population purity was



assessed by flow cytometry (**Fig. S1C**) using a Fortessa flow cytometer (Becton Dickinson). Genomic DNA from 24 million of cells was extracted using phenol/chloroform and subjected to PCR (1x (95°C for 2min)/ 28-31X (95°C for 15s, 65°C for 20s, 72°C for 30s) and 72°C for 3 min) to amplify gRNA sequences using staggered primers having Illumina adaptors and barcodes (see Table S2 for primers). Multiplexed samples were submitted to high-throughput sequencing (1x50 bp) on an Illumina HiSeq4000 sequencer at the GenomEast sequencing platform of IGBMC. gRNA counts were extracted from raw data using the PoolQ software from the Broad Institute. gRNA counts were normalized according as described<sup>16</sup>. Ranking of candidate genes was performed using MaGeCK<sup>34</sup>. Candidate genes ( $p < 0.05$ ) and having more than 2 effective gRNAs were selected.

### ***Cell culture.***

Bosc23 cells were cultured in DMEM supplemented with glucose (4,5g/L), 10% of heat-inactivated fetal calf serum, penicillin-streptomycin (100U/mL) and sodium pyruvate (1mM). CH12 cells and primary B cells were cultured in RPMI supplemented with 10% of heat-inactivated fetal calf serum, HEPES (10 mM), penicillin-streptomycin (100 U/mL), Sodium Pyruvate (1 mM) and  $\beta$ -mercaptoethanol (50  $\mu$ M).

### ***CSR assays.***

CH12 cells were cultured for 72h in the presence of TGF- $\beta$  (1 ng/ml; R&D Systems Europe), IL-4 (5 ng/ml; Peprotech) and an anti-CD40 antibody (200 ng/ml; eBioscience). Cells were then stained with an anti-IgA-PE antibody (Southern Biotech) to assess CSR by flow cytometry. Prior to analysis, DAPI was added to discriminate dead cells. Samples were analyzed using a Fortessa flow cytometer (Becton Dickinson) and the FlowJo software.

### ***Primary B cell cultures.***

Splenic resting B cells were purified using anti-CD43 magnetic beads (Miltenyi), labeled with CFSE (COMPANY) and cultured from 4 days with a combination of LPS (25  $\mu$ g/ml, Sigma-Aldrich), IL-4 (25  $\mu$ g/ml, Preprotech), anti-IgD-Dextran (6 ng/ml, Fina Biosolutions), IFN- $\gamma$  (100 ng/ml, Preprotech), IL-5 (5 ng/ml, R&D Systems), TGF- $\beta$  (3 ng/ml, R&D Systems) and Retinoic Acid (RA; 0.3 ng/ml, Sigma-Aldrich).

### ***RTq-PCR.***

RNA and cDNA were prepared using standard techniques. qPCR was performed in triplicates using Roche LightCycler 480 Probes Master mix Universal Probe Library (UPL) in combination with appropriate UPL probes (**Table S2**). Transcript quantities were calculated relative to standard curves and normalized to HPRT or Ig $\beta$  mRNA.

### ***Retroviral transductions.***

CH12 cells were transduced with retroviral supernatants obtained by transfecting Bosc23 cells with an empty retrovirus (pMX-PIE; Puromycin-IRES-EGFP) or expressing mFam72a (pMX-Fam72a), o mFam72a<sup>W125R</sup> (pMX-mFam72a<sup>W125R</sup>), mFam72a<sup>W125A</sup> (pMX-Fam72a<sup>W125A</sup>), mUng2 (pMX-mUng2) or Ugi (pMX-Ugi). Transduced cells were then selected with puromycin (1  $\mu$ g/ml) for 7 days and submitted to CSR assays.

### ***Generation of CH12 knockout clones.***

CH12 cells were transfected by electroporation using the Neon transfection System (ThermoFisher) with a plasmid expressing one or two gRNAs targeting a critical exon (see Table S2 for gRNA sequences) and co-expressing the high-fidelity Cas9 nuclease<sup>17</sup> coupled to EGFP. 24h after transfection, individual EGFP-positive cells were sorted into 96-well plates using a cell sorter FACS ARIA Fusion (Becton Dickinson) and cultured for 10 days. Clones were then genotyped by PCR and sequencing.

### ***S $\mu$ and JH4i somatic hypermutation analysis.***

For S $\mu$  mutation analysis, genomic DNA was extracted from CH12 cells cultured for 3 days with TGF- $\beta$  (1 ng/ml; R&D Systems Europe), IL-4 (5 ng/ml; Peprotech) and an anti-CD40 antibody (200 ng/ml; eBioscience). For JH4i mutation analysis, genomic DNA was extracted from B cells isolated from the Peyer's patches and sorted by flow cytometry using anti-B220-PE-Cy7 (eBiosciences), anti-Fas-PE (BD Pharmingen) and anti-GL7-Pacific Blue (BioLegend) antibodies. S $\mu$  and JH4i sequences were amplified by PCR (1x 98°C for 30s, 35x (98°C for 10s, 70°C for 10s and 72°C for 30s) and 1x 72°C for 5 min) using the Q5 polymerase (New England BioLabs) and cloned into pUC57 using the MEGAWHOP<sup>35</sup> method. Inserts were sequenced (Sanger sequencing) using the M13 Forward universal primer. Sequences were aligned with Lasergene (DNASTAR) and analyzed with the SHMTool<sup>36</sup> server.

### ***Cloning.***

All gRNA/Cas9HF1 plasmids were generated through golden gate cloning<sup>37</sup>. cDNAs were cloned into the pMX-PIE retrovirus through SLICE<sup>38</sup>.

### ***Mice.***

*Fam72a*<sup>-/-</sup> mice (C57BL6) were generated by the knockout mouse program (KOMP)<sup>18</sup>. Mice were bred under pathogen-free (SPF) conditions. In all experiments 8-12 week old age-matched littermates were used. Animal work was performed under protocols approved by an ethics committee (APAFIS#23104-2019112915476749).

### ***IgH FISH.***

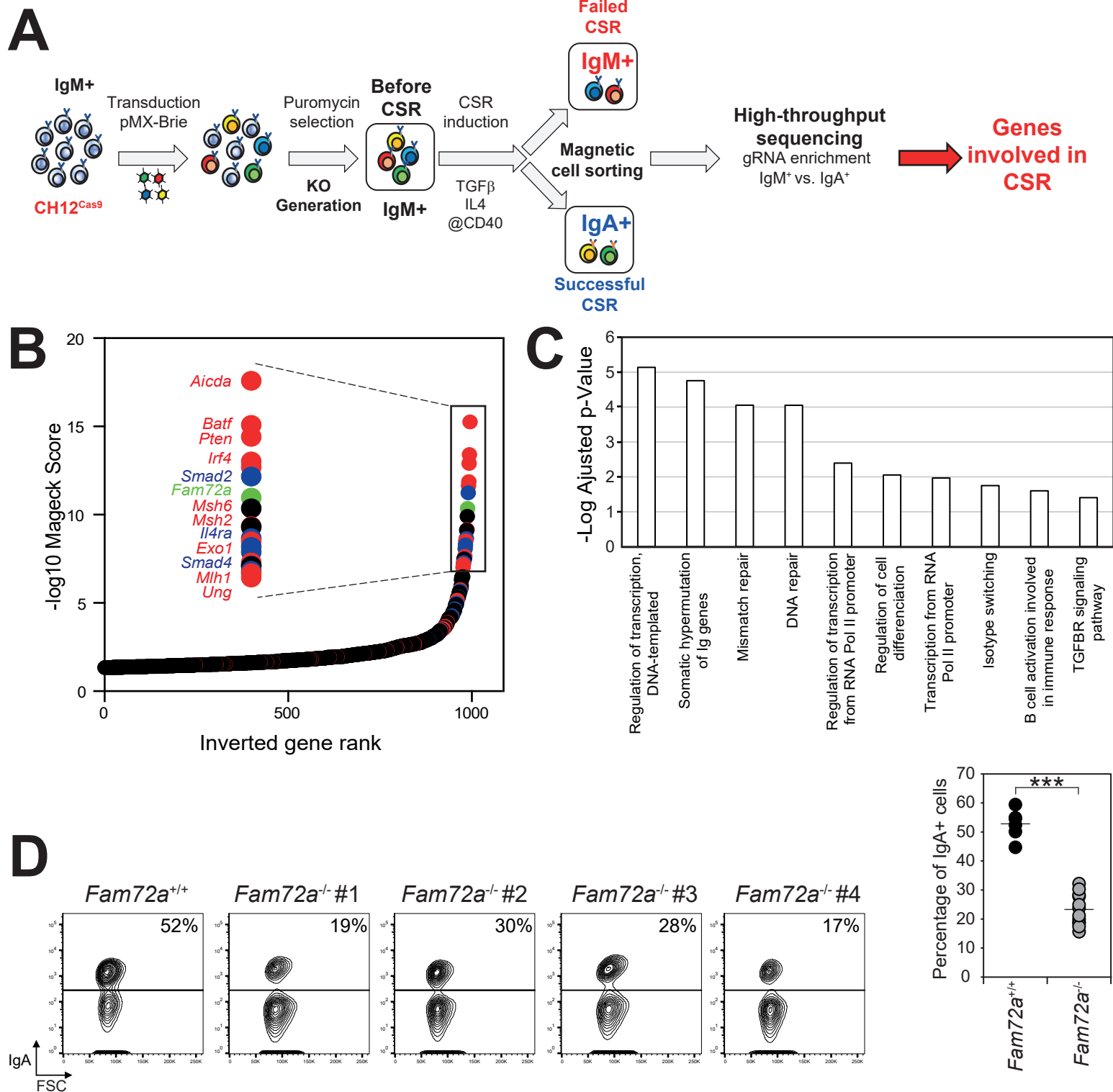
Metaphases were prepared using standard procedures<sup>39</sup>. DNA FISH on metaphases spreads was performed as previously described<sup>39</sup> using BAC probes RP24-134G24 (5' Igh C) and RP24-386J17 (3' Igh V) and XCyting Mouse Chromosome 12 (Orange) paint from MetaSystems. Metaphases were imaged using a ZEISS Axiolmager.Z2 microscope and the Metafer automated capture system (MetaSystems), and counted manually.

### ***Western blot analysis.***

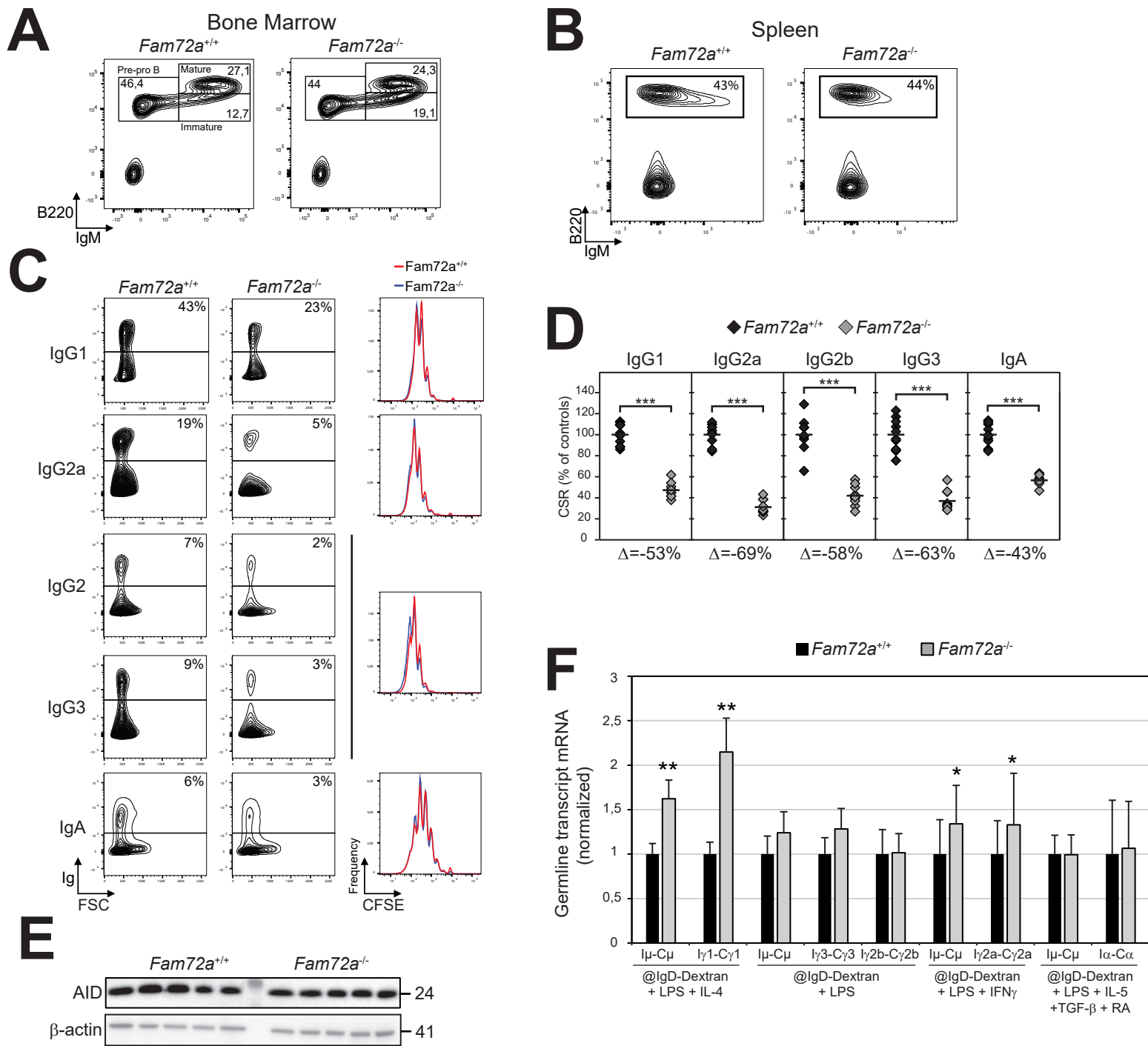
Proteins extracts were prepared using standard techniques. Proteins were separated by SDS-PAGE using gradient gels (4-12%; Invitrogen), transferred to PVDF membranes (Immobilon; Millipore) and analyzed using anti-AID<sup>40</sup>, anti- $\beta$ -Actin (Sigma), anti-UNG1/2 (Gift from B. Kavli) antibodies. Cell fractionation experiments were performed as previously described<sup>41</sup>.

### ***B cell development.***

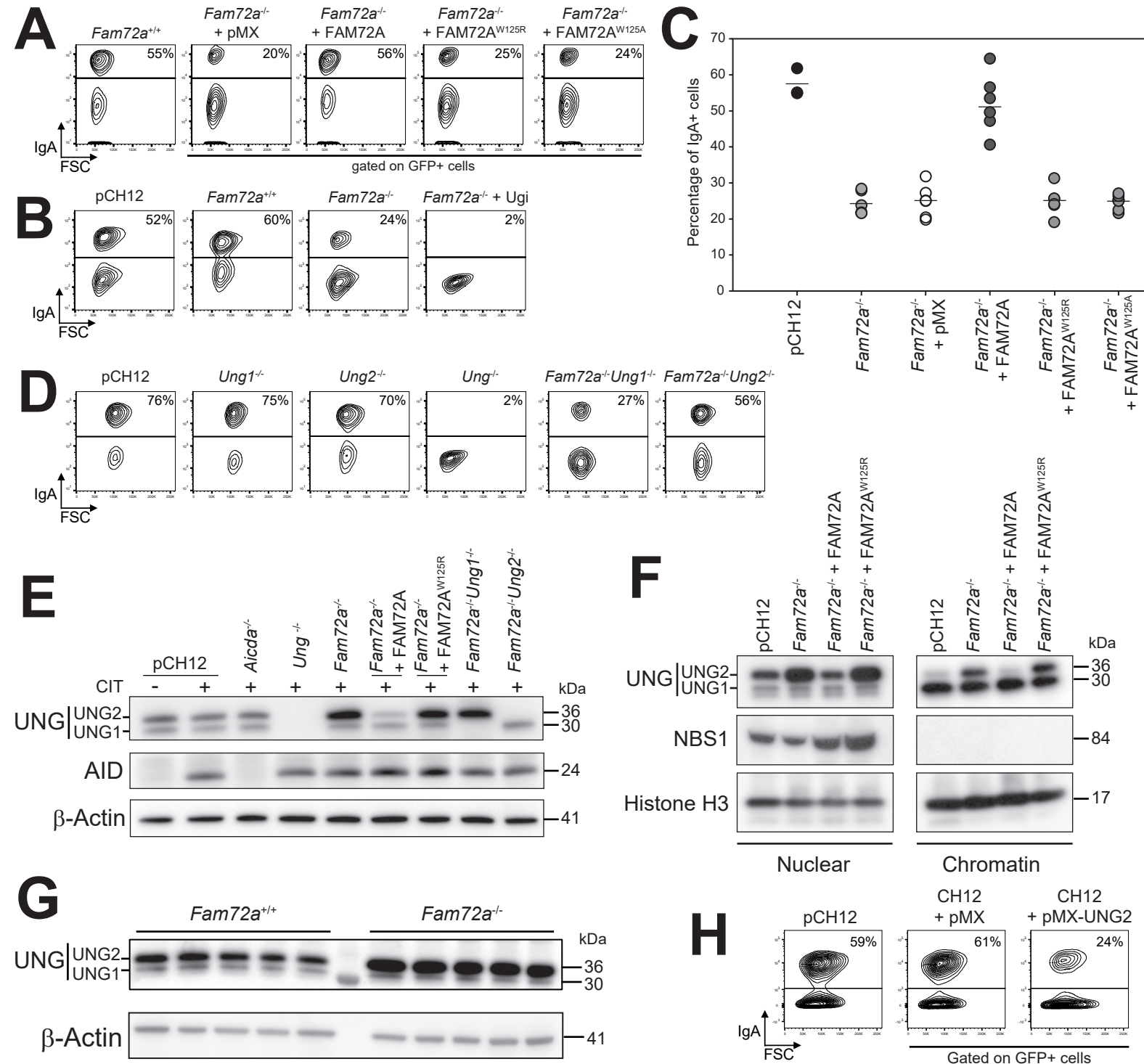
Total splenic or bone marrow cells were labeled with suitable antibodies and analyzed on an FACS-Fortessa flow cytometer (Becton Dickinson).



**Figure 1. A genome-wide CRISPR/Cas9 knockout screen for genes involved in CSR identifies *Fam72a*.** (A) Schematic overview of the CRISPR/Cas9 screen. (B) Graph depicting the gene significance score of the ranking ( $-\log_{10}$  Score Mageck) calculated with the MAGECK algorithm, plotted against the logarithmic value to the base 2 of the ratio of IgM over IgA read counts (LFC) for each gene represented by a dot. Genes known to be involved in CSR are depicted in red, TGF- $\beta$ , IL-4, CD40 and BCR signaling pathways are depicted in blue. Genes with no known function in CSR are in black. (C) Gene ontology analysis of genes identified in the screen. (D) Flow cytometry analysis of IgA expression in *Fam72a*<sup>+/+</sup> and four independent *Fam72a*<sup>-/-</sup> CH12 cell clones cultured for 3 days with TGF- $\beta$ , anti-CD40 antibody and IL4. The percentage of IgA-expressing cells from 3 independent experiments is shown on the right. p-value was determined using two-tailed Student's t-test; \*\*\*p<0,0005.



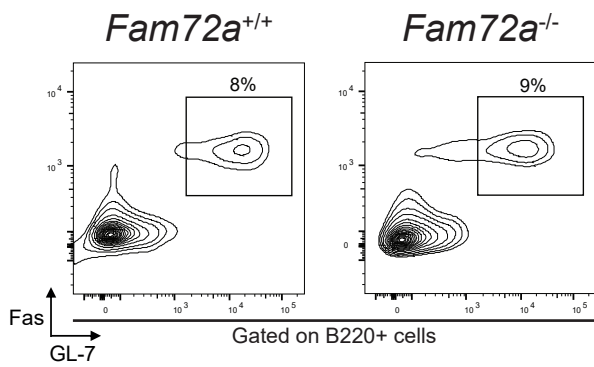
**Figure 2. CSR is defective in B cells from *Fam72a*<sup>-/-</sup> mice.** Flow cytometry analysis of bone marrow (**A**) and spleen (**B**) cells from *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> mice. Staining antibodies and percentage of cells within gates are indicated. (**C**) Flow cytometry analysis of Ig expression in *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> splenic B cells cultured for 96h with either LPS + IL-4 + anti-IgD-Dextran (CSR to IgG1), LPS + IFN- $\gamma$  + anti-IgD-Dextran (CSR to IgG2a) or LPS + anti-IgD-Dextran (CSR to IgG2b and IgG3) or LPS + IL-5 + TGF- $\beta$  + RA + anti-IgD-Dextran (CSR to IgA). The percentage of Ig-expressing cells is indicated. Representative plots from 2 experiments with 5 mice of each genotype are shown. CFSE dilution analysis is shown on the right. (**D**) Percentage of CSR in *Fam72a*<sup>-/-</sup> B cells relative to *Fam72a*<sup>+/+</sup> B cells for the different isotypes from 10 independent mice is shown on the right. A horizontal line indicates mean values. The difference ( $\Delta$ ) in the percentage of CSR for each isotype between *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> B cells is indicated below the plot. p-values were determined using two-tailed Student's t test (\*\*\*: p<0.0001). (**E**) Western blot analysis for AID and  $\beta$ -Actin in *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> splenic B cells cultured with LPS, IL-4 and anti-IgD-Dextran for 72h. (**F**) Real-time qPCR analysis for germline transcripts at donor (GLT $\mu$ ) and acceptor switch regions (GLT $\gamma$ 3, GLT $\gamma$ 1, GLT $\gamma$ 2b, GLT $\gamma$ 2a and GLT $\alpha$ ) in *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> splenic B cells cultured for 96h as in (C). Expression is normalized to Ig $\beta$  and is presented relative to expression in *Fam72a*<sup>+/+</sup> B cells, set as 1. Mean of 3 independent experiments and the SEM were calculated following the rules for error propagation while calculating a ratio. Statistical analysis was performed using two-tailed Student's t test (\*p < 0,05; \*\*p<0,005).



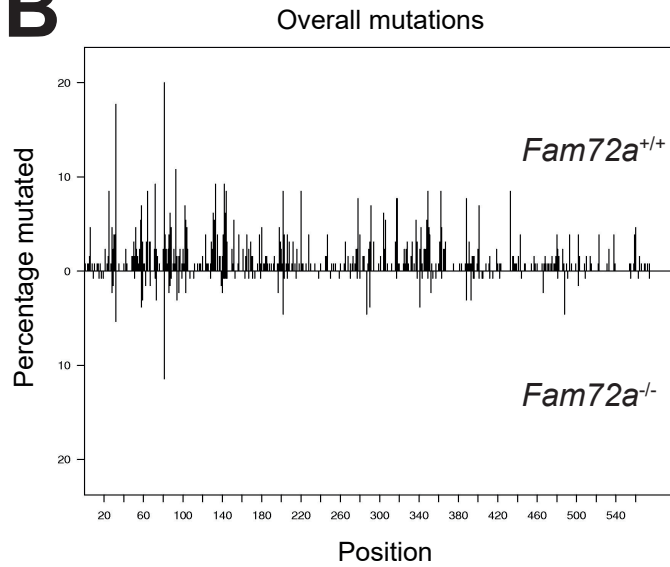
**Figure 3. The CSR defect in *Fam72a*<sup>-/-</sup> B cells is due to the specific up-regulation of *Ung2*.** (A) Flow cytometry analysis of IgA expression in *Fam72a*<sup>-/-</sup> CH12 cells transduced with an empty retrovirus (pMX) or expressing FAM72A, FAM72A<sup>W125A</sup> or FAM72A<sup>W125R</sup> and cultured for 3 days with TGF- $\beta$ , anti-CD40 antibody and IL-4. Plots are gated on EGFP expression. The percentage of IgA-expressing cells is indicated. (B) Flow cytometry analysis of IgA expression in *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> CH12 cells expressing (or not) an UNG inhibitor (Ugi) and cultured for 3 days with TGF- $\beta$ , anti-CD40 antibody and IL-4. The parental cell line (pCH12) was included as a positive control. Representative plots are shown. The percentage of IgA-expressing cells is indicated. (C) Plot showing the percentage of IgA-expressing cells from 3 independent experiments. p-value was determined using two-tailed Student's t-test; \*\*\*p<0.0005. (D) Flow cytometry analysis of IgA expression in *Fam72a*<sup>-/-</sup>, *Ung1*<sup>-/-</sup>, *Ung2*<sup>-/-</sup>, *Ung*<sup>-/-</sup>, *Fam72a*<sup>-/-</sup> *Ung1*<sup>-/-</sup> and *Fam72a*<sup>-/-</sup> *Ung2*<sup>-/-</sup> CH12 cells cultured for 72h with TGF- $\beta$ , IL-4 and anti-CD40 antibody. The percentage of IgA-expressing cells is indicated. (E) Western blot analysis for UNG (UNG1 and UNG2), AID and  $\beta$ -Actin in *Fam72a*<sup>-/-</sup>, *Ung1*<sup>-/-</sup>, *Ung2*<sup>-/-</sup>, *Ung*<sup>-/-</sup>, *Fam72a*<sup>-/-</sup> *Ung1*<sup>-/-</sup>, *Fam72a*<sup>-/-</sup> *Ung2*<sup>-/-</sup> and *Fam72a*<sup>-/-</sup> CH12 cells transduced with a retrovirus expressing FAM72A or FAM72A<sup>W125R</sup> and cultured for 72h with TGF- $\beta$ , IL-4 and anti-CD40 antibody (CIT). (F) Western blot analysis for UNG (UNG1 and UNG2), NBS1 and Histone H3 on nuclear and chromatin fractions prepared from CH12 cells (pCH12) and *Fam72a*<sup>-/-</sup> B cells expressing FAM72A or FAM72A<sup>W125R</sup>. (G) Western blot analysis for UNG (UNG1 and UNG2) and  $\beta$ -Actin protein expression levels in *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> splenic B cells cultured with LPS, IL-4 and anti-IgD-Dextran for 72h. (H) Flow cytometry analysis of IgA expression in CH12 cells transduced with an empty retrovirus or expressing UNG2. Representative plots of three experiments are shown. The percentage of IgA-expressing cells is indicated.



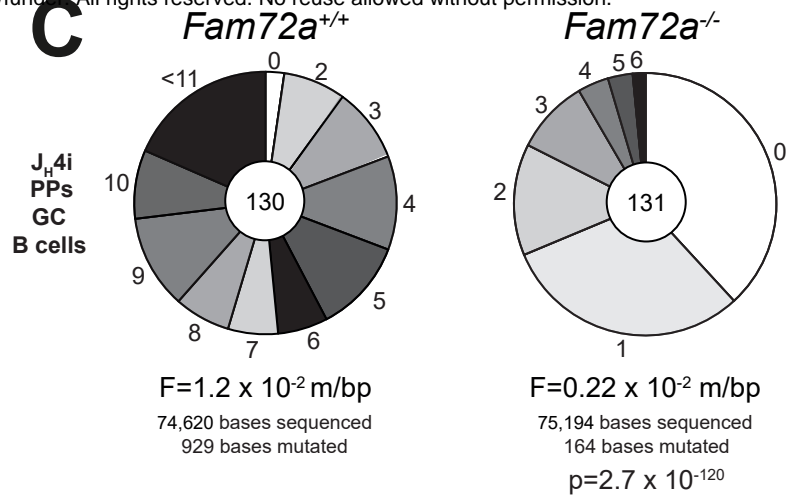
**A**



**B**



**C**



to	G	C	A	T	Total
G		97	81	35	213
C	22		22	51	95
A	192	67		108	367
T	43	128	83		254
					929

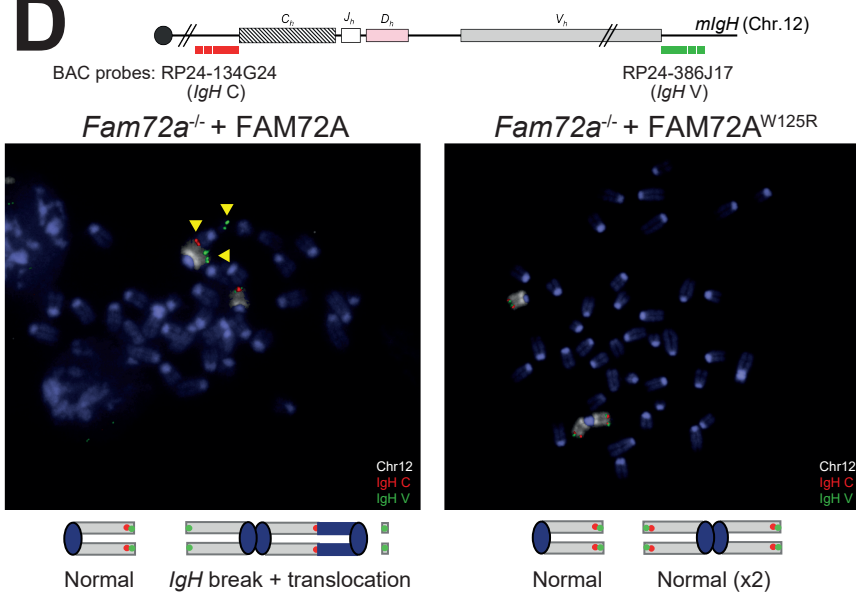
to	G	C	A	T	%
G		10,4	8,7	3,8	22,9
C	2,4		2,4	5,5	10,2
A	20,7	7,2		11,6	39,5
T	4,6	13,8	8,9		27,3
					100

to	G	C	A	T	Total
G		17	6	14	37
C	0		10	2	12
A	37	20		20	77
T	2	11	25		38
					164

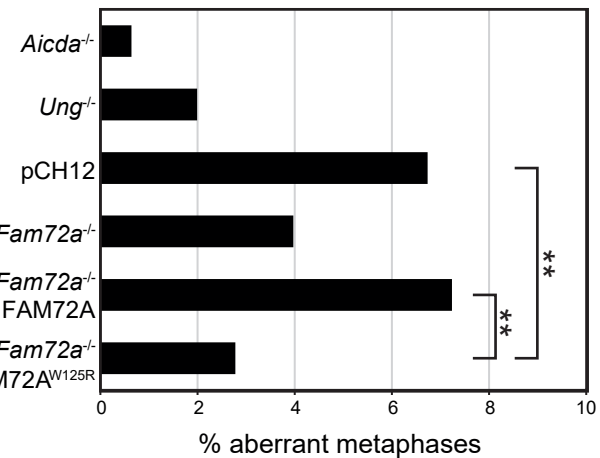
to	G	C	A	T	%
G		10,4	3,7	8,5	22,6
C	0,0		6,1	1,2	7,3
A	22,6	12,2		12,2	47,0
T	1,2	6,7	15,2		23,2
					100

	<i>Fam72a</i> <sup>+/+</sup>	<i>Fam72a</i> <sup>-/-</sup>
Mutations at dC/dG:Mutations at dA/dT	33:67	30:70
Transitions:Transversions at dC/dG	57:43	16:84
Transitions:Transversions at dA/dT	48:52	58:42

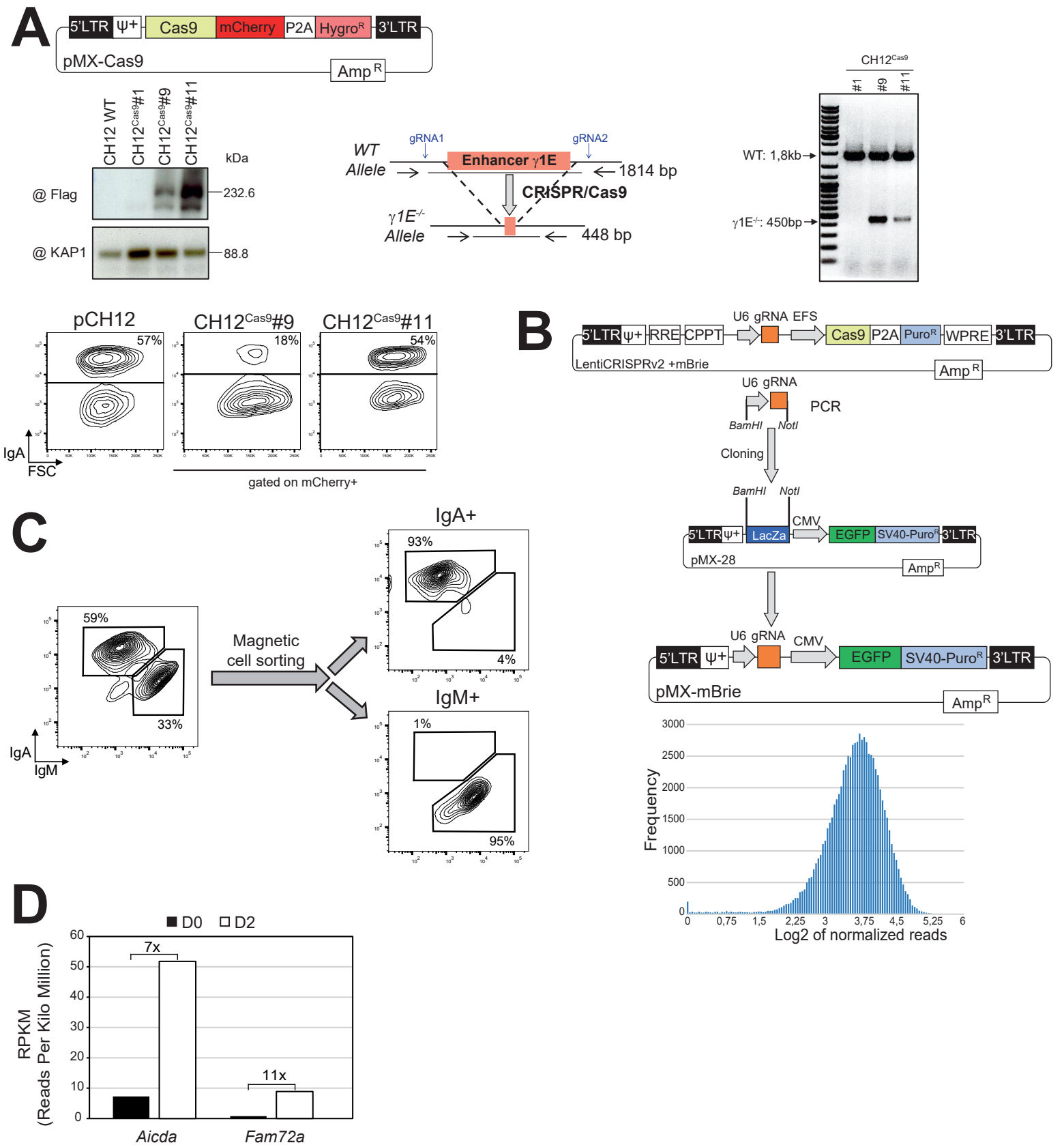
**D**



**E**

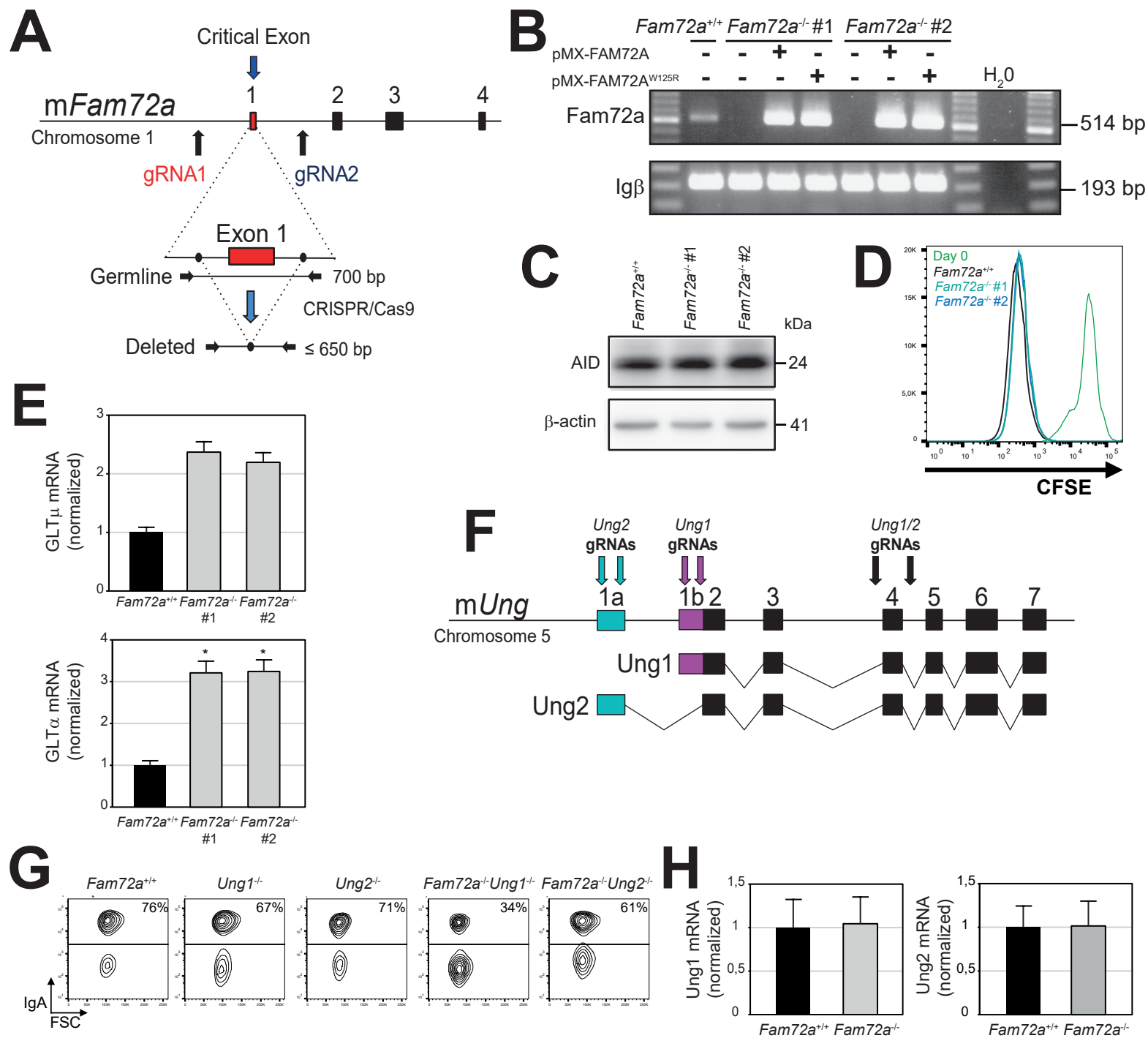


**Figure 4. Deficiency in *Fam72a* shifts the balance toward error-free DNA repair.** (A) Flow cytometry analysis of germinal center B cells (B220+ Fas+ GL-7+) isolated from the Peyer's patches of unimmunized *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> mice. Plots are gated on B220+ cells. (B) Distribution of mutations at the J<sub>H</sub>4i intron (J<sub>H</sub>4i) in *Fam72a*<sup>+/+</sup> (top) and *Fam72a*<sup>-/-</sup> (bottom) sequences obtained from germinal center B cells (B220+ Fas+ GL-7+) isolated from the Peyer's patches (PPs) of unimmunized *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> mice. (C) Pie charts depict the proportion of J<sub>H</sub>4i sequences with the indicated number of mutations. The total number of sequences analyzed is indicated in the center and the mutation frequency below. Tables depicting the mutation profiles are shown below. Statistical significance was determined using the SHMtool server. (D) Upper panel: Schematic representation of the *Igh* locus with positions of the BACs used for generation of DNA FISH probes. Lower panel: DNA-FISH on representative metaphases from day 2 stimulated *Fam72a*<sup>-/-</sup> cells complemented with FAM72A or FAM72A<sup>W125R</sup> using the *Igh* C BAC probe (red) combined with *Igh* V BAC probe (green) and chromosome 12 paint (white). Yellow arrowheads point to broken or translocated chromosome 12. (E) Percentage of aberrant metaphases from day 2 stimulated cells of the indicated genotype harboring *Igh* locus aberrations. See also Table S1. \*\* p<0.01.

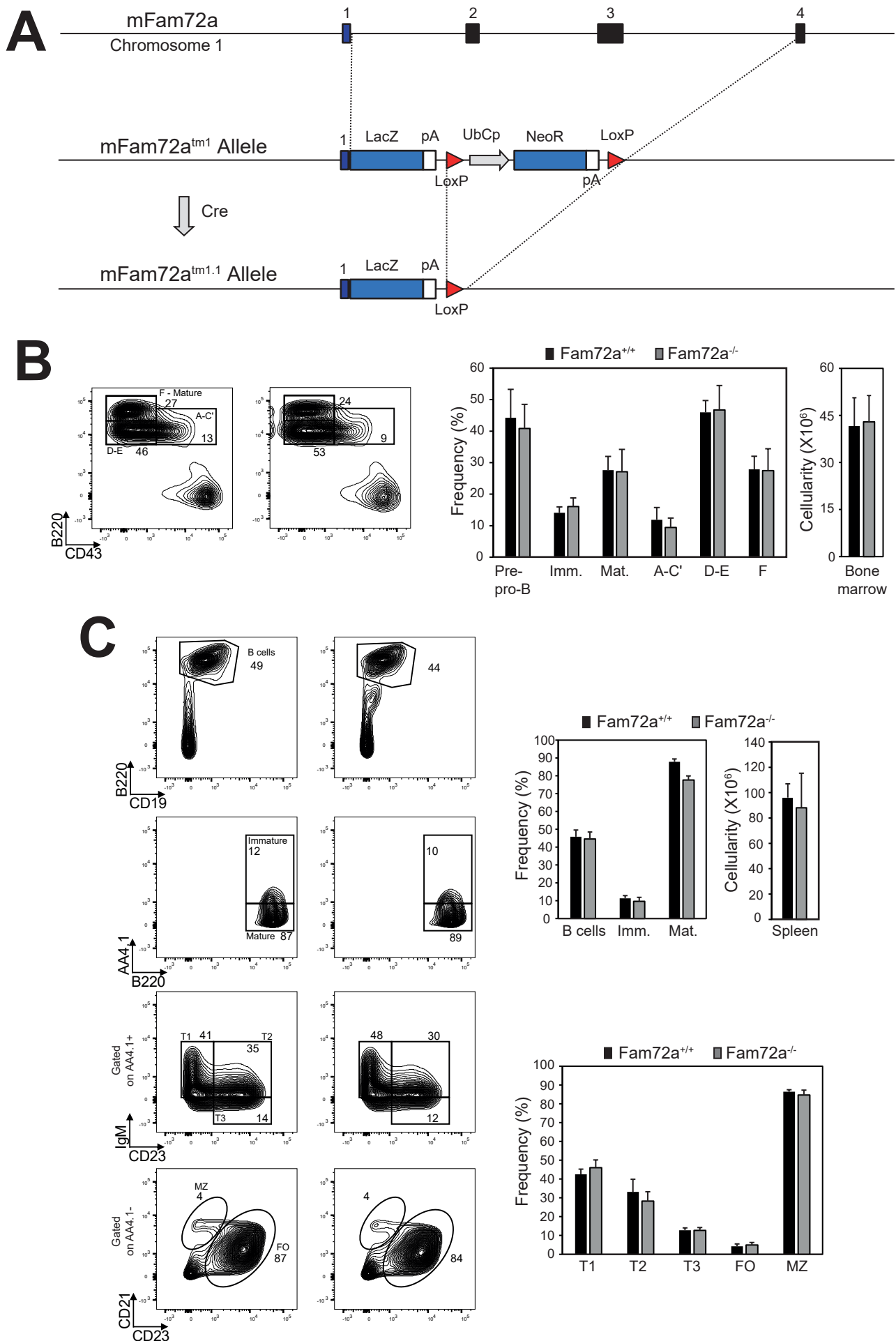


**Figure S1.** (A) Generation and validation of Cas9-expressing (CH12Cas9) cells. Clones were verified by Western blot for Cas9 expression, by PCR for Cas9-induced deletion of the enhancer  $\gamma 1$  and by flow cytometry for CSR. Clone #11 was chosen for the screen. (B) The U6-gRNA sequence was amplified by PCR from the mBrie library and subcloned into pMX-28 using BamHI and NotI restriction enzymes to generate the pMX-mBrie gRNA library. gRNA representation was analyzed by high-throughput sequencing. (C) Purity of IgM<sup>+</sup> and IgA<sup>+</sup> sorted populations was verified by flow cytometry. (D) Number of reads at the *Aicda* and *Fam72a* genes in wildtype splenic B cells before and after 48h in culture with LPS and IL-4, as determined by mRNA-Seq.

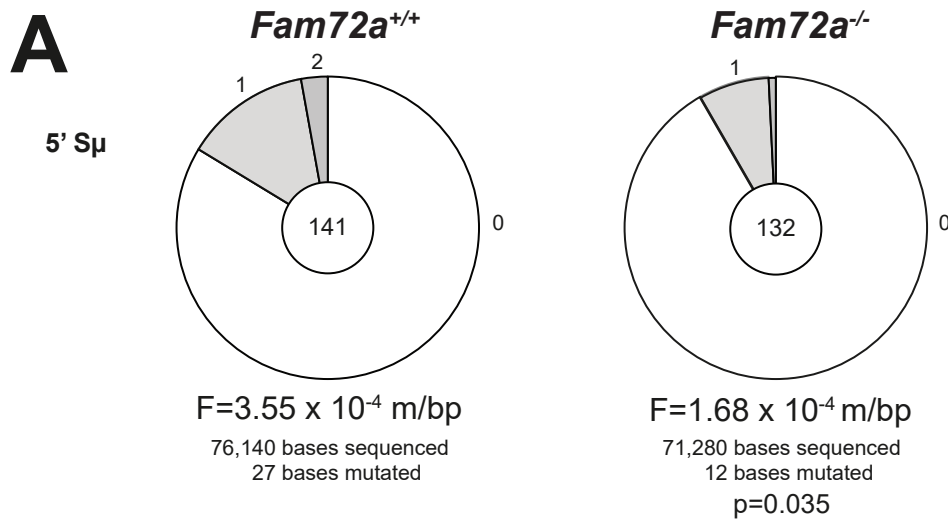




**Figure S2.** (A) Schematic representation of the murine *Fam72a* locus and location of the gRNAs used to generate *Fam72a*<sup>-</sup> clones using CRISPR/Cas9-HF1. (B) RT-PCR analysis for *Fam72a* and Igβ in *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-</sup> CH12 cells transduced (or not) with a retrovirus expressing FAM72A, FAM72A<sup>W125R</sup> or FAM72A<sup>W125A</sup>. (C) Western blot analysis for AID and β-actin in *Fam72a*<sup>+/+</sup> and two independent *Fam72a*<sup>-</sup> CH12 cell clones cultured for 3 days with TGF-β, IL-4 and anti-CD40 antibody. (D) CFSE dye-dilution analysis by flow cytometry in *Fam72a*<sup>+/+</sup> and two independent *Fam72a*<sup>-</sup> CH12 cell clones cultured for 3 days with TGF-β, IL-4 and anti-CD40 antibody. (E) RT-qPCR analysis for GLTμ and GLTα in *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-</sup> CH12 cells cultured for 3 days with TGF-β, IL-4 and anti-CD40 antibody. Triplicates were normalized to the abundance of Igβ and are expressed relative to *Fam72a*<sup>+/+</sup> controls. Statistical significance was determined by a two-tailed Student's t-test (\*p-value<0,05). Data are representative of 3 experiments. (F) Schematic representation of the murine *Ung* locus and location of the gRNAs targeting *Ung2* (exon 1a; blue), *Ung1* (exon 1b; purple) or *Ung* (exon4; black) used to generate *Ung1*<sup>-</sup>, *Ung2*<sup>-</sup>, *Ung*<sup>-</sup>, *Fam72a*<sup>-</sup> *Ung1*<sup>-</sup> and *Fam72a*<sup>-</sup> *Ung2*<sup>-</sup> CH12 cell clones using CRISPR/Cas9-HF1. (G) Flow cytometry analysis of IgA expression in additional *Ung1*<sup>-</sup>, *Ung2*<sup>-</sup>, *Fam72a*<sup>-</sup> *Ung1*<sup>-</sup> and *Fam72a*<sup>-</sup> *Ung2*<sup>-</sup> CH12 cell clones cultured for 3 days with TGF-β, IL-4 and anti-CD40 antibody. (H) RT-qPCR analysis for *Ung1*, *Ung2* and Igβ in *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-</sup> splenic B cells cultured for 4 days with LPS, IL-4 and anti-IgD-Dextran. Triplicates were normalized to the abundance of Igβ and are expressed relative to *Fam72a*<sup>+/+</sup> controls.



**Figure S3. (A)** Schematic representation of the murine *Fam72a* locus and strategy for the generation of *Fam72a*<sup>-/-</sup> mouse model. Flow cytometry and cellular analysis of B cell populations in the bone marrow (**B**) or in the spleen (**C**) from *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> mice, using the indicated cell surface markers. The data are representative of 5 mice per genotype.



**B**

ALL MUTATIONS	WT	<i>Fam72a</i> <sup>-/-</sup>	SITES	bp WT	bp <i>Fam72a</i> <sup>-/-</sup>	Frequency WT	Frequency <i>Fam72a</i> <sup>-/-</sup>	p value
G->C	97	17	157	20410	20567	0.004752572	0.0008265668	9.308466e-14
G->A	81	6	157	20410	20567	0.003968643	0.0002917295	1.488416e-15
G->T	35	14	157	20410	20567	0.001714846	0.0006807021	0.00390532
sum:G	213	37	157	20410	20567	0.01043606	0.001798998	6.205097e-29
C->G	22	0	91	11830	11921	0.001859679	0	6.882109e-06
C->A	22	10	91	11830	11921	0.001859679	0.0008388558	0.04911894
C->T	51	2	91	11830	11921	0.004311074	0.0001677712	3.387735e-11
sum:C	95	12	91	11830	11921	0.008030431	0.001006627	1.405943e-15
A->G	192	37	152	19760	19912	0.009716599	0.001858176	1.020279e-24
A->C	67	20	152	19760	19912	0.003390688	0.001004419	6.59427e-07
A->T	108	20	152	19760	19912	0.005465587	0.001004419	9.505953e-15
sum:A	367	77	152	19760	19912	0.01857287	0.003867015	9.102467e-44
T->G	43	2	174	22620	22794	0.001900973	8.774239e-05	2.078337e-09
T->C	128	11	174	22620	22794	0.005658709	0.0004825831	4.187917e-23
T->A	83	25	174	22620	22794	0.003669319	0.00109678	3.179272e-08
sum:T	254	38	174	22620	22794	0.011229	0.001667105	6.862086e-37
sum:GC	308	49	248	32240	32488	0.00955335	0.001508249	4.120884e-43
sum:AT	621	115	326	42380	42706	0.01465314	0.00269283	7.535157e-79
sum:ALL	929	164	574	74620	75194	0.01244975	0.002181025	2.703369e-120
Tv:GC	176	41	248	32240	32488	0.005459057	0.001262004	4.800845e-20
Tr:GC	132	8	248	32240	32488	0.004094293	0.0002462448	1.432898e-25
Tv:AT	301	67	326	42380	42706	0.007102407	0.001568866	1.763476e-34
Tr:AT	320	48	326	42380	42706	0.007550731	0.001123964	5.861283e-46
Tv:ALL	477	108	574	74620	75194	0.006392388	0.001436285	4.281994e-53
Tr:ALL	452	56	574	74620	75194	0.006057357	0.0007447403	1.178364e-69

**Figure S4. (A)** S $\mu$  mutation analysis in *Fam72a*<sup>+/+</sup> and *Fam72a*<sup>-/-</sup> CH12 cells cultured for 3 days with TGF- $\beta$ , IL-4 and anti-CD40 antibody. Pie charts depict the proportion of S $\mu$  sequences with the indicated number of mutations. The mutation frequency per base pair sequenced is shown below. The total number of sequences analyzed is indicated in the center. Statistical significance was determined with the Student's t test. **(B)** Mutation analysis at J $\mu$ 4i sequences performed with the SHMTool server.

**Table S1. IgH FISH analysis.**

Genotype	Expmt #	Metaphases with <i>Igh</i> breaks	Metaphases with <i>Igh</i> translocations	Total aberrant metaphases	Total normal metaphases	Total metaphases analyzed	% abnormal metaphases
<i>Aicda</i> <sup>-/-</sup>	1	2	0	2	248	250	0,8
	2	1	0	1	219	220	0,5
	<b>TOTAL</b>	<b>3</b>	<b>0</b>	<b>3</b>	<b>467</b>	<b>470</b>	<b>0,6</b>
<i>Ung</i> <sup>-/-</sup>	1	4	0	4	235	239	1,7
	2	4	1	5	209	213	2,3
	<b>TOTAL</b>	<b>8</b>	<b>1</b>	<b>9</b>	<b>444</b>	<b>452</b>	<b>2,0</b>
pCH12	1	11	6	16	233	249	6,4
	2	17	2	19	252	271	7,0
	<b>TOTAL</b>	<b>28</b>	<b>8</b>	<b>35</b>	<b>485</b>	<b>520</b>	<b>6,7</b>
<i>Fam72a</i> <sup>-/-</sup>	1	6	4	10	242	252	4,0
	<b>TOTAL</b>	<b>6</b>	<b>4</b>	<b>10</b>	<b>242</b>	<b>252</b>	<b>4,0</b>
<i>Fam72a</i> <sup>-/-</sup> +FAM72A	1	10	3	13	241	254	5,1
	2	15	0	15	118	133	11,3
	<b>TOTAL</b>	<b>25</b>	<b>3</b>	<b>28</b>	<b>359</b>	<b>387</b>	<b>7,2</b>
<i>Fam72a</i> <sup>-/-</sup> +FAM72A <sup>W125R</sup>	1	3	0	3	157	160	1,9
	2	7	1	8	228	236	3,4
	<b>TOTAL</b>	<b>10</b>	<b>1</b>	<b>11</b>	<b>385</b>	<b>396</b>	<b>2,8</b>

**Statistical analysis (Fisher exact test)**

			p value	Level
<i>Fam72a</i> <sup>+/+</sup>	vs	<i>Fam72a</i> <sup>-/-</sup>	0.14188	ns
<i>Fam72a</i> <sup>+/+</sup>	vs	<i>Fam72a</i> <sup>-/-</sup> + FAM72A	0.79264	ns
<i>Fam72a</i> <sup>+/+</sup>	vs	<i>Fam72a</i> <sup>-/-</sup> + FAM72A <sup>W125R</sup>	0.00873	**
<i>Fam72a</i> <sup>-/-</sup>	vs	<i>Fam72a</i> <sup>-/-</sup> + FAM72A	0.12234	ns
<i>Fam72a</i> <sup>-/-</sup>	vs	<i>Fam72a</i> <sup>-/-</sup> + FAM72A <sup>W125R</sup>	0.49575	ns
<i>Fam72a</i> <sup>-/-</sup> + FAM72A	vs	<i>Fam72a</i> <sup>-/-</sup> + FAM72A <sup>W125R</sup>	0.00481	**

**Table S2. Primers and gRNAs.**

Primers	Orientation	Sequence	UPL
<b>Transcripts</b>			
Ig $\beta$	Fwd	TGGTGCTGTCTTCCATGC	18
	Rev	TTGCTGGTACCGGCTCAC	
I $\mu$ -C $\mu$	Fwd	CCCAGACCTGGGAATGTATG	29
	Rev	GGAAGACATTTGGGAAGGACT	
I $\mu$ -C $\mu$ 2	Fwd	ACCTGGGAATGTATGGTTGTGGCTT	
	Rev	TCTGAACCTTCAAGGATGCTCTTG	
I $\gamma$ 3-C $\gamma$ 3	Fwd	GCAGAAATCTGCAGGACTAACA	71
	Rev	ACCGAGGATCCAGATGTGTC	
I $\gamma$ 2b-C $\gamma$ 2b	Fwd	TGGGCCTTCCAGACCTAAT	88
	Rev	GGGCTGATCTGTCAACTCCT	
I $\gamma$ 2a-C $\gamma$ 2a	Fwd	CAGCCTGGGATCAAGCAG	109
	Rev	TGGGGCTGTTGTTTTGGT	
I $\gamma$ 1-C $\gamma$ 1	Fwd	GGCCCTTCCAGATCTTTGAG	
	Rev	ATGGAGTTAGTTTGGGCAGCA	
I $\alpha$ -C $\alpha$	Fwd	GGAGACTCCCAGGCTAGACA	27
	Rev	CGGAAGGGAAGTAATCGTGA	
Ung1	Fwd	CTGCTCGGCTGGACCAT	
	Rev	GCGCCAACCGCAAAGAC	
Ung2	Fwd	AGTGGCGGCCGAGATC	
	Rev	CCACCCGGCCTTCTTG	
<b>Sub-cloning mBrie library</b>			
5'-BamHI site	Fwd	CGAATGCATCTAGATATCGGATCCCTTCCCATGATTCCTTCATATTTGC	
3'-NotI site	Rev	GCAGGCCTCTGCAGTCGACGGGCCCGCGCCGCAAAAAAGCACCGACTC GGTGCCAC	
<b>High-throughput sequencing</b>			
Stagger	P5 ARGON #1	AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTT CCGATCTTTGTGGAAAGGACGAAAC*A*C*C*G	
	P5 ARGON #2	AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTT CCGATCTCTTGTGGAAAGGACGAAAC*A*C*C*G	
	P5 ARGON #3	AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTT CCGATCTGCTTGTGGAAAGGACGAAAC*A*C*C*G	
	P5 ARGON #4	AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTT CCGATCTAGCTTGTGGAAAGGACGAAAC*A*C*C*G	
	P5 ARGON #5	AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTT CCGATCTCAACTTGTGGAAAGGACGAAAC*A*C*C*G	
	P5 ARGON #6	AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTT CCGATCTTGACCTTGTGGAAAGGACGAAAC*A*C*C*G	
	P5 ARGON #7	AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTT CCGATCTACGCAACTTGTGGAAAGGACGAAAC*A*C*C*G	
	P5 ARGON #8	AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGACGCTCTT CCGATCTGAAGACCTTGTGGAAAGGACGAAAC*A*C*C*G	
Index	1	CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGT GTGCTCTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
	2	CAAGCAGAAGACGGCATAACGAGATACATCGGTGACTGGAGTTCAGACGT GTGCTCTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	

	3	CAAGCAGAAGACGGCATAACGAGATGCCTAAGTACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
	4	CAAGCAGAAGACGGCATAACGAGATTGGTCAGTACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
	5	CAAGCAGAAGACGGCATAACGAGATCACTGTGTACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
	6	CAAGCAGAAGACGGCATAACGAGATATTGGCGTACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
	7	CAAGCAGAAGACGGCATAACGAGATGATCTGGTACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
	8	CAAGCAGAAGACGGCATAACGAGATTCAAGTGTACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
<b>Genotyping of CH12 KO clones</b>			
<i>Fam72a</i> <sup>-/-</sup>	Fwd	CGGGGTCGGAGCTGAAGAGTAGAAG	
	Rev	TCAACATAAGGGATGCACAGACACATG	
<i>Ung1</i> <sup>-/-</sup>	Fwd	CAATCAGAAGCGGCAAGGGGCGGGAGGTG	
	Rev	CTTCCCGAACTCCCCGCACAGC	
<i>Ung2</i> <sup>-/-</sup>	Fwd	CTCAGCCCCCTCCACACATG	
	Rev	CTAAGGTGTCCCAGTCTGTCCG	
<i>Ung</i> <sup>-/-</sup>	Fwd	GAGGCCGGATGTGGGGTGGGTGAGAC	
	Rev	TGAAACAGGAGCGGCCAAGGATAACAG	
<b>Sμ mutations</b>			
Sμ	Fwd	CGAATGCATCTAGATATCGGATCCCAATGGATACCTCAGTGGTTTTTAATG GTGGGTTTA	
	Rev	GCAGGCCTCTGCAGTCGACGGGCCAGCGGCCAGCTCATTCCAGTTCAT TACAG	
<b>SHM</b>			
VH588/FR3	Fwd	CGAATGCATCTAGATATCGGATCCCGCCTGACATCTGAGGACTCTGC	
J <sub>H</sub> 4 intron	Rev	GCAGGCCTCTGCAGTCGACGGGCCGACTTTTGCAGGCTCCACCAGACC	
<b>gRNAs</b>			
<i>Fam72a</i>	gRNA 1 (5' - 3')	AGAGAGTCCTATTTTCCTGG	
	gRNA 2 (5' - 3')	TACCAAGTAAGTCTGACGGG	
<i>Ung1</i>	gRNA 1 (5' - 3')	CCGCCGCCCAAGACGCCCA	
	gRNA 2 (5' - 3')	TGCGGTTGGCGCGGAGAGCG	
<i>Ung2</i>	gRNA 1 (5' - 3')	TCTTCTCTCCGACCCCCACA	
	gRNA 2 (5' - 3')	GCCGGTCCCAGGGAGCGGAG	
<i>Ung</i>	gRNA 1 (5' - 3')	ACCTCCACCTACGTGCAGAG	
	gRNA 2 (5' - 3')	AGCACACTCTATAATCACAG	

**Table S1. Top relevant genes for CSR found enriched in the CRISPR/Cas9 screen**

<b>Gene</b>	<b>p-value</b>	<b>Rank</b>	<b>Active gRNAs</b>
Aicda	2,3947E-07	1	6
Batf	2,3947E-07	2	4
Pten	2,3947E-07	3	3
Foxo1	2,3947E-07	4	4
Irf4	2,3947E-07	5	4
Smad2	2,3947E-07	6	4
Fam72a	2,3947E-07	7	4
Plekha2	2,3947E-07	8	4
Msh6	2,3947E-07	9	4
Atxn7l3	2,3947E-07	10	4
Runx3	2,3947E-07	11	4
Msh2	2,3947E-07	12	4
Junb	2,3947E-07	13	4
Il4ra	2,3947E-07	14	3
Usp22	2,3947E-07	15	4
Smad4	2,3947E-07	16	4
Tgfbr1	2,3947E-07	17	4
Cbfb	2,3947E-07	18	4
Stat6	2,3947E-07	19	4
Exo1	2,3947E-07	20	4
Ercc6l2	7,184E-07	23	3
Mlh1	4,0709E-06	27	4
Il2rg	4,5499E-06	28	4
Huwe1	0,000014129	32	4
Ung	0,000016523	35	4
Kdm1a	0,000043822	39	3
Ep300	0,000075432	42	3
Supt4a	0,00022821	52	4
Pms2	0,00044038	62	3
Rcor1	0,00046002	64	3
Batf3	0,00054	65	3
Med23	0,00065973	69	3
Trim28	0,00073301	70	3
Men1	0,00086136	74	2
Usp34	0,00087334	75	3
Brd4	0,00098205	77	3
Smad3	0,0010256	80	3
Xrcc5	0,0011478	81	3
Tgfbr2	0,0011789	83	2
Jak3	0,0021392	103	2
Cd40	0,0031076	121	3
Zmynd8	0,0058164	165	3
Supt6	0,0058813	166	2
Med4	0,0071723	197	2
Exosc8	0,0088576	225	2
Exosc10	0,011368	257	2
Polr2a	0,013108	284	2
Runx1	0,013817	300	2
Trp53bp1	0,015733	336	2

**Table S2. Primers used in this study**

<b>Primers</b>	<b>Orientation</b>	<b>Sequence</b>	<b>UPL References</b>
<b>Transcripts</b>			
I $\gamma$ $\beta$	Fwd	TGGTGCTGTCTTCCATGC	18
	Rev	TTGCTGGTACCGGCTCAC	
GLT $\mu$	Fwd	CCCAGACCTGGGAATGTATG	29
	Rev	GGAAGACATTTGGGAAGGACT	
GLT $\alpha$	Fwd	GGAGACTCCCAGGCTAGACA	27
	Rev	CGGAAGGGAAGTAATCGTGA	
<b>Sub-cloning mBrie library</b>			
5'-BamHI site	Fwd	CGAATGCATCTAGATATCGGATCCCTTTCCCATGATTCCTTC ATATTTGC	
3'-NotI site	Rev	GCAGGCCTCTGCAGTCGACGGGCCCGCGGCCGCAAAAAA GCACCGACTCGGTGCCAC	
<b>Genotyping of CH12 KO clones</b>			
Fam72a <sup>-/-</sup>	Fwd	CGGGGTCGGAGCTGAAGAGTAGAAG	
	Rev	TCAACATAAGGGATGCACAGACACATG	
Ung1 <sup>-/-</sup>	Fwd	CAATCAGAAGCGGCAAGGGGCGGGAGGTG	
	Rev	CTTCCCGAACTCCCCGCACAGC	
Ung2 <sup>-/-</sup>	Fwd	CTCAGCCCCTCCCACACATG	
	Rev	CTAAGGTGTCCCAGTCTGTCCG	
Ung <sup>-/-</sup>	Fwd	GAGGCCGGATGTGGGGTGGGTGAGAC	
	Rev	TGAAACAGGAGCGGCCAAGGATAACAG	
<b>S<math>\mu</math> mutations</b>			
S $\mu$	Fwd	CGAATGCATCTAGATATCGGATCCCAATGGATACCTCAGTG GTTTTTAATGGTGGGTTTA	
	Rev	GCAGGCCTCTGCAGTCGACGGGCCCAGCGGCCCAGCTCA TTCCAGTTCATTACAG	
<b>SHM</b>			
VH588/FR3	Fwd	CGAATGCATCTAGATATCGGATCCCGCCTGACATCTGAGG ACTCTGC	
J $\mu$ 4 intron	Rev	GCAGGCCTCTGCAGTCGACGGGCCCGACTTTTGCAGGCTC CACCAGACC	



**Table S3. Antibodies used in this study**

<b>Antibody</b>	<b>Clone/Name</b>	<b>Origin</b>	<b>Usage</b>	<b>References</b>
Anti FLAG	A8592	Sigma	Western	
Anti KAP1	1TB1A9	IGBMC	Western	
Anti IgA-Biotin	1040-08	Southern Biotech	Magnetic cell sorting, Flow cytometry	
Anti UNG	6103	Gift from Bodil Kavli	Western	18
Anti AID	Strasbg9, AID-2E11	IGBMC	Western	
Anti b-actin	A1978	Sigma	Western	
Anti IgG1-Biotin	A 85-1	BD Pharmigen	Flow cytometry	
Anti IgG2a-Biotin	RMG2A	Biologend	Flow cytometry	
Anti IgG2b-Biotin	RMG2B1	Biologend	Flow cytometry	
Anti IgG3-Biotin	R40-82	BD Pharmigen	Flow cytometry	
Anti IgA-PE	1040-09	Southern Biotech	Flow cytometry	
Anti GL-7-Pacific Blue		BioLegend	Flow cytometry	
Anti B220-PE-Cy7	RA3-6B2	eBiosciences	Flow cytometry	
Anti Fas(CD95)-PE	Jo2	BD Pharmigen	Flow cytometry	

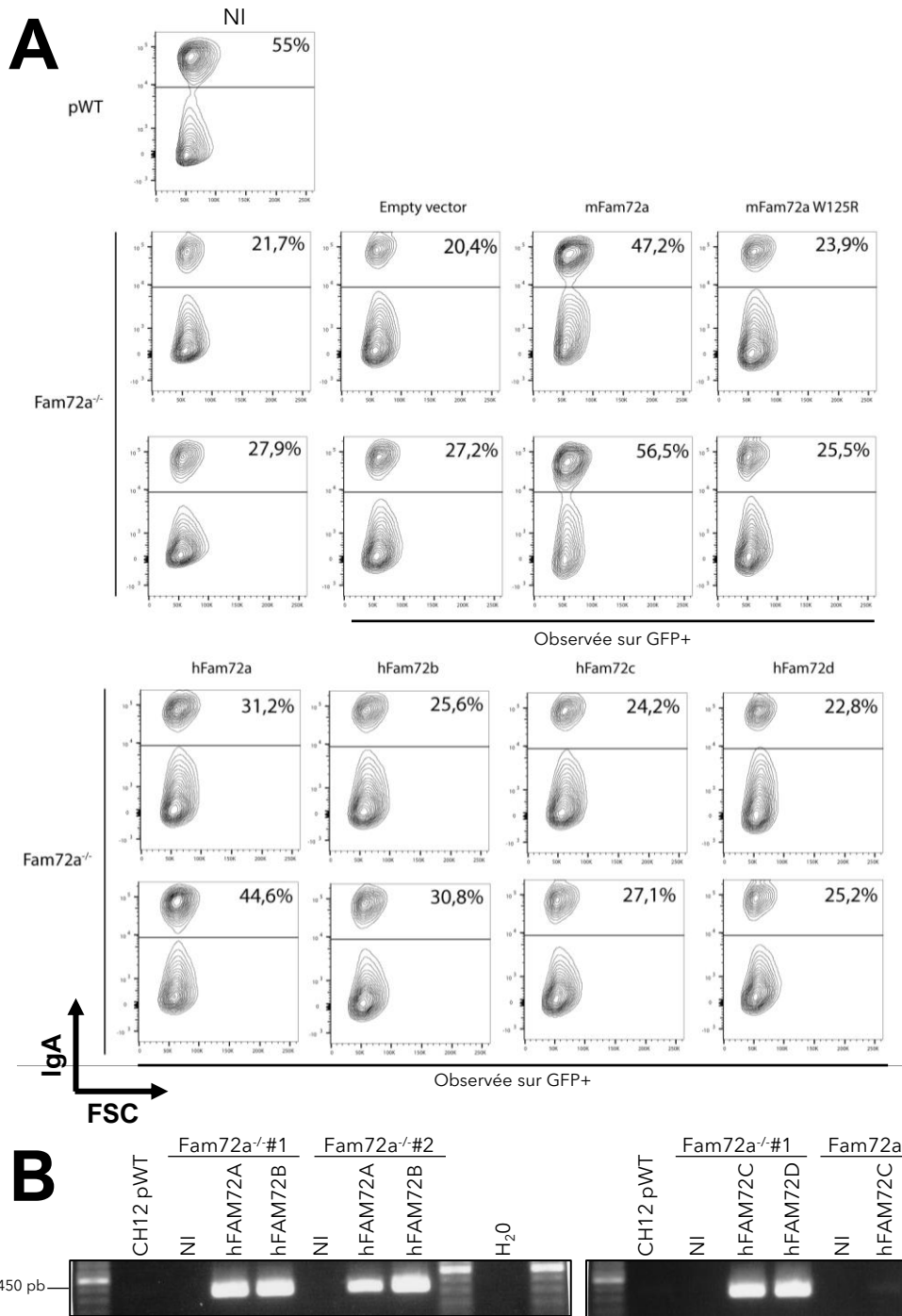
## II- Résultats supplémentaires :

### 1. Reconstitution de l'expression de mFam72a et mFam72<sup>W125R</sup> dans les cellules CH12 Fam72a<sup>-/-</sup>

Afin de montrer que le défaut de CI décrit dans les cellules CH12 Fam72a<sup>-/-</sup> est bien dû à la perte de Fam72a, nous avons tout d'abord réalisé des expériences de reconstitution avec des versions totales de mFam72a et mFam72a<sup>W125R</sup> comportant une étiquette Flag dans le but de voir l'expression de Fam72a par western blot. Je tiens à préciser que nous avons testé un anticorps commercial anti-Fam72a ainsi que trois autres anticorps anti-Fam72 synthétisés au service d'anticorps de l'IGBMC qui n'ont donné aucune spécificité pour Fam72a. Afin de réaliser ces expériences de reconstitution, nous avons donc infecté les cellules Fam72a<sup>-/-</sup> à l'aide de rétrovirus exprimant Flag-mFam72a ou Flag-mFam72a<sup>W125R</sup>. Ces cellules ont ensuite été sélectionnées puis stimulées pendant 72h avec du TGF- $\beta$ , de l'IL-4 et de l'anti-CD40 afin de réaliser le mécanisme de CI. L'efficacité d'infection et de CI a été analysée par cytométrie en flux en regardant le pourcentage de cellules GFP<sup>+</sup> et l'expression d'IgA à la surface des cellules. L'efficacité d'infection était d'environ 90% avec une absence de reconstitution pour la version totale Flag-mFam72a (27 à 30%) ainsi que Flag-mFam72a<sup>W125R</sup> (25 à 32%) comparée aux cellules Fam72a<sup>-/-</sup> non infectées (25 à 30%) et aux cellules WT non infectées et stimulées (62,5%) (données non montrées). Suite à ces résultats négatifs, nous avons analysé l'expression de mFam72a par western blot à l'aide d'un anticorps anti-Flag qui a révélé que les versions Flag-mFam72a et Flag-mFam72a<sup>W125R</sup> n'étaient pas exprimées dans les différents extraits cellulaires correspondants aux différentes conditions citées ci-dessus. Ces expériences de reconstitution ont été répétées avec Flag en C-ter des deux versions de mFam72a qui également ne sauvaient pas le phénotype et n'étaient pas exprimées sur western blot à l'aide d'un anticorps Flag et enfin, avec les versions mFam72a et mFam72a<sup>W125R</sup> sans tag (résultats Figure 2E du papier) où l'on a observé une reconstitution du phénotype de la CI avec la version mFam72a mais pas avec la version mutante mFam72a<sup>W125R</sup> comme décrit dans le papier. L'expression de mFam72a et mFam72a<sup>W125R</sup> a été confirmée par RT-PCR (Figure 2E du papier).

### 2. Reconstitution de l'expression des versions humaines hFAM72A-D dans les cellules CH12 Fam72a<sup>-/-</sup>

Fam72a est un gène très conservé au sein des espèces mais, de manière surprenante chez l'humain il existe 4 gènes paralogues hFAM72A, hFAM72B, hFAM72C et hFAM72D qui ont une identité de séquence avec mFam72a de 90%. L'alignement des séquences de ces gènes



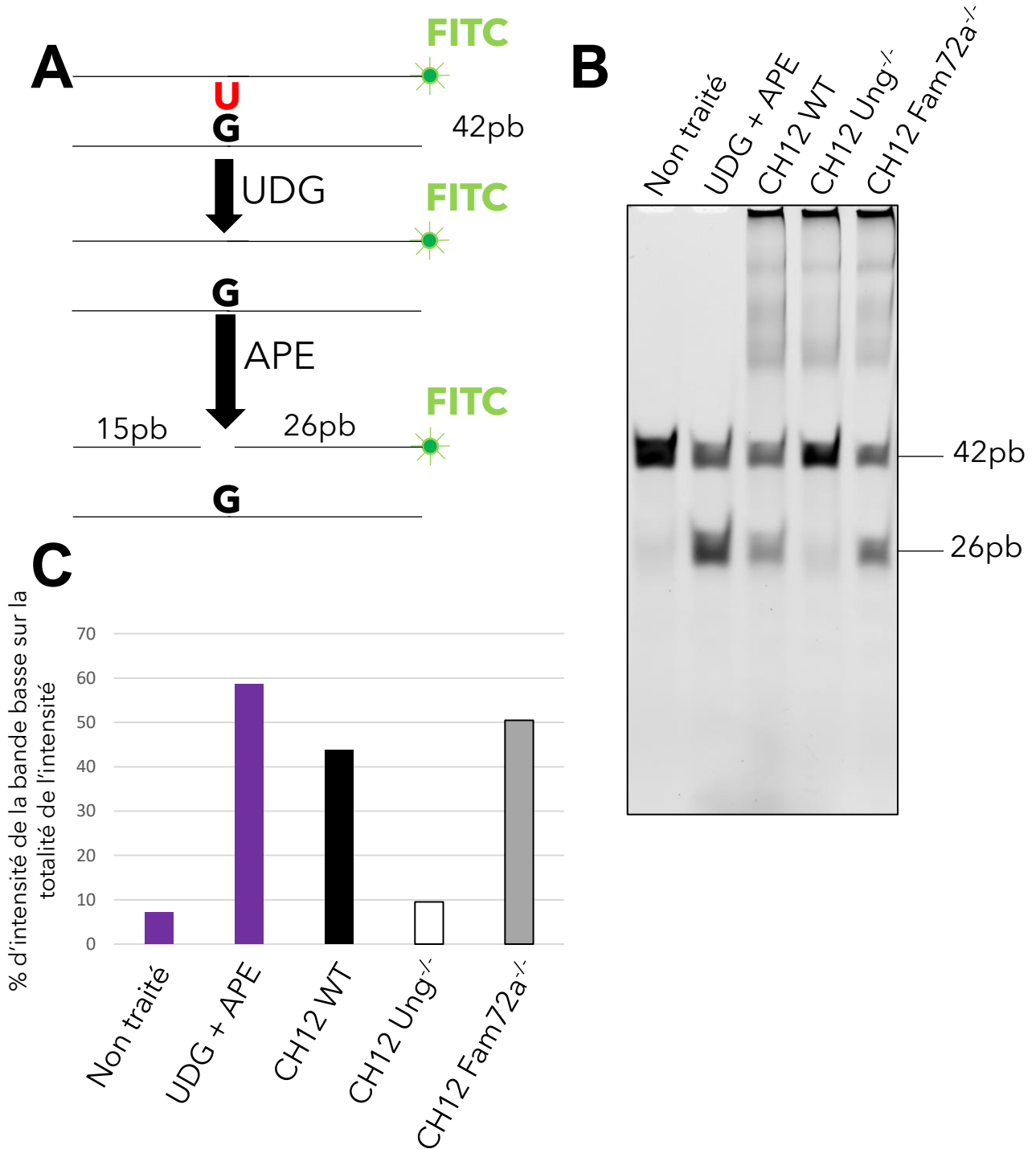
**Figure 27: Reconstitution des cellules *Fam72a*<sup>-/-</sup> avec les formes souris, mFam72a et mFam72a<sup>W125R</sup> et les formes humaines hFAM72A-D.** **A.** Les lignées clonales CH12 *Fam72a*<sup>-/-</sup> ont été infectées à l'aide de rétrovirus exprimant la version totale de mFam72a, le mutant mFam72a<sup>W125R</sup> et les versions humaines hFAM72A-D. Elles ont été ensuite sélectionnées et stimulées pendant 72h avec du TGFβ, de l'IL-4 et de l'anti-CD40 afin de réaliser la CI. Enfin, l'expression d'IgA à la surface des cellules a été analysée par cytométrie en flux. Des cellules CH12 pWT ainsi que les lignées clonales CH12 *Fam72a*<sup>-/-</sup> non infectées (NI) ont été ajoutées comme contrôle lors de la CI. Dans les lignées clonales CH12 *Fam72a*<sup>-/-</sup>, la première ligne correspond au clone KO#1 et la seconde ligne au clone KO#2. **B.** L'expression des différentes versions hFAM72A-D a été confirmée par RT-PCR. NI: cellules non infectées

paralogues montre que plusieurs acides aminés diffèrent. Il est intéressant de noter que la séquence des gènes hFAM72A et hFAM72B comporte le W125 tandis que le W125 est remplacé par la R125 dans la séquence des gènes hFAM72C et hFAM72D. Etant donné que le W125 est important pour l'association de Fam72a avec Ung2, nous avons essayé de sauver le phénotype de la CI avec les différentes versions hFAM72A-D afin de valider que seules les formes A et B comportant le W125 peuvent reconstituer le défaut de CI.

Comme ci-dessus nous avons donc infecté les deux lignées cellulaires clonales CH12 Fam72a<sup>-/-</sup> à l'aide de rétrovirus exprimant les différentes formes hFAM72A-D ainsi que les versions mFam72a et mFam72a<sup>W125R</sup> en contrôle. Ces cellules ont ensuite été sélectionnées puis stimulées pendant 72h avec du TGF-β, de l'IL-4 et de l'anti-CD40 afin de réaliser le mécanisme de CI. L'efficacité d'infection et de CI a été analysée par cytométrie en flux en regardant le pourcentage de cellules GFP+ et l'expression d'IgA à la surface des cellules. L'efficacité d'infection était d'environ 90% (données non montrées), et nous avons observé une légère reconstitution de l'efficacité de la CI avec la version hFAM72A (+10 et +17%), et une très légère reconstitution avec la version hFAM72B (+3 et +4%) qui ne sont pas vraiment significatives. Toutefois, nous n'observons pas de sauvetage du phénotype de la CI avec les formes hFAM72C et hFAM72D (**Figure 27A**). L'expression des différentes versions hFAM72A-D a été confirmée par RT-PCR (**Figure 27B**) où nous pouvons voir une expression faible de la forme hFAM72C dans la lignée cellulaire Fam72a<sup>-/-</sup> #2 qui est due à un problème d'extraction d'ARN. Nous en concluons qu'il y a une tendance non significative de reconstitution du phénotype de CI avec les versions humaines comportant le W125 (hFAM72A et hFAM72B), donnant un peu plus de poids dans l'importance du W125 dans le rôle de Fam72a.

### 3. L'activité d'Ung est-elle modifiée en l'absence de Fam72a ?

Sachant que l'action de Fam72a est dépendante de son interaction avec Ung2, nous avons décidé de regarder si l'activité d'Ung2 est modifiée en absence de Fam72a. Nous avons donc réalisé un test d'activité d'Ung *in vitro* ([Cortizas et al., 2016b](#); [Di Noia and Neuberger, 2002](#)) à l'aide d'un substrat d'ADNdb contenant un mésappariement interne dU :dG pour analyser l'activité de glycosylation d'Ung (**Figure 28A**) dans les extraits de cellules B CH12 WT, Ung<sup>-/-</sup> et Fam72a<sup>-/-</sup> qui ont été stimulées avec du TGF-β, de l'IL-4 et de l'anti-CD40. Des contrôles avec les oligonucléotides hybridés non traités et le substrat seul traité (UDG+APE) ont été ajoutés. Tandis que l'extrait de cellules Ung<sup>-/-</sup> affiche une activité de glycosylation résiduelle minimale (9%), certainement due à l'expression de Smug1, l'extrait de cellules CH12 WT montre une activité de



**Figure 28: Test d'activité d'Ung en absence de Fam72a.** **A.** Schéma explicatif du test d'activité d'UNG. Deux oligonucléotides sont hybridés afin de générer un mésappariement U:G (42 pb) qui peut être reconnu par UNG pour créer un site abasique pris en charge par APE qui va générer une cassure simple brin (26pb). L'oligonucléotide contenant l'uracile est couplé à la fluorescéine (FITC), ce qui permet de distinguer sur gel de polyacrylamide 15% TBE-Urée la quantité de mésappariements pris en charge par UNG et donc l'activité d'UNG. **B.** Oligonucléotides hybridés incubés avec différents extraits cellulaires (10 µg) des lignées CH12 indiquées au dessus digérées avec UDG et APE et des contrôles tels que les oligonucléotides hybridés non traités (Non traités) ou digérés (UDG+APE) qui ont été résolus sur un gel de polyacrylamide 15% TBE-Urée. **C.** Quantification du % pourcentage de glycosylation pour les différents extraits résolus sur le gel en (B) en regardant l'intensité de la bande basse (26 pb) comparée à la totalité de l'intensité.

glycosylation d'Ung abondante (43,6%). En revanche, l'extrait de cellules Fam72a<sup>-/-</sup> montrent une activité de glycosylation d'Ung plus importante (50,4%) comparée à l'extrait de cellules CH12 WT (**Figure 28B-C**). Cette expérience préliminaire montre une activité de glycosylation d'Ung plus importante en absence de Fam72a, cependant nous avons aussi observé que nous avons plus d'Ung2 total dans ces cellules. Avec ces résultats préliminaires, nous ne pouvons donc pas dire que l'absence de Fam72a améliore l'activité de glycosylation d'Ung.

### III- DISCUSSION

Lors des mécanismes de diversification des anticorps, les voies du BER et du MMR sont détournées de leur mode de réparation habituellement fidèle pour permettre l'apparition de mutations et de DSBs dans le cadre de l'HMS et de la CI respectivement ([Methot and Di Noia, 2017a](#)). Même si des acteurs clés de cette réparation mutagène ont été identifiés, les mécanismes dictant le choix entre une réparation fidèle ou non-fidèle restent aujourd'hui encore mal compris. Cette étude nous a permis d'identifier Fam72a comme un facteur faisant pencher la balance en faveur d'une réparation non-fidèle dans les loci Ig. Il est ainsi requis pour réaliser efficacement la CI et l'HMS. Nos résultats indiquent que l'interaction entre Fam72a et Ung2 est nécessaire pour cet effet, et que Fam72a semble réguler directement la quantité d'Ung2 dans les cellules B activées. En effet, l'expression de Fam72a semble être négativement corrélée à la quantité d'Ung2 dans les cellules B. De plus, la délétion de Fam72a entraîne un défaut important de la CI et une baisse de la fréquence de mutations dans les loci Ig. Ainsi, nous émettons l'hypothèse que Fam72a entraîne la dégradation d'Ung2 pour promouvoir une réparation mutagène requise pour réaliser efficacement l'HMS et la CI.

#### 1. Ung2 et la commutation isotypique : une affaire de quantité ?

##### 1.1 Une trop grande quantité d'Ung2 nuit-elle à la CI ?

Il peut paraître difficile au premier abord d'expliquer en quoi une baisse de la quantité d'Ung2 permet d'améliorer l'efficacité de CI. En effet, on pourrait émettre l'hypothèse qu'avoir plus d'Ung2 pourrait permettre de générer plus de sites abasiques, et donc potentiellement plus de cassures simple brin grâce à APE. Ces cassures simple brin plus nombreuses permettraient d'aboutir logiquement à l'apparition de plus de DSBs, dont la CI est dépendante. Cette hypothèse est à l'opposé de ce que nous observons dans nos expériences, où une surexpression d'Ung2 par transduction rétrovirale semble inhiber la CI dans les CH12 (Figure 3D du manuscrit). Il est cependant important de noter que ces résultats vont à l'encontre de ce qui a été observé dans d'autres études. En effet, il a été montré que la surexpression d'Ung2 WT ou de versions mutantes ayant perdu presque toute activité catalytique permet de restaurer la CI à des niveaux normaux dans des cellules B primaires Ung<sup>-/-</sup> ([Begum, 2004](#); [Di Noia et al., 2007a](#)).

L'hypothèse selon laquelle la transduction rétrovirale des cellules B primaires dans ces études ne permettrait pas d'exprimer assez fortement Ung2 pour observer l'inhibition de la CI que nous observons semble peu probable. En effet, Di Noia et al. ont montré que le niveau d'expression d'Ung2 suite à sa surexpression la rend détectable et très bien visible par WB, contrairement au

niveau d'expression trop faible des cellules Ung<sup>+/+</sup> pour qu'elle soit visible dans ce cas ([Di Noia et al., 2007a](#)). Ceci montre donc bien qu'il existe une grosse différence d'expression entre les cellules transduites et les cellules Ung<sup>+/+</sup>.

Enfin, il se peut que l'effet d'une diminution de l'efficacité de CI en cas de surexpression d'Ung2 soit spécifique à la lignée CH12, qui posséderait alors des particularités que n'ont pas les cellules B primaires pour expliquer cet effet. Nous prévoyons ainsi d'entraîner la surexpression d'Ung2 dans des cellules B primaires et d'observer quel effet cela aura sur l'efficacité de la CI vers différents isotypes.

## 1.2 Et une quantité trop faible ?

A l'inverse, une étude récente semble indiquer qu'une quantité trop faible d'Ung2 ne suffise pas à réaliser la CI efficacement ([Eldin et al., 2020](#)). En effet, une étude récente, montre que dans des cellules CH12 exprimant la protéine virale Vpr, la quantité totale est réduite de plus de 80% par rapport à des CH12 n'exprimant pas Vpr ([Eldin et al., 2020](#)). Or, après stimulation de ces cellules, une perte d'environ 50% de l'efficacité de CI peut être observée dans les cellules exprimant Vpr, ainsi qu'une augmentation de la quantité d'uraciles présents dans l'ADN ([Eldin et al., 2020](#)). La diminution de CI observée est de plus dépendante de la dose de virus utilisée pour infecter les cellules, et donc potentiellement dépendante de la quantité de Vpr exprimée. Cette étude confirme de plus ce même effet dans des cellules B primaires, ce qui exclut la possibilité d'un mécanisme propre aux CH12.

Contrairement à ce que semble supposer cette étude, nous n'observons pas de relation dose-effet entre une faible quantité d'Ung2 et une baisse d'efficacité de la CI. En effet, nous observons que la surexpression de Fam72a par transduction rétrovirale de cellules Fam72a<sup>-/-</sup> entraîne une plus forte dégradation d'Ung2, mais permet tout de même de restaurer complètement la déficience en CI (Figures 2E et 3B du manuscrit). Cela peut indiquer que soit Vpr a un rôle inhibiteur de la CI indépendamment du fait qu'elle dégrade Ung2, soit c'est Fam72a qui promeut une CI efficace indépendamment du fait qu'elle entraînerait la dégradation d'Ung2.

## 2. Fam72a et Ung2 dans la CI : d'autres régulations possibles

Si cette dernière hypothèse est exacte, il se pourrait que Fam72a agisse sur d'autres aspects d'Ung2, ou agisse même sur d'autres facteurs de la réparation pour promouvoir une CI efficace.

Parmi les autres actions qu'elle pourrait jouer sur Ung2, on peut par exemple penser au fait que Fam72a régule peut-être l'interaction entre Ung2 et Rev1, puisque cette dernière s'avère importante pour le ciblage d'Ung2 aux régions S et pour l'augmentation de son activité catalytique ([Zan et al., 2012](#)).



Il est aussi possible que Fam72a régule la présence et/ou l'activité d'Ung2 en fonction du cycle cellulaire. En effet, il a été montré qu'AID désamine les cytosines en uraciles en phase G1 ([Faili et al., 2002](#); [Wang et al., 2017](#)), et que ces uraciles sont également réparés lors de cette phase ([Le and Maizels, 2015](#); [Sharbeen et al., 2012](#)). Or, cette réparation en G1 est particulièrement importante pour aboutir à une réparation mutagène dans les régions V lors de l'HMS ([Sharbeen et al., 2012](#)), mais peut également l'être dans les régions S afin de promouvoir la réparation des cassures par NHEJ. L'activation des lymphocytes B permet d'augmenter à la fois l'expression et l'activité d'Ung2, autant chez l'humain ([Imai et al., 2003b](#)) que chez la souris ([Di Noia et al., 2006](#); [Doseth et al., 2011](#)), et fait qu'Ung2 est présent et actif tout au long du cycle cellulaire ([Schrader et al., 2007a](#)). Les mécanismes permettant cette régulation sont cependant mal connus.

Il a cependant été montré qu'une partie de la régulation de la version humaine d'UNG2 se fait par la présence de modifications post-traductionnelles telles que des phosphorylations, des ubiquitinations et de la sumoylation ([Hagen et al., 2008a](#); [Lu et al., 2004](#); [Ma et al., 2009](#)). Ainsi, il est légitime de se demander si Fam72a ne joue pas également un rôle dans la présence ou l'absence de ces modifications, pour permettre par exemple à Ung2 de persister durant la phase G1 du cycle cellulaire et d'opérer une réparation mutagène des uraciles.

Pour répondre à cette question, il peut être envisagé de stopper des cellules CH12 en G1 à l'aide de thymidine, et d'observer s'il existe des différences de quantité et/ou d'activité d'Ung2 entre des cellules Fam72a<sup>+/+</sup> et des cellules Fam72a<sup>-/-</sup>.

### 3. Et Ung1 dans tout ça ?

Les expériences que nous avons menées ont également permis de valider la récente observation qui montre qu'Ung1 est capable de soutenir des niveaux normaux de CI en absence d'Ung2. Cela soulève la question de savoir à quel point elle contribue au mécanisme de CI dans des conditions normales. L'ablation d'Ung1 n'entraîne pas de perte de CI dans les CH12, mais cela ne veut pas dire qu'elle ne joue pas de rôle dans le mécanisme normal de CI, puisqu'il est possible qu'Ung2 soit capable de compenser la perte d'Ung1.

Le domaine d'interaction de Fam72a dans Ung2 étant situé dans la partie N-terminale qui la différencie d'Ung1, il est peu probable que Fam72a régule l'activité d'Ung1. Quelle est la raison pour laquelle il est nécessaire que Fam72a soit exprimée et interagisse avec Ung2 pour obtenir une CI efficace, alors qu'Ung1 est capable seule d'initier la CI avec la même efficacité ? Il n'y a qu'une faible proportion d'Ung1 qui est localisée dans le noyau à un moment donné ([Sarno et al., 2019a](#)), et ceci pourrait être la raison pour laquelle Ung1 est capable de soutenir des niveaux normaux de CI. Cette idée va de pair avec le fait qu'une surexpression d'Ung1 entraîne également

une baisse de CI (Figure 3D du manuscrit). A nouveau, cela donne du poids au modèle qui énonce que c'est la quantité d'Ung1 ou d'Ung2 qui dicte l'efficacité de la CI. Ung1 et Ung2 auraient alors des fonctionnements similaires, et l'augmentation du niveau de l'un ou de l'autre aboutirait à la baisse du niveau de CI, probablement du fait d'une réparation fidèle accrue.

#### 4. Fam72a et Ung2 dans l'HMS

Conformément à ce que nous observons dans les cellules *Fam72a<sup>-/-</sup>*, il a été observé que la surexpression d'Ung2 entraîne une réduction relativement importante du nombre de transitions à G:C et des mutations à A:T, ainsi qu'une baisse plus faible des transversions en G:C ([Pérez-Durán et al., 2012](#); [Thientosapol et al., 2017](#)). Ces résultats permettent de conforter l'hypothèse selon laquelle une quantité accrue d'Ung2 fait pencher la balance en faveur d'une réparation fidèle. Ils montrent également que la réparation des uraciles par le MMR est affectée par la surexpression d'Ung2, puisqu'on observe une réduction du nombre de mutations à A:T. Cette réduction pourrait être le fait d'une inhibition de la réparation des uraciles par le MMR seul, mais pourrait également être due à une collaboration plus limitée entre BER et MMR, qui peuvent agir ensemble pour initier des mutations à A:T ([Frieder et al., 2009a](#); [Krijger et al., 2009](#); [Schanz et al., 2009](#)). Il reste alors à expliquer en quoi une quantité supplémentaire d'Ung2 empêche cette collaboration mais également pourquoi est-ce que cela entraîne une réparation plus fidèle. Ung2 n'est évidemment pas le seul acteur requis pour la réparation classique par le BER, et du fait de sa constante catalytique élevée, n'a pas l'air non plus d'être le facteur limitant de cette réparation ([Kavli et al., 1996](#)). Au contraire, une étude implique l'ADN polymérase  $\beta$  comme étant le facteur limitant dans le BER classique ([Nakamura and Swenberg, 1999](#)). De façon intéressante, une étude récente met ainsi également en cause une baisse *in vivo* du niveau d'ADN polymérase  $\beta$  dans des cellules B pour permettre la mutagenèse à A:T pendant l'HMS ([Bahjat et al., 2020](#)). Cela montre qu'il n'y a pas que la quantité d'Ung2 qui est déterminante pour faire pencher la réparation du côté mutagène ou fidèle. Il se pourrait ainsi qu'une quantité accrue d'Ung2 change la façon dont l'ADN polymérase  $\beta$  est recrutée aux loci Ig.

Enfin, cette question de la balance entre réparation fidèle et mutagène par Ung2 ne se limite évidemment pas qu'aux loci Ig. A ce titre, le rôle d'Ung2 comme protecteur ou aggravateur des mutations induites par AID en dehors des loci Ig est assez ambivalent. Il a par exemple été montré qu'elle protège les télomères de l'action d'AID et de la réparation mutagène du MMR lors des mécanismes de diversification ([Cortizas et al., 2016a](#)). En accord avec ce rôle protecteur d'Ung2, il n'est pas rare d'observer de l'hyperplasie lymphoïde dans les souris *Ung<sup>-/-</sup>*. Celles-ci vont même jusqu'à développer relativement tardivement dans leur vie (à partir de 18

mois) des lymphomes B de type folliculaire (FL) ou diffus à grandes cellules B (DLBCL) à une fréquence supérieure à plus de 20 fois celle des souris Ung<sup>+/+</sup> ([Andersen et al., 2005](#); [Nilsen et al., 2003](#)). De la même façon, l'absence d'Ung entraîne l'apparition de 4 fois plus de mutations dans le contexte de *kataegis* ([Taylor et al., 2013](#)), et aboutit ainsi à une toxicité accrue dans des tumeurs où APOBEC3B est surexprimé ([Serebrenik et al., 2019](#)). Cependant, une autre étude montre que la déficience d'Ung2 confère une meilleure protection contre le développement de DLBCL dans des souris où Bcl6 est dérégulé, ce qui implique qu'Ung2 est également à l'origine d'une réparation mutagène des désaminations d'AID en dehors des loci Ig, contrairement au MMR dans ce cas ([Gu et al., 2016](#)).

## 5. Perspectives

Plusieurs expériences sont encore à réaliser pour confirmer les effets que nous observons et pour tenter de les comprendre. Celles-ci sont regroupées par la suite sous forme de questions que nous cherchons à résoudre.

### 5.1 Ung2 est-elle bien surexprimée en absence de Fam72a ?

Au vu des apparentes contradictions entre les résultats que nous obtenons et ceux obtenus par d'autres groupes, il est nécessaire de confirmer nos observations et de valider nos hypothèses en utilisant plusieurs techniques. Ainsi, une des premières expériences à faire serait de confirmer qu'Ung2 est bien surexprimée en absence de Fam72a. Pour cela, nous avons généré des lignées CH12 possédant un KI de la séquence des tags Flag et BioTag au début de la séquence codante d'Ung2, pour aboutir à l'expression de Flag-Biotag-Ung2. De plus, nous prévoyons de générer le KO de Fam72a dans cette lignée KI (Ung2<sup>Flag-Biotag</sup> Fam72a<sup>-/-</sup>). Le BioTag est un peptide reconnu et biotinylé de façon spécifique par la biotine ligase d'*E.coli*, BirA (WT) ([Beckett et al., 1999](#)), qui sera exprimé dans ces 2 lignées (Ung2<sup>Flag-Biotag</sup> Fam72a<sup>+/+</sup> et Ung2<sup>Flag-Biotag</sup> Fam72a<sup>-/-</sup>) à l'aide d'une transduction rétrovirale. Une fois biotinylé, le peptide BioTag nous permettra de repérer facilement le niveau d'expression de Flag-BioTag-Ung2 par WB avec de la streptavidine couplée à l'HRP, afin d'évaluer efficacement la différence de quantité d'Ung2 contenue dans des cellules Fam72a<sup>+/+</sup> par rapport à des cellules Fam72a<sup>-/-</sup>. Ceci permettra de valider les résultats obtenus du WB à l'aide de l'anticorps anti-Ung.

### 5.2 Fam72a a-t-il un autre rôle que celui de dégrader Ung2 ?

Grâce à ces 2 lignées Ki précédemment évoquées, nous pourrons réaliser une immunoprécipitation à la chromatine (ChIP ; de l'anglais chromatin immunoprecipitation) à l'aide de streptavidine pour évaluer l'efficacité du recrutement d'Ung2 aux régions S en présence ou en

absence de Fam72a. Sans ce KI, cette expérience n'est pas réalisable puisque nous ne possédons pas d'anticorps spécifique anti-Ung2 utilisable pour réaliser une immunoprécipitation. Enfin, ce KI peut également être utilisé pour réaliser une co-immunoprécipitation (co-IP) des protéines interagissant avec Ung2, en présence ou en absence de Fam72a. Pour cela, une purification des protéines dans des conditions natives à l'aide d'un anticorps anti-Flag suivie d'une autre purification à l'aide de streptavidine pour éliminer les protéines contaminantes sera réalisée. Les protéines partenaires purifiées seront alors identifiées par spectrométrie de masse. Cette expérience devrait nous permettre de savoir si Fam72a régule l'interaction d'Ung2 avec d'autres partenaires protéiques.

Nous envisageons également de voir si la déficience de CI causée par la surexpression d'Ung2 peut être inversée en abaissant artificiellement le niveau d'expression d'Ung2. Plusieurs possibilités existent, comme, par exemple, l'expression d'Ung2 à partir d'un promoteur plus faible que celui contenu dans les long tandem repeats (LTR) du pMX-PIE ou encore le couplage de la surexpression d'Ung2 avec l'utilisation d'un système de dégradation inductible à l'auxine ([Nishimura et al., 2009](#)). Ainsi, une expression faible d'Ung2 dans des cellules Fam72a<sup>-/-</sup> nous permettra de savoir si le rôle que joue Fam72a se limite à limiter l'expression d'Ung2 ou pas.

### **5.3 Est-ce que le recrutement d'AID à S<sub>μ</sub> est altéré en absence de Fam72a ?**

Du fait que nous observons une baisse de la fréquence de mutations aux loci Ig en absence de Fam72a, nous souhaitons vérifier que celle-ci n'est pas due à un défaut de recrutement ou d'activité d'AID. Nous prévoyons donc de réaliser une expérience de ChIP à l'aide d'un anticorps anti-AID pour évaluer l'efficacité du recrutement d'AID à S<sub>μ</sub> en présence et en absence de Fam72a. De plus, des contrôles pWT et AID<sup>-/-</sup> seront ajoutés.

### **5.4 Est-ce que l'activité d'Ung est altérée en absence de Fam72a ?**

Des données préliminaires présentées dans le manuscrit nous permettent d'évaluer que l'activité d'Ung2 est peu touchée par la présence de Fam72a ([Figure 28](#)).

Il serait néanmoins important de répéter l'expérience de quantification de l'activité d'Ung2 *in vitro*, en réalisant notamment des dilutions en cascade pour mieux la quantifier à la manière de ce qui a été fait dans l'étude impliquant Rev1 dans le ciblage d'Ung2 pour la CI ([Zan et al., 2012](#)).

### **5.5 Y a-t-il moins d'uracile dans l'ADN en absence de Fam72a ?**

Une autre façon de confirmer que la surexpression d'Ung2 est couplée à une réparation plus efficace des uraciles dans l'ADN est de quantifier de façon globale la quantité d'uraciles dans l'ADN en absence de Fam72a. Cependant, il se peut que nous nous heurtions à des problèmes

techniques puisque nous ne sommes pas sûrs que le seuil de détection soit assez bas pour pouvoir quantifier et comparer correctement la quantité d'uraciles dans l'ADN (même en utilisant la spectrométrie de masse ([Sarno et al., 2019a](#))).

### 5.6 La dégradation d'Ung2 est-elle dépendante du protéasome ?

Fam72a pourrait être à l'origine de la dégradation active d'Ung2 par le protéasome. Afin de le savoir, nous prévoyons de mettre en culture des cellules CH12 pWT, Ung<sup>-/-</sup>, Fam72a<sup>-/-</sup>, Fam72a<sup>-/-</sup> reconstituées avec mFam72a et mFam72a<sup>W125R</sup> en présence ou en absence de l'inhibiteur du protéasome MG132. Nous réaliserons alors un western blot pour Ung2 afin de quantifier une éventuelle différence entre les conditions avec ou sans traitement avec MG132.

### 5.7 Y a-t-il plus ou moins de DSBs non réparées induites par AID en l'absence de Fam72a ?

La CI est dépendante de la quantité de DSBs générées dans les régions S, qui sont des intermédiaires obligatoires (Petersen et al., 2001). Le défaut de CI observé dans les cellules B Fam72a<sup>-/-</sup> pourrait s'expliquer par la réparation fidèle imposée en raison de la régulation à la hausse d'Ung2, qui à son tour entraînerait la génération d'une quantité moindre de DSBs dans les régions S. Pour explorer cette possibilité, nous allons tout d'abord évaluer par western blot l'état de phosphorylation des facteurs clés de la réponse aux dommages de l'ADN qui sont ATM (pATM<sup>S1981</sup>) et l'histone H2AX ( $\gamma$ -H2AX), et ce, dans les cellules CH12 Fam72a<sup>+/+</sup>, AID<sup>-/-</sup> et Fam72a<sup>-/-</sup> stimulées. Puis, pour appuyer ce résultat et dans le but de quantifier le nombre de DSBs induites par AID dans les régions S des cellules CH12 Fam72a<sup>+/+</sup>, AID<sup>-/-</sup>, Fam72a<sup>-/-</sup> et les cellules CH12 Fam72a<sup>-/-</sup> reconstituées avec mFam72a et mFam72a<sup>W125R</sup> stimulées, une expérience d'End-Seq (Canela et al., 2016) est en cours. Très brièvement, cette technique de haute résolution permet de mesurer de manière robuste les cassures de l'ADN *in vivo*, même si celles-ci ont une fréquence de faible niveau (Canela et al., 2016). De plus, une expérience de FISH sur les propagations en métaphase permettant d'observer et de quantifier les cassures au sein du locus IgH sur ces mêmes conditions cellulaires viendra compléter les résultats obtenus.

### 5.8 Est-ce que Ung2 est supprimée dans les lignées de cellules cancéreuses humaines surexprimant Fam72a ?

Fam72a a été identifiée comme étant surexprimée dans de nombreux cancers courants tels que ceux du colon, du sein, du poumon, de l'utérus et de l'ovaire ([Guo et al., 2008](#); [Heese, 2013](#); [Pramanik et al., 2015](#)). Le fait que Fam72a soit surexprimé dans ces cancers nous conforte dans l'idée qu'il pourrait avoir un rôle oncogénique en promouvant la réparation non-fidèle des uraciles dans l'ADN. Etant donné que nous avons montré que la surexpression de Fam72a entraîne une dégradation de la quantité d'Ung2, nous souhaitons examiner si on observe également une faible expression d'UNG2 dans des cellules cancéreuses humaines surexprimant FAM72A. Pour cela,

nous prévoyons de réaliser un western blot pour Ung sur les lignées cellulaires humaines cancéreuses HeLa, K562, HL-60, A549, MOLT-4, SW-480 et RAJI, en utilisant les lignées cellulaires U2OS et NIH 3T3 comme contrôles négatifs et les cellules CH12 pWT et Fam72a<sup>-/-</sup> comme contrôles positifs.

## 5.9 Est-ce que la transcription des régions S est affectée dans les cellules B primaires

### Fam72a<sup>-/-</sup> ?

D'après les résultats obtenus dans les cellules CH12 Fam72a<sup>-/-</sup>, la transcription des régions S acceptrice S<sub>μ</sub> et donneuse S<sub>α</sub> n'est pas affectée. Cependant, dans le but de confirmer ces résultats dans les cellules B primaires, nous proposons de réaliser des qRT-PCR pour les GLTs S<sub>μ</sub>, S<sub>γ3</sub>, S<sub>γ2b</sub>, S<sub>γ2a</sub> et S<sub>α</sub>, ainsi que pour Fam72a dans les cellules B primaires Fam72a<sup>-/-</sup> et Fam72a<sup>+/+</sup>.

## 6. Conclusion

En résumé, il nous reste à valider un certain nombre d'expériences avant de pouvoir ériger un modèle qui corresponde avec les résultats observés dans d'autres études s'intéressant au rôle d'Ung2 dans la CI et dans l'HMS. Il se pourrait cependant que certaines de ces expériences permettent de découvrir un nouveau rôle pour Fam72a, ce qui permettrait d'expliquer l'effet de son absence sur l'efficacité de la CI.

Quoiqu'il en soit, la caractérisation plus poussée de ce facteur prometteur permettra de mieux comprendre comment une réparation mutagène est induite aux loci Ig lors de la diversification des anticorps. De même, connaître par quels mécanismes Ung2 promeut une réparation fidèle ou mutagène permettra ainsi de comprendre comment ce facteur peut protéger certains loci de l'action d'AID, et échoue à en protéger d'autres.

## Conclusion générale

Malgré le fort potentiel oncogénique d'AID, la plupart des cellules B exprimant AID ne présente pas de mutations répandues ou d'instabilité génomique. Il apparaît donc que des mécanismes spécifiques de régulation restreignent le potentiel mutagène et recombino-gène d'AID. Ceux-ci comprennent différents niveaux de contrôle de l'activité d'AID et de son ciblage approprié aux gènes Ig, ainsi que des mécanismes régulant la réparation efficace des lésions qu'elle induit. Ces niveaux de contrôle promeuvent une diversification des Ig efficace et robuste et sont garants du maintien de l'intégrité du génome dans les cellules B. Mon travail de thèse s'est efforcé de mettre à jour ces mécanismes de régulation afin de pouvoir les étudier et les comprendre.

Ainsi, mon premier projet s'est articulé autour de l'identification de nouveaux partenaires d'AID pour mieux comprendre comment son ciblage et son activité sont régulés. Pour ce faire, j'ai utilisé la technologie BioID dans les lymphocytes B primaires, ce qui a généré une longue liste de protéines qui interagissent potentiellement avec AID. Afin de pouvoir faire ressortir un nombre plus restreint de gènes candidats à étudier, cette liste a été croisée avec celle obtenue à l'aide d'un crible CRISPR/Cas9 fonctionnel sur la CI. Les données préliminaires de la caractérisation de *Jmjd1c* nous montrent que ce recoupement est valide et peut servir de base pour de futures investigations. La caractérisation plus poussée de ces candidats devrait permettre de mieux comprendre comment l'action d'AID est régulée, afin qu'elle puisse jouer son rôle d'initiatrice de la CI et de l'HMS sans toutefois menacer l'intégrité génomique des cellules B.

Dans un second projet, j'ai participé à l'étude du rôle de *Fam72a* dans la CI et l'HMS. Les résultats que nous obtenons montrent que ce facteur promeut une réparation mutagène des uraciles dans les loci Ig. Avec la caractérisation complète du mode d'action de *Fam72a*, nous espérons mieux comprendre comment se fait le choix entre réparation fidèle ou mutagène sur l'ensemble du génome. Cette problématique est cruciale pour comprendre comment les lymphocytes B sont capables de promouvoir la diversification des anticorps tout en se protégeant de l'action mutagène d'AID. Ces résultats pourraient également avoir une portée thérapeutique, en désignant *Fam72a* comme une nouvelle cible à inhiber pour lutter contre l'apparition de lymphomes B.

# Matériel et méthodes



## Western blot

Des extraits cellulaires totaux sont séparés sur un gel de polyacrylamide Bis-Tris avec un gradient de 4-12% (NuPage ; Invitrogen) par électrophorèse (SDS-PAGE), puis transférés sur une membrane PVDF (Millipore). La membrane est bloquée 1h à température ambiante (RT) ou sur la nuit (O/N) à 4°C avec du TBS-T supplémenté avec du lait (50mM Tris pH 7,5, 150mM NaCl, 0,1% Tween-20, 5% m/v lait en poudre). La membrane est incubée avec l'anticorps ou le réactif primaire dilué dans du TBS-T-BSA 5% pendant 2h à RT ou O/N à 4°C sous agitation douce. Après 3 lavages de 15 minutes avec du TBS-T, les membranes sont incubées avec un anticorps secondaire anti-Fragment constant (Fc) de souris couplé à la peroxidase (HRP) pendant 2h à RT. Après à nouveau 3 lavages de 15 minutes avec du TBS-T, les membranes sont incubées avec un substrat chimioluminescent (SuperSignal West Pico, ThermoScientific ou Immobilon™ western, Millipore) transformé en produit coloré par l'HRP. La présence de ce produit coloré est observée à l'aide d'un Amersham imager 600.

Liste des anticorps primaires / réactifs utilisés

Anticorps/Réactif	Référence	Dilution WB
Flag	Sigma (A8592)	1/1000
AID	IGBMC (Str9)	1/10000
Streptavidine-HRP	ThermoFisher (21126)	1/2500
Actine $\beta$	Sigma (A1978)	1/200000

## Purification des protéines biotinylées dans les cellules B primaires.

50.10<sup>6</sup> cellules B primaires par condition sont récoltées et lavées 2 fois avec du PBS 1x. Les cellules sont ensuite lysées dans 5 mL de tampon RIPA (NP-40 1%, 50 mM Tris-HCl pH 7,5, 150 mM NaCl, 1mM EDTA, 1mM EGTA, 0,1% SDS, 0,5% sodium deoxycholate et Protease Inhibitor Cocktail (PIC) 1X (Roche)). Les échantillons sont ensuite incubés à 4°C pendant 30 min sous rotation. 125 U de benzonase sont ajoutées pendant les 15 dernières minutes de rotation. Les échantillons sont ensuite soniqués (amplitude 50%, ON:10s; OFF: 20s; temps de sonication total : 20s ; appareil QSonica Q700) puis centrifugés à 11 000g pendant 30 mins. Les surnageants correspondant aux extraits totaux sont récupérés. 40 mL de chaque surnageant sont mis de côté en tant qu'input. 90 mL de billes magnétiques couplées à la streptavidine (Dynabeads MyOne T1 ; ThermoFischer) par échantillon sont lavées 2 fois avec du PBS-Tween (0,02%), puis 1 fois avec du tampon RIPA et sont ajoutées à chaque échantillon. Les mix échantillons et billes magnétiques sont incubés à 4°C sous rotation pendant 3 heures. Par la suite, le surnageant (fraction non liée) est éliminé et les billes sont lavées 3 fois avec du tampon RIPA puis 3 fois avec du tampon TAP (10% glycerol, 50mM Hepes NaOH pH8, 150mM NaCl, 2mM EDTA, 0,1%NP40, PIC 1x). Les protéines biotinylées sont ensuite éluées avec 30 mL de tampon Laemmli 2x (100 mM Tris-HCl pH 6,8 ; 4% SDS ; 12% glycerol ; 0,004 % bleu de bromophenol ; 16 mM DTT) par échantillon. 25 mL de chaque échantillon sont déposés sur un gel Bis-Tris 10% (NP 0301 ; Invitrogen NuPage) pour être analysés par spectrométrie de masse. Les protéines contenues dans le gel sont alors digérées à l'aide de LysC pendant 4h à 37°C dans du tampon Tris 0,1M pH8.5 et CaCl<sub>2</sub> 2 mM, puis avec de la trypsine sur la nuit à 37°C. Les peptides sont ensuite analysés grâce à un système nano-LC-MS/MS (nano-LC RSLC U3000 couplé à un Orbitrap ELITE (Thermo Scientific)).

## Méthode d'analyse des données de spectrométrie de masse par Perseus.

Les données d'abondance issues du traitement des données par le logiciel MaxQuant version 1.6.6.0 (Tyanova et al., 2016a) sont analysées à l'aide du logiciel Perseus version 1.6.12.0 (Tyanova et al., 2016b). Les séquences des protéines inversées, des contaminants et des protéines qui ne sont identifiées que dans un seul réplicat technique d'un seul échantillon sont tout d'abord éliminées de l'analyse. Le log en base 2 des valeurs d'abondance est ensuite calculé et utilisé pour l'analyse. Les valeurs d'abondance manquantes sont remplacées par des valeurs faibles d'abondance suivant la distribution d'une loi normale, à l'aide de l'outil d'imputation de Perseus (width = 0.3, down shift = 1.8). La valeur médiane d'abondance de chaque réplicat technique peut également être soustraite, afin de réaliser une normalisation des données. Les volcano plots servant à la comparaison de 2 conditions sont également générés à l'aide de Perseus, en préservant les différents réplicats techniques d'un échantillon de la randomisation. Les valeurs de FDR et de S0 utilisées sont de 0,05 et 0,1 respectivement.

## Cytométrie en flux et tri cellulaire.

1 à  $2 \cdot 10^6$  cellules sont marquées dans 50  $\mu$ L de la dilution appropriée d'anticorps pendant 15 minutes sur glace. Pour les CH12, l'anticorps utilisé est l'anti-IgA couplé à la phycoérythrine (SouthernBiotech ; dilution 1/500e) et ce sont l'anti-IgG1 couplé à la biotine (BD ; 1/500e) et la streptavidine-PE (Beckman Coulter ; 1/1000e) qui sont utilisés pour les lymphocytes B primaires. Le lavage des anticorps non fixés est réalisé par ajout de PBS-BSA 0,5%. Avant d'être analysées, les cellules sont reprises dans du PBS-BSA 0,5% contenant 50 nM de Topro-3 (Invitrogen), un agent fluorescent et intercalant de l'ADN permettant de discriminer les cellules vivantes et les cellules mortes. Le cytomètre en flux utilisé pour l'analyse de l'efficacité de CI est FACS Fortessa (BD) et les résultats sont analysés à l'aide du logiciel FlowJo (Tree Star). Le tri cellulaire des cellules GFP positives se fait à l'aide du FACS ARIA II (BD).

## Liste des ARNg utilisés

Type d'ARNg	Gène ciblé	Exon ciblé	Séquence de complémentarité de l'ARNg (5' vers 3')	Rôle
pG-86	mAID	Exon 1	CTGGAGACCGATATGGAC	KI Nter AID
pG-270	mAID	Exon 5	AGATGCATTTTCGTATGT	KI Cter AID
pGH-150	mAscc3	5' Exon 5	AATGTGATAAATCTACACAG	KO Ascc3
pGH-151	mAscc3	3' Exon 5	TAATCTCAGAGAAAAATCTG	KO Ascc3
pGH-206	mDhx29	5' Exon 2	TGAGTGTGCTTTTCATCAGTG	KO Dhx29
pGH-207	mDhx29	3' Exon 2	TCTATGGCGAAAAGCCACTG	KO Dhx29
pGH-208	mJmjd1c	5' Exon 6	ATGTATTTGATGTTACACTG	KO Jmjd1c
pGH-209	mJmjd1c	3' Exon 7	GATGGTTGTGAGCCACCATG	KO Jmjd1c

Liste des oligonucléotides utilisés pour le crible PCR de clones KI ou KO.

Oligonucléotide PCR	Séquence (5' - 3')
mAID - Exon 1 - 5'	GGACCCAACCCAGGAGGCAGATGT
mAID - Exon 1 - 3'	CCTCTAAGGCTTCGCTGTTATTACCAC
mAID - Exon 5 - 5'	CAATTCTCCGACCTCTGCCTTCT
mAID - Exon 5 - 3'	AAATCCCAACATACGAAATGCATCTC
mAscc3 - Exon 5 - 5'	TGGTGTTTGGTCTTTTTGGGTCTGTTG
mAscc3 - Exon 5 - 3'	CATAATTTGGTGATGGTGGTGGTG
mDhx29 - Exon 2 - 5'	GTGGTTAAGAGCGCTGACTGC
mDhx29 - Exon 2 - 3'	TCCAACCATCTACCGAGTGAC
mJmjd1c - Exon 6 - 5'	ACGCCTTCCATGTGCTGAGTGC
mJmjd1c - Exon 7 - 3'	CAGGTCTAGCTATGCAGGATGAGG

# Références bibliographiques

- Adelman, K., and J.T. Lis. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* 13:720-731.
- Ahnesorg, P., P. Smith, and S.P. Jackson. 2006. XLF Interacts with the XRCC4-DNA Ligase IV Complex to Promote DNA Nonhomologous End-Joining. *Cell* 124:301-313.
- Aida, M., N. Hamad, A. Stanlie, N.A. Begum, and T. Honjo. 2013. Accumulation of the FACT complex, as well as histone H3.3, serves as a target marker for somatic hypermutation. *Proceedings of the National Academy of Sciences* 110:7784-7789.
- Akira, S., K. Okazaki, and H. Sakano. 1987. Two pairs of recombination signals are sufficient to cause immunoglobulin V-(D)-J joining. *238:1134-1138*.
- Alexandrov, L.B., S. Nik-Zainal, D.C. Wedge, S.A.J.R. Aparicio, S. Behjati, A.V. Biankin, G.R. Bignell, N. Bolli, A. Borg, A.-L. Børresen-Dale, S. Boyault, B. Burkhardt, A.P. Butler, C. Caldas, H.R. Davies, C. Desmedt, R. Eils, J.E. Eyfjörd, J.A. Foekens, M. Greaves, F. Hosoda, B. Hutter, T. Ilcic, S. Imbeaud, M. Imielinski, N. Jäger, D.T.W. Jones, D. Jones, S. Knappskog, M. Kool, S.R. Lakhani, C. López-Otín, S. Martin, N.C. Munshi, H. Nakamura, P.A. Northcott, M. Pajic, E. Papaemmanuil, A. Paradiso, J.V. Pearson, X.S. Puente, K. Raine, M. Ramakrishna, A.L. Richardson, J. Richter, P. Rosenstiel, M. Schlesner, T.N. Schumacher, P.N. Span, J.W. Teague, Y. Totoki, A.N.J. Tutt, R. Valdés-Mas, M.M. Van Buuren, L. Van 'T Veer, A. Vincent-Salomon, N. Waddell, L.R. Yates, J. Zucman-Rossi, P. Andrew Futreal, U. McDermott, P. Lichter, M. Meyerson, S.M. Grimmond, R. Siebert, E. Campo, T. Shibata, S.M. Pfister, P.J. Campbell, and M.R. Stratton. 2013. Signatures of mutational processes in human cancer. *Nature* 500:415-421.
- Allen, C.D., K.M. Ansel, C. Low, R. Lesley, H. Tamamura, N. Fujii, and J.G. Cyster. 2004. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat Immunol* 5:943-952.
- Alt, Frederick W., Y. Zhang, F.-L. Meng, C. Guo, and B. Schwer. 2013. Mechanisms of Programmed DNA Lesions and Genomic Instability in the Immune System. *Cell* 152:417-429.
- Alvarez-Prado, A.F., P. Perez-Duran, A. Perez-Garcia, A. Benguria, C. Torroja, V.G. de Yébenes, and A.R. Ramiro. 2018. A broad atlas of somatic hypermutation allows prediction of activation-induced deaminase targets. *J Exp Med* 215:761-771.
- Andersen, S., M. Ericsson, H.Y. Dai, J. Pena-Diaz, G. Slupphaug, H. Nilsen, H. Aarset, and H.E. Krokan. 2005. Monoclonal B-cell hyperplasia and leukocyte imbalance precede development of B-cell malignancies in uracil-DNA glycosylase deficient mice. *DNA Repair (Amst)* 4:1432-1441.
- Anindya, R., O. Aygun, and J.Q. Svejstrup. 2007. Damage-induced ubiquitylation of human RNA polymerase II by the ubiquitin ligase Nedd4, but not Cockayne syndrome proteins or BRCA1. *Mol Cell* 28:386-397.
- Aoufouchi, S., A. Faily, C. Zober, O. D'Orlando, S. Weller, J.C. Weill, and C.A. Reynaud. 2008. Proteasomal degradation restricts the nuclear lifespan of AID. *J Exp Med* 205:1357-1368.
- Arakawa, H., and J.-M. Buerstedde. 2004. Immunoglobulin gene conversion: Insights from bursal B cells and the DT40 cell line. *Developmental Dynamics* 229:458-464.
- Arakawa, H., S. Furusawa, S. Ekino, and H. Yamagishi. 1996. Immunoglobulin gene hyperconversion ongoing in chicken splenic germinal centers. *EMBO J* 15:2540-2546.
- Audebert, M., B. Salles, and P. Calsou. 2004. Involvement of Poly(ADP-ribose) Polymerase-1 and XRCC1/DNA Ligase III in an Alternative Route for DNA Double-strand Breaks Rejoining. *Journal of Biological Chemistry* 279:55117-55126.

- Bahjat, M., M. Stratigopoulou, B. Pilzecker, T.P. Van Dam, S. Mobach, R.J. Bende, C.J.M. Van Noesel, H. Jacobs, and J.E.J. Guikema. 2020. DNA polymerase  $\beta$  prevents AID-instigated mutagenic non-canonical mismatch DNA repair. In Cold Spring Harbor Laboratory.
- Bardwell, P.D., A. Martin, E. Wong, Z. Li, W. Edelmann, and M.D. Scharff. 2003. Cutting Edge: The G-U Mismatch Glycosylase Methyl-CpG Binding Domain 4 Is Dispensable for Somatic Hypermutation and Class Switch Recombination. *170*:1620-1624.
- Bardwell, P.D., C.J. Woo, K. Wei, Z. Li, A. Martin, S.Z. Sack, T. Parris, W. Edelmann, and M.D. Scharff. 2004. Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1–mutant mice. *5*:224-229.
- Barreto, V., B. Reina-San-Martin, A.R. Ramiro, K.M. McBride, and M.C. Nussenzweig. 2003a. C-Terminal Deletion of AID Uncouples Class Switch Recombination from Somatic Hypermutation and Gene Conversion. *Molecular Cell* *12*:501-508.
- Barreto, V., B. Reina-San-Martin, A.R. Ramiro, K.M. McBride, and M.C. Nussenzweig. 2003b. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol Cell* *12*:501-508.
- Bassing, C.H., H. Suh, D.O. Ferguson, K.F. Chua, J. Manis, M. Eckersdorff, M. Gleason, R. Bronson, C. Lee, and F.W. Alt. 2003. Histone H2AX:A Dosage-Dependent Suppressor of Oncogenic Translocations and Tumors. *114*:359-370.
- Basso, K., C. Schneider, Q. Shen, A.B. Holmes, M. Setty, C. Leslie, and R. Dalla-Favera. 2012. BCL6 positively regulates AID and germinal center gene expression via repression of miR-155. *J Exp Med* *209*:2455-2465.
- Basu, U., J. Chaudhuri, C. Alpert, S. Dutt, S. Ranganath, G. Li, J.P. Schrum, J.P. Manis, and F.W. Alt. 2005. The AID antibody diversification enzyme is regulated by protein kinase A phosphorylation. *Nature* *438*:508-511.
- Basu, U., F.L. Meng, C. Keim, V. Grinstein, E. Pefanis, J. Eccleston, T. Zhang, D. Myers, C.R. Wasserman, D.R. Wesemann, K. Januszyk, R.I. Gregory, H. Deng, C.D. Lima, and F.W. Alt. 2011. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* *144*:353-363.
- Batista, F.D., and M.S. Neuberger. 2000. B cells extract and present immobilized antigen: implications for affinity discrimination. *EMBO J* *19*:513-520.
- Beckett, D., E. Kovaleva, and P.J. Schatz. 1999. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci* *8*:921-929.
- Begum, N.A. 2004. Uracil DNA Glycosylase Activity Is Dispensable for Immunoglobulin Class Switch. *Science* *305*:1160-1163.
- Begum, N.A., N. Izumi, M. Nishikori, H. Nagaoka, R. Shinkura, and T. Honjo. 2007. Requirement of non-canonical activity of uracil DNA glycosylase for class switch recombination. *J Biol Chem* *282*:731-742.
- Begum, N.A., K. Kinoshita, N. Kakazu, M. Muramatsu, H. Nagaoka, R. Shinkura, D. Biniszkiewicz, L.A. Boyer, R. Jaenisch, and T. Honjo. 2004. Uracil DNA glycosylase activity is dispensable for immunoglobulin class switch. *Science* *305*:1160-1163.
- Begum, N.A., A. Stanlie, M. Nakata, H. Akiyama, and T. Honjo. 2012. The histone chaperone Spt6 is required for activation-induced cytidine deaminase target determination through H3K4me3 regulation. *J Biol Chem* *287*:32415-32429.
- Bennardo, N., A. Cheng, N. Huang, and J.M. Stark. 2008. Alternative-NHEJ Is a Mechanistically Distinct Pathway of Mammalian Chromosome Break Repair. *PLoS Genetics* *4*:e1000110.
- Bensimon, A., R. Aebersold, and Y. Shiloh. 2011. Beyond ATM: the protein kinase landscape of the DNA damage response. *FEBS Lett* *585*:1625-1639.
- Betz, A.G., C. Rada, R. Pannell, C. Milstein, and M.S. Neuberger. 1993. Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hot spots. *90*:2385-2388.

- Betz, B.C., K.L. Jordan-Williams, C. Wang, S.G. Kang, J. Liao, M.R. Logan, C.H. Kim, and E.J. Taparowsky. 2010. Batf coordinates multiple aspects of B and T cell function required for normal antibody responses. *J Exp Med* 207:933-942.
- Bharati, S., H.E. Krokan, L. Kristiansen, M. Otterlei, and G. Slupphaug. 1998. Human mitochondrial uracil-DNA glycosylase preform (UNG1) is processed to two forms one of which is resistant to inhibition by AP sites. *Nucleic Acids Res* 26:4953-4959.
- Bjorkman, A., L. Du, K. Felgentreff, C. Rosner, R. Pankaj Kamdar, G. Kokaraki, Y. Matsumoto, E.G. Davies, M. Van Der Burg, L.D. Notarangelo, L. Hammarstrom, and Q. Pan-Hammarstrom. 2015. DNA-PKcs Is Involved in Ig Class Switch Recombination in Human B Cells. 195:5608-5615.
- Blagodatski, A., V. Batrak, S. Schmidl, U. Schoetz, R.B. Caldwell, H. Arakawa, and J.M. Buerstedde. 2009. A cis-acting diversification activator both necessary and sufficient for AID-mediated hypermutation. *PLoS Genet* 5:e1000332.
- Boboila, C., F.W. Alt, and B. Schwer. 2012a. Chapter One - Classical and Alternative End-Joining Pathways for Repair of Lymphocyte-Specific and General DNA Double-Strand Breaks. *Advances in Immunology* 116:1-49.
- Boboila, C., F.W. Alt, and B. Schwer. 2012b. Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. *Adv Immunol* 116:1-49.
- Boboila, C., C. Yan, D.R. Wesemann, M. Jankovic, J.H. Wang, J. Manis, A. Nussenzweig, M. Nussenzweig, and F.W. Alt. 2010. Alternative end-joining catalyzes class switch recombination in the absence of both Ku70 and DNA ligase 4. 207:417-427.
- Borchert, G.M., N.W. Holton, and E.D. Larson. 2011. Repression of human activation induced cytidine deaminase by miR-93 and miR-155. *BMC Cancer* 11:347.
- Bothmer, A., Davide, Michela, Samuel, Isaac, N. Feldhahn, J. Barlow, H.-T. Chen, D. Bosque, E. Callen, A. Nussenzweig, and Michel. 2011. Regulation of DNA End Joining, Resection, and Immunoglobulin Class Switch Recombination by 53BP1. *Molecular Cell* 42:319-329.
- Bothmer, A., D.F. Robbiani, N. Feldhahn, A. Gazumyan, A. Nussenzweig, and M.C. Nussenzweig. 2010. 53BP1 regulates DNA resection and the choice between classical and alternative end joining during class switch recombination. *The Journal of Experimental Medicine* 207:855-865.
- Bottaro, A., R. Lansford, L. Xu, J. Zhang, P. Rothman, and F.W. Alt. 1994. S region transcription per se promotes basal IgE class switch recombination but additional factors regulate the efficiency of the process. *EMBO J* 13:665-674.
- Bransteitter, R., P. Pham, M.D. Scharff, and M.F. Goodman. 2003a. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc Natl Acad Sci U S A* 100:4102-4107.
- Bransteitter, R., P. Pham, M.D. Scharff, and M.F. Goodman. 2003b. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proceedings of the National Academy of Sciences of the United States of America* 100:4102-4107.
- Brar, S.S., M. Watson, and M. Diaz. 2004. Activation-induced Cytosine Deaminase (AID) Is Actively Exported out of the Nucleus but Retained by the Induction of DNA Breaks. 279:26395-26401.
- Buerstedde, J.M., J. Alinikula, H. Arakawa, J.J. McDonald, and D.G. Schatz. 2014. Targeting of somatic hypermutation by immunoglobulin enhancer and enhancer-like sequences. *PLoS Biol* 12:e1001831.
- Burns, M.B., L. Lackey, M.A. Carpenter, A. Rathore, A.M. Land, B. Leonard, E.W. Refsland, D. Kotandeniya, N. Tretyakova, J.B. Nikas, D. Yee, N.A. Temiz, D.E. Donohue, R.M. McDougale, W.L. Brown, E.K. Law, and R.S. Harris. 2013. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature* 494:366-370.
- Calis, J.J., and B.R. Rosenberg. 2014. Characterizing immune repertoires by high throughput sequencing: strategies and applications. *Trends Immunol* 35:581-590.

- Callebaut, I., L. Malivert, A. Fischer, J.P. Mornon, P. Revy, and J.P. De Villartay. 2006. Cernunnos Interacts with the XRCC4{middle dot}DNA-ligase IV Complex and Is Homologous to the Yeast Nonhomologous End-joining Factor Nej1. 281:13857-13860.
- Canugovi, C., M. Samaranayake, and A.S. Bhagwat. 2009. Transcriptional pausing and stalling causes multiple clustered mutations by human activation-induced deaminase. *FASEB J* 23:34-44.
- Casellas, R., U. Basu, W.T. Yewdell, J. Chaudhuri, D.F. Robbiani, and J.M. Di Noia. 2016a. Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. *Nature Reviews Immunology* 16:164-176.
- Casellas, R., U. Basu, W.T. Yewdell, J. Chaudhuri, D.F. Robbiani, and J.M. Di Noia. 2016b. Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. *Nat Rev Immunol* 16:164-176.
- Casellas, R., A. Nussenzweig, R. Wuerffel, R. Pelanda, A. Reichlin, H. Sush, X.-F. Qin, E. Besmer, A. Kenter, K. Rajewsky, and M. C.Nussenzweig. 1998. Ku80 is required for immunoglobulin isotype switching. *The EMBO Journal* 17:2404–2411.
- Celeste, A., S. Petersen, P.J. Romanienko, O. Fernandez-Capetillo, H.T. Chen, O.A. Sedelnikova, B. Reina-San-Martin, V. Coppola, E. Meffre, M.J. Difilippantonio, C. Redon, D.R. Pilch, A. Olaru, M. Eckhaus, R.D. Camerini-Otero, L. Tessarollo, F. Livak, K. Manova, W.M. Bonner, M.C. Nussenzweig, and A. Nussenzweig. 2002. Genomic Instability in Mice Lacking Histone H2AX. *Science* 296:922-927.
- Chahwan, R., W. Edelmann, M.D. Scharff, and S. Roa. 2012. AIDing antibody diversity by error-prone mismatch repair. *Seminars in Immunology* 24:293-300.
- Chandra, A., F. van Maldegem, S. Andrews, M.S. Neuberger, and C. Rada. 2013. Deficiency in spliceosome-associated factor CTNNB1 does not affect ongoing cell cycling but delays exit from quiescence and results in embryonic lethality in mice. *Cell Cycle* 12:732-742.
- Chang, H.H.Y., and M.R. Lieber. 2016. Structure-Specific nuclease activities of Artemis and the Artemis: DNA-PKcs complex. *Nucleic Acids Research* 44:4991-4997.
- Chang, H.H.Y., N.R. Pannunzio, N. Adachi, and M.R. Lieber. 2017. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature Reviews Molecular Cell Biology* 18:495-506.
- Chaudhuri, J., and F.W. Alt. 2004. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nature Reviews Immunology* 4:541-552.
- Chaudhuri, J., U. Basu, A. Zarrin, C. Yan, S. Franco, T. Perlot, B. Vuong, J. Wang, R.T. Phan, A. Datta, J. Manis, and F.W. Alt. 2007. Evolution of the Immunoglobulin Heavy Chain Class Switch Recombination Mechanism. In *AID for Immunoglobulin Diversity*. 157-214.
- Chaudhuri, J., M. Tian, C. Khuong, K. Chua, E. Pinaud, and F.W. Alt. 2003. Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature* 422:726-730.
- Chen, J., Z. Cai, M. Bai, X. Yu, C. Zhang, C. Cao, X. Hu, L. Wang, R. Su, D. Wang, L. Wang, Y. Yao, R. Ye, B. Hou, Y. Yu, S. Yu, J. Li, and Y. Xue. 2018. The RNA-binding protein ROD1/PTBP3 cotranscriptionally defines AID-loading sites to mediate antibody class switch in mammalian genomes. *Cell Res* 28:981-995.
- Chen, M., N. Zhu, X. Liu, B. Laurent, Z. Tang, R. Eng, Y. Shi, S.A. Armstrong, and R.G. Roeder. 2015. JMJD1C is required for the survival of acute myeloid leukemia by functioning as a coactivator for key transcription factors. *Genes & Development* 29:2123-2139.
- Cheng, H.L., B.Q. Vuong, U. Basu, A. Franklin, B. Schwer, J. Astarita, R.T. Phan, A. Datta, J. Manis, F.W. Alt, and J. Chaudhuri. 2009. Integrity of the AID serine-38 phosphorylation site is critical for class switch recombination and somatic hypermutation in mice. 106:2717-2722.
- Chevillard, C., J. Ozaki, C.D. Herring, and R. Riblet. 2002. A Three-Megabase Yeast Artificial Chromosome Contig Spanning the C57BL Mouse Igh Locus. 168:5659-5666.
- Choi, J.E., A.J. Matthews, G. Michel, and B.Q. Vuong. 2020. AID Phosphorylation Regulates Mismatch Repair–Dependent Class Switch Recombination and Affinity Maturation. *The Journal of Immunology* 204:13-22.



- Chojnowski, A., P.F. Ong, E.S. Wong, J.S. Lim, R.A. Motalif, R. Navasankari, B. Dutta, H. Yang, Y.Y. Liow, S.K. Sze, T. Boudier, G.D. Wright, A. Colman, B. Burke, C.L. Stewart, and O. Dreesen. 2015. Progerin reduces LAP2alpha-telomere association in Hutchinson-Gilford progeria. *Elife* 4: Ciubotaru, M., A.J. Trexler, L.N. Spiridon, M.D. Surleac, E. Rhoades, A.J. Petrescu, and D.G. Schatz. 2013. RAG and HMGB1 create a large bend in the 23RSS in the V(D)J recombination synaptic complexes. 41:2437-2454.
- Cobb, O. treich, Osipovic, and Oltz. 2006. Accessibility Control of V(D)J Recombination. *advances in immunology* 91:
- Coticello, S.G., K. Ganesh, K. Xue, M. Lu, C. Rada, and M.S. Neuberger. 2008. Interaction between antibody-diversification enzyme AID and spliceosome-associated factor CTNNB1. *Mol Cell* 31:474-484.
- Cook, A.J.L., L. Oganessian, P. Harumal, A. Basten, R. Brink, and C.J. Jolly. 2003. Reduced Switching in SCID B Cells Is Associated with Altered Somatic Mutation of Recombined S Regions. *The Journal of Immunology* 171:6556-6564.
- Core, L.J., and J.T. Lis. 2008. Transcription Regulation Through Promoter-Proximal Pausing of RNA Polymerase II. *Science* 319:1791-1792.
- Cortizas, E.M., A. Zahn, M.E. Hajjar, A.-M. Patenaude, J.M.D. Noia, and R.E. Verdun. 2013. Alternative End-Joining and Classical Nonhomologous End-Joining Pathways Repair Different Types of Double-Strand Breaks during Class-Switch Recombination. *The Journal of Immunology* 191:5751-5763.
- Cortizas, E.M., A. Zahn, S. Safavi, J.A. Reed, F. Vega, J.M. Di Noia, and R.E. Verdun. 2016a. UNG protects B cells from AID-induced telomere loss. *J Exp Med* 213:2459-2472.
- Cortizas, E.M., A. Zahn, S. Safavi, J.A. Reed, F. Vega, J.M.D. Noia, and R.E. Verdun. 2016b. UNG protects B cells from AID-induced telomere loss. *Journal of Experimental Medicine* 213:2459-2472.
- Crouch, E.E., Z. Li, M. Takizawa, S. Fichtner-Feigl, P. Gourzi, C. Montano, L. Feigenbaum, P. Wilson, S. Janz, F.N. Papavasiliou, and R. Casellas. 2007. Regulation of AID expression in the immune response. *J Exp Med* 204:1145-1156.
- Cyster, J.G. 2010. B cell follicles and antigen encounters of the third kind. *Nat Immunol* 11:989-996.
- D'Addabbo, P., M. Scascitelli, V. Giambra, M. Rocchi, and D. Frezza. 2011. Position and sequence conservation in Amniota of polymorphic enhancer HS1.2 within the palindrome of IgH 3'Regulatory Region. 11:71.
- Dango, S., N. Mosammamarast, M.E. Sowa, L.J. Xiong, F. Wu, K. Park, M. Rubin, S. Gygi, J.W. Harper, and Y. Shi. 2011. DNA unwinding by ASCC3 helicase is coupled to ALKBH3-dependent DNA alkylation repair and cancer cell proliferation. *Mol Cell* 44:373-384.
- Daniel, J.A., M.A. Santos, Z. Wang, C. Zang, K.R. Schwab, M. Jankovic, D. Filsuf, H.T. Chen, A. Gazumyan, A. Yamane, Y.W. Cho, H.W. Sun, K. Ge, W. Peng, M.C. Nussenzweig, R. Casellas, G.R. Dressler, K. Zhao, and A. Nussenzweig. 2010. PTIP promotes chromatin changes critical for immunoglobulin class switch recombination. *Science* 329:917-923.
- Daniels, G.A., and M.R. Lieber. 1995. RNA:DNA complex formation upon transcription of immunoglobulin switch regions: implications for the mechanism and regulation of class switch recombination. *Nucleic Acids Res* 23:5006-5011.
- De Silva, N.S., and U. Klein. 2015. Dynamics of B cells in germinal centres. *Nat Rev Immunol* 15:137-148.
- de Yebenes, V.G., L. Belver, D.G. Pisano, S. Gonzalez, A. Villasante, C. Croce, L. He, and A.R. Ramiro. 2008. miR-181b negatively regulates activation-induced cytidine deaminase in B cells. *J Exp Med* 205:2199-2206.
- Dedeoglu, F., B. Horwitz, J. Chaudhuri, F.W. Alt, and R.S. Geha. 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.

- Delbos, F.D.R., S. Aoufouchi, A. Faili, J.-C. Weill, and C.-A.S. Reynaud. 2007. DNA polymerase  $\eta$  is the sole contributor of A/T modifications during immunoglobulin gene hypermutation in the mouse. *Journal of Experimental Medicine* 204:17-23.
- Delgado, P., Á.F. Álvarez-Prado, E. Marina-Zárate, I.V. Sernandez, S.M. Mur, J. De La Barrera, F. Sanchez-Cabo, M. Cañamero, A. De Molina, L. Belver, V.G. De Yébenes, and A.R. Ramiro. 2020. Interplay between UNG and AID governs intratumoral heterogeneity in mature B cell lymphoma. In Cold Spring Harbor Laboratory.
- Dhote, V., T.R. Sweeney, N. Kim, C.U.T. Hellen, and T.V. Pestova. 2012. Roles of individual domains in the function of DHX29, an essential factor required for translation of structured mammalian mRNAs. 109:E3150-E3159.
- Di Noia, J., and M.S. Neuberger. 2002. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature* 419:43-48.
- Di Noia, J.M., C. Rada, and M.S. Neuberger. 2006. SMUG1 is able to excise uracil from immunoglobulin genes: insight into mutation versus repair. *The EMBO Journal* 25:585-595.
- Di Noia, J.M., G.T. Williams, D.T. Chan, J.M. Buerstedde, G.S. Baldwin, and M.S. Neuberger. 2007a. Dependence of antibody gene diversification on uracil excision. *J Exp Med* 204:3209-3219.
- Di Noia, J.M., G.T. Williams, D.T.Y. Chan, J.-M. Buerstedde, G.S. Baldwin, and M.S. Neuberger. 2007b. Dependence of antibody gene diversification on uracil excision. 204:3209-3219.
- Dickerson, S.K., E. Market, E. Besmer, and F.N. Papavasiliou. 2003a. AID mediates hypermutation by deaminating single stranded DNA. *J Exp Med* 197:1291-1296.
- Dickerson, S.K., E. Market, E. Besmer, and F.N. Papavasiliou. 2003b. AID Mediates Hypermutation by Deaminating Single Stranded DNA. *Journal of Experimental Medicine* 197:1291-1296.
- Dinesh, R.K., B. Barnhill, A. Ilanges, L. Wu, D.A. Michelson, F. Senigl, J. Alinikula, J. Shabanowitz, D.F. Hunt, and D.G. Schatz. 2020. Transcription factor binding at Ig enhancers is linked to somatic hypermutation targeting. *Eur J Immunol* 50:380-395.
- Dingler, F.A., K. Kemmerich, M.S. Neuberger, and C. Rada. 2014. Uracil excision by endogenous SMUG1 glycosylase promotes efficient Ig class switching and impacts on A:T substitutions during somatic mutation. *European Journal of Immunology* 44:1925-1935.
- Dinkelmann, M., E. Spehalski, T. Stoneham, J. Buis, Y. Wu, J.M. Sekiguchi, and D.O. Ferguson. 2009a. Multiple functions of MRN in end-joining pathways during isotype class switching. *Nat Struct Mol Biol* 16:808-813.
- Dinkelmann, M., E. Spehalski, T. Stoneham, J. Buis, Y. Wu, J.M. Sekiguchi, and D.O. Ferguson. 2009b. Multiple functions of MRN in end-joining pathways during isotype class switching. *Nature Structural & Molecular Biology* 16:808-813.
- Dobbs, T.A., J.A. Tainer, and S.P. Lees-Miller. 2010. A structural model for regulation of NHEJ by DNA-PKcs autophosphorylation. *DNA Repair* 9:1307-1314.
- Doi, T., L. Kato, S. Ito, R. Shinkura, M. Wei, H. Nagaoka, J. Wang, and T. Honjo. 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 U S A* 106:2758-2763.
- Doil, C., N. Mailand, S. Bekker-Jensen, P. Menard, D.H. Larsen, R. Pepperkok, J. Ellenberg, S. Panier, D. Durocher, J. Bartek, J. Lukas, and C. Lukas. 2009. RNF168 Binds and Amplifies Ubiquitin Conjugates on Damaged Chromosomes to Allow Accumulation of Repair Proteins. *Cell* 136:435-446.
- Dominguez-Sola, D., G.D. Victora, C.Y. Ying, R.T. Phan, M. Saito, M.C. Nussenzweig, and R. Dalla-Favera. 2012. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. *Nat Immunol* 13:1083-1091.
- Dong, J., R.A. Panchakshari, T. Zhang, Y. Zhang, J. Hu, S.A. Volpi, R.M. Meyers, Y.J. Ho, Z. Du, D.F. Robbiani, F. Meng, M. Gostissa, M.C. Nussenzweig, J.P. Manis, and F.W. Alt. 2015. Orientation-specific joining of AID-initiated DNA breaks promotes antibody class switching. *Nature* 525:134-139.

- Donnelly, M.L.L., G. Luke, A. Mehrotra, X. Li, L.E. Hughes, D. Gani, and M.D. Ryan. 2001. Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *J Gen Virol* 82:1013-1025.
- Dorsett, Y., K.M. McBride, M. Jankovic, A. Gazumyan, T.H. Thai, D.F. Robbani, M. Di Virgilio, B. Reina San-Martin, G. Heidkamp, T.A. Schwickert, T. Eisenreich, K. Rajewsky, and M.C. Nussenzweig. 2008. MicroRNA-155 suppresses activation-induced cytidine deaminase-mediated Myc-Igh translocation. *Immunity* 28:630-638.
- Doseth, B., C. Ekre, G. Slupphaug, H.E. Krokan, and B. Kavli. 2012. Strikingly different properties of uracil-DNA glycosylases UNG2 and SMUG1 may explain divergent roles in processing of genomic uracil. *DNA Repair (Amst)* 11:587-593.
- Doseth, B., T. Visnes, A. Wallenius, I. Ericsson, A. Sarno, H.S. Pettersen, A. Flatberg, T. Catterall, G. Slupphaug, H.E. Krokan, and B. Kavli. 2011. Uracil-DNA glycosylase in base excision repair and adaptive immunity: species differences between man and mouse. *J Biol Chem* 286:16669-16680.
- Drolet, M., S. Broccoli, F. Rallu, C. Hraiky, C. Fortin, E. Masse, and I. Baaklini. 2003. The problem of hypernegative supercoiling and R-loop formation in transcription. *Front Biosci* 8:d210-221.
- Dudley, D.D., J.P. Manis, A.A. Zarrin, L. Kaylor, M. Tian, and F.W. Alt. 2002. Internal IgH class switch region deletions are position-independent and enhanced by AID expression. *Proceedings of the National Academy of Sciences of the United States of America* 99:9984-9989.
- Dunnick, W.A., J.T. Collins, J. Shi, G. Westfield, C. Fontaine, P. Hakimpour, and F.N. Papavasiliou. 2009. Switch recombination and somatic hypermutation are controlled by the heavy chain 3' enhancer region. 206:2613-2623.
- Dunnick, W.A., J. Shi, C. Fontaine, and J.T. Collins. 2013. Transgenes of the mouse immunoglobulin heavy chain locus, lacking distal elements in the 3' regulatory region, are impaired for class switch recombination. *PLoS One* 8:e55842.
- Duquette, M.L., P. Handa, J.A. Vincent, A.F. Taylor, and N. Maizels. 2004. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes Dev* 18:1618-1629.
- Eccleston, J., C. Yan, K. Yuan, F.W. Alt, and E. Selsing. 2011. Mismatch Repair Proteins MSH2, MLH1, and EXO1 Are Important for Class-Switch Recombination Events Occurring in B Cells That Lack Nonhomologous End Joining. *The Journal of Immunology* 186:2336-2343.
- Ehrenstein, M.R., C. Rada, A.M. Jones, C. Milstein, and M.S. Neuberger. 2001. Switch junction sequences in PMS2-deficient mice reveal a microhomology-mediated mechanism of Ig class switch recombination. 98:14553-14558.
- Eldin, P., S. Peron, A. Galashevskaya, N. Denis-Lagache, M. Cogne, G. Slupphaug, and L. Briant. 2020. Impact of HIV-1 Vpr manipulation of the DNA repair enzyme UNG2 on B lymphocyte class switch recombination. *J Transl Med* 18:310.
- Endo, Y., H. Marusawa, T. Kou, H. Nakase, S. Fujii, T. Fujimori, K. Kinoshita, T. Honjo, and T. Chiba. 2008. Activation-induced cytidine deaminase links between inflammation and the development of colitis-associated colorectal cancers. *Gastroenterology* 135:889-898, 898 e881-883.
- Ettinger, R., G.P. Sims, A.M. Fairhurst, R. Robbins, Y.S. da Silva, R. Spolski, W.J. Leonard, and P.E. Lipsky. 2005. IL-21 induces differentiation of human naive and memory B cells into antibody-secreting plasma cells. *J Immunol* 175:7867-7879.
- Fagarasan, S., K. Kinoshita, M. Muramatsu, K. Ikuta, and T. Honjo. 2001. In situ class switching and differentiation to IgA-producing cells in the gut lamina propria. 413:639-643.
- Faili, A., S. Aoufouchi, Q. Gueranger, C. Zober, A. Leon, B. Bertocci, J.C. Weill, and C.A. Reynaud. 2002. AID-dependent somatic hypermutation occurs as a DNA single-strand event in the BL2 cell line. *Nat Immunol* 3:815-821.
- Falck, J., J. Coates, and S.P. Jackson. 2005. Conserved modes of recruitment of ATM, ATR and DNA-PKcs to sites of DNA damage. *Nature* 434:605-611.

- Feng, Y., N. Seija, D.I.N. JM, and A. Martin. 2020. AID in Antibody Diversification: There and Back Again. *Trends Immunol* 41:586-600.
- Fernandez, D., M. Ortiz, L. Rodriguez, A. Garcia, D. Martinez, and I. Moreno de Alboran. 2013. The proto-oncogene c-myc regulates antibody secretion and Ig class switch recombination. *J Immunol* 190:6135-6144.
- Franco, S., M. Gostissa, S. Zha, D.B. Lombard, M.M. Murphy, A.A. Zarrin, C. Yan, S. Tepsuporn, J.C. Morales, M.M. Adams, Z. Lou, C.H. Bassing, J.P. Manis, J. Chen, P.B. Carpenter, and F.W. Alt. 2006. H2AX Prevents DNA Breaks from Progressing to Chromosome Breaks and Translocations. *Molecular Cell* 21:201-214.
- Franco, S., M.M. Murphy, G. Li, T. Borjeson, C. Boboila, and F.W. Alt. 2008. DNA-PKcs and Artemis function in the end-joining phase of immunoglobulin heavy chain class switch recombination. *205:557-564.*
- Frey, S., B. Bertocci, F. Delbos, L. Quint, J.-C. Weill, and C.-A. Reynaud. 1998. Mismatch Repair Deficiency Interferes with the Accumulation of Mutations in Chronically Stimulated B Cells and Not with the Hypermutation Process. *9:127-134.*
- Frieder, D., M. Larijani, C. Collins, M. Shulman, and A. Martin. 2009a. The concerted action of Msh2 and UNG stimulates somatic hypermutation at A · T base pairs. *Mol Cell Biol* 29:5148-5157.
- Frieder, D., M. Larijani, C. Collins, M. Shulman, and A. Martin. 2009b. The Concerted Action of Msh2 and UNG Stimulates Somatic Hypermutation at A · T Base Pairs. *Molecular and Cellular Biology* 29:5148-5157.
- Frit, P., V. Ropars, M. Modesti, J.B. Charbonnier, and P. Calsou. 2019. Plugged into the Ku-DNA hub: The NHEJ network. *Prog Biophys Mol Biol* 147:62-76.
- Fritz, E.L., B.R. Rosenberg, K. Lay, A. Mihailović, T. Tuschl, and F.N. Papavasiliou. 2013. A comprehensive analysis of the effects of the deaminase AID on the transcriptome and methylome of activated B cells. *Nature Immunology* 14:749-755.
- Fujimura, S., T. Matsui, K. Kuwahara, K. Maeda, and N. Sakaguchi. 2008. Germinal center B-cell-associated DNA hypomethylation at transcriptional regions of the AID gene. *Mol Immunol* 45:1712-1719.
- Gaidano, G., A. Carbone, C. Pastore, D. Capello, A. Migliazza, A. Gloghini, S. Roncella, M. Ferrarini, G. Saglio, and R. and Dalla-Favera. 1997. Frequent Mutation of the 5' Noncoding Region of the BCL-6 Gene in Acquired Immunodeficiency Syndrome-Related Non-Hodgkin's Lymphomas. *Blood* 89:3755-3762.
- Gan, H., T. Shen, D.P. Chupp, J.R. Taylor, H.N. Sanchez, X. Li, Z. Xu, H. Zan, and P. Casali. 2020. B cell Sirt1 deacetylates histone and non-histone proteins for epigenetic modulation of AID expression and the antibody response. *Sci Adv* 6:eaay2793.
- Ganesh, K., S. Adam, B. Taylor, P. Simpson, C. Rada, and M. Neuberger. 2011. CTNNB1 is a novel nuclear localization sequence-binding protein that recognizes RNA-splicing factors CDC5L and Prp31. *J Biol Chem* 286:17091-17102.
- Garside, P., E. Ingulli, R.R. Merica, J.G. Johnson, R.J. Noelle, and M.K. Jenkins. 1998. Visualization of specific B and T lymphocyte interactions in the lymph node. *Science* 281:96-99.
- Gazumyan, A., K. Timachova, G. Yuen, E. Siden, M. Di Virgilio, E.M. Woo, B.T. Chait, B. Reina San-Martin, M.C. Nussenzweig, and K.M. McBride. 2011. Amino-terminal phosphorylation of activation-induced cytidine deaminase suppresses c-myc/IgH translocation. *Mol Cell Biol* 31:442-449.
- Geisberger, R., C. Rada, and M.S. Neuberger. 2009. The stability of AID and its function in class-switching are critically sensitive to the identity of its nuclear-export sequence. *Proc Natl Acad Sci U S A* 106:6736-6741.
- Ghezraoui, H., C. Oliveira, J.R. Becker, K. Bilham, D. Moralli, C. Anzilotti, R. Fischer, M. Deobagkar-Lele, M. Sanchiz-Calvo, E. Fueyo-Marcos, S. Bonham, B.M. Kessler, S. Rottenberg, R.J. Cornall, C.M. Green, and J.R. Chapman. 2018. 53BP1 cooperation with the REV7–shieldin complex underpins DNA structure-specific NHEJ. *Nature* 560:122-127.

- Girelli Zubani, G., M. Zivojnovic, A. De Smet, O. Albagli-Curiel, F. Huetz, J.-C. Weill, C.-A. Reynaud, and S. Storck. 2017. Pms2 and uracil-DNA glycosylases act jointly in the mismatch repair pathway to generate Ig gene mutations at A-T base pairs. *jem*.20161576.
- Gitlin, A.D., C.T. Mayer, T.Y. Oliveira, Z. Shulman, M.J.K. Jones, A. Koren, and M.C. Nussenzweig. 2015. T cell help controls the speed of the cell cycle in germinal center B cells. *Science* 349:643-646.
- Gitlin, A.D., Z. Shulman, and M.C. Nussenzweig. 2014. Clonal selection in the germinal centre by regulated proliferation and hypermutation. *Nature* 509:637-640.
- Goenka, R., A.H. Matthews, B. Zhang, P.J. O'Neill, J.L. Scholz, T.S. Migone, W.J. Leonard, W. Stohl, U. Hershberg, and M.P. Cancro. 2014. Local BlyS production by T follicular cells mediates retention of high affinity B cells during affinity maturation. *J Exp Med* 211:45-56.
- Gonda, H., M. Sugai, Y. Nambu, T. Katakai, Y. Agata, K.J. Mori, Y. Yokota, and A. Shimizu. 2003. The balance between Pax5 and Id2 activities is the key to AID gene expression. *J Exp Med* 198:1427-1437.
- Gordon, M.S., C.M. Kanegai, J.R. Doerr, and R. Wall. 2003. Somatic hypermutation of the B cell receptor genes B29 (Ig , CD79b) and mb1 (Ig , CD79a). 100:4126-4131.
- Graham, T.G.W., J.C. Walter, and a.J.J. Loparo. 2016. Two-Stage Synapsis of DNA Ends during Non-homologous End Joining. *Molecular Cell* 61:850-858.
- Grawunder, U., X.W. M Wilm, P. Kulesza, T.E. Wilson, M. Mann, and M.R. Lieber. 1997. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* 31:492-495.
- Gu, X., C.J. Booth, Z. Liu, and M.P. Strout. 2016. AID-associated DNA repair pathways regulate malignant transformation in a murine model of BCL6-driven diffuse large B-cell lymphoma. *Blood* 127:102-112.
- Guo, C., X. Zhang, S.P. Fink, P. Platzer, K. Wilson, J.K. Willson, Z. Wang, and S.D. Markowitz. 2008. Ugene, a newly identified protein that is commonly overexpressed in cancer and binds uracil DNA glycosylase. *Cancer Res* 68:6118-6126.
- Gupta, R., K. Somyajit, T. Narita, E. Maskey, A. Stanlie, M. Kremer, D. Typas, M. Lammers, N. Mailand, A. Nussenzweig, J. Lukas, and C. Choudhary. 2018. DNA Repair Network Analysis Reveals Shieldin as a Key Regulator of NHEJ and PARP Inhibitor Sensitivity. *Cell* 173:972-988 e923.
- Hackney, J.A., S. Misaghi, K. Senger, C. Garris, Y. Sun, M.N. Lorenzo, and A.A. Zarrin. 2009. Chapter 5 DNA Targets of AID. In 163-189.
- Hagen, L., B. Kavli, M.M. Sousa, K. Torseth, N.B. Liabakk, O. Sundheim, J. Pena-Diaz, M. Otterlei, O. Horning, O.N. Jensen, H.E. Krokan, and G. Slupphaug. 2008a. Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA. *EMBO J* 27:51-61.
- Hagen, L., B. Kavli, M.M. Sousa, K. Torseth, N.B. Liabakk, O. Sundheim, J. Peña-Diaz, M. Otterlei, O. Hørning, O.N. Jensen, H.E. Krokan, and G. Slupphaug. 2008b. Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA. *The EMBO Journal* 27:51-61.
- Haince, J.F., D. McDonald, A. Rodrigue, U. Dery, J.Y. Masson, M.J. Hendzel, and G.G. Poirier. 2008. PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites. *J Biol Chem* 283:1197-1208.
- Han, L., S. Masani, and K. Yu. 2010. Cutting Edge: CTNNB1 Is Dispensable for Ig Class Switch Recombination. 185:1379-1381.
- Han, S., K. Hathcock, B. Zheng, T.B. Kepler, R. Hodes, and G. Kelsoe. 1995. Cellular interaction in germinal centers. Roles of CD40 ligand and B7-2 in established germinal centers. *J Immunol* 155:556-567.
- Harriman, G.R., A. Bradley, S. Das, P. Rogers-Fani, and A.C. Davis. 1996. IgA class switch in I alpha exon-deficient mice. Role of germline transcription in class switch recombination. 97:477-485.

- Harris, R.S., J.E. Sale, S.K. Petersen-Mahrt, and M.S. Neuberger. 2002. AID Is Essential for Immunoglobulin V Gene Conversion in a Cultured B Cell Line. *Current Biology* 12:435-438.
- Hasham, M.G., N.M. Donghia, E. Coffey, J. Maynard, K.J. Snow, J. Ames, R.Y. Wilpan, Y. He, B.L. King, and K.D. Mills. 2010. Widespread genomic breaks generated by activation-induced cytidine deaminase are prevented by homologous recombination. *Nat Immunol* 11:820-826.
- Hasler, J., C. Rada, and M.S. Neuberger. 2011. Cytoplasmic activation-induced cytidine deaminase (AID) exists in stoichiometric complex with translation elongation factor 1 (eEF1A). *108:18366-18371*.
- He, M., E.M. Cortizas, R.E. Verdun, and E. Severinson. 2015. Cyclin-dependent kinases regulate Ig class switching by controlling access of AID to the switch region. *J Immunol* 194:4231-4239.
- Heese, K. 2013. The protein p17 signaling pathways in cancer. *Tumour Biol* 34:4081-4087.
- Heesters, B.A., R.C. Myers, and M.C. Carroll. 2014. Follicular dendritic cells: dynamic antigen libraries. *Nat Rev Immunol* 14:495-504.
- Hein, K., M.G. Lorenz, G. Siebenkotten, K. Petry, R. Christine, and A. Radbruch. 1998. Processing of switch transcripts is required for targeting of antibody class switch recombination. *J Exp Med* 188:2369-2374.
- Helleday, T. 2003. Pathways for mitotic homologous recombination in mammalian cells. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 532:103-115.
- Hozumi, N., and S. Tonegawa. 1976. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proceedings of the National Academy of Sciences* 73:3628-3632.
- Hu, W., N.A. Begum, S. Mondal, A. Stanlie, and T. Honjo. 2015. Identification of DNA cleavage- and recombination-specific hnRNP cofactors for activation-induced cytidine deaminase. *Proc Natl Acad Sci U S A* 112:5791-5796.
- Hu, Y., I. Ericsson, K. Torseth, S.P. Methot, O. Sundheim, N.B. Liabakk, G. Slupphaug, J.M. Di Noia, H.E. Krokan, and B. Kavli. 2013. A combined nuclear and nucleolar localization motif in activation-induced cytidine deaminase (AID) controls immunoglobulin class switching. *J Mol Biol* 425:424-443.
- Huen, M.S.Y., R. Grant, I. Manke, K. Minn, X. Yu, M.B. Yaffe, and J. Chen. 2007. RNF8 Transduces the DNA-Damage Signal via Histone Ubiquitylation and Checkpoint Protein Assembly. *Cell* 131:901-914.
- Huong le, T., M. Kobayashi, M. Nakata, G. Shioi, H. Miyachi, T. Honjo, and H. Nagaoka. 2013. In vivo analysis of Aicda gene regulation: a critical balance between upstream enhancers and intronic silencers governs appropriate expression. *PLoS One* 8:e61433.
- Hustedt, N., and D. Durocher. 2016. The control of DNA repair by the cell cycle. *Nat Cell Biol* 19:1-9.
- Hwang, J.K., F.W. Alt, and L.S. Yeap. 2015. Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination. *Microbiol Spectr* 3:MDNA3-0037-2014.
- Ijspeert, H., P.A. Van Schouwenburg, I. Pico-Knijnenburg, J. Loeffen, L. Brugieres, G.J. Driessen, C. Blattmann, M. Suerink, D. Januszkiewicz-Lewandowska, A.A. Azizi, M.G. Seidel, H. Jacobs, and M. Van Der Burg. 2019. Repertoire Sequencing of B Cells Elucidates the Role of UNG and Mismatch Repair Proteins in Somatic Hypermutation in Humans. *Frontiers in Immunology* 10:
- Imai, K., G. Slupphaug, W.-I. Lee, P. Revy, S. Nonoyama, N. Catalan, L. Yel, M. Forveille, B. Kavli, H.E. Krokan, H.D. Ochs, A. Fischer, and A. Durandy. 2003a. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. *Nature Immunology* 4:1023-1028.
- Imai, K., G. Slupphaug, W.I. Lee, P. Revy, S. Nonoyama, N. Catalan, L. Yel, M. Forveille, B. Kavli, H.E. Krokan, H.D. Ochs, A. Fischer, and A. Durandy. 2003b. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. *Nat Immunol* 4:1023-1028.
- Incorvaia, E., L. Sicouri, S.K. Petersen-Mahrt, and K.M. Schmitz. 2013. Hormones and AID: balancing immunity and autoimmunity. *Autoimmunity* 46:128-137.

- Ise, W., M. Kohyama, B.U. Schraml, T. Zhang, B. Schwer, U. Basu, F.W. Alt, J. Tang, E.M. Oltz, T.L. Murphy, and K.M. Murphy. 2011. The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. *Nat Immunol* 12:536-543.
- Ito, S., H. Nagaoka, R. Shinkura, N. Begum, M. Muramatsu, M. Nakata, and T. Honjo. 2004a. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc Natl Acad Sci U S A* 101:1975-1980.
- Ito, S., H. Nagaoka, R. Shinkura, N. Begum, M. Muramatsu, M. Nakata, and T. Honjo. 2004b. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proceedings of the National Academy of Sciences* 101:1975-1980.
- Jacob, J., R. Kassir, and G. Kelsoe. 1991. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. I. The architecture and dynamics of responding cell populations. *The Journal of Experimental Medicine* 173:1165-1175.
- Jansen, J.G., P. Langerak, A. Tsaalbi-Shtylik, P. Van Den Berk, H. Jacobs, and N. De Wind. 2006. Strand-biased defect in C/G transversions in hypermutating immunoglobulin genes in Rev1-deficient mice. *The Journal of Experimental Medicine* 203:319-323.
- Jeevan-Raj, B.P., I. Robert, V. Heyer, A. Page, J.H. Wang, F. Cammas, F.W. Alt, R. Losson, and B. Reina-San-Martin. 2011. Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. *J Exp Med* 208:1649-1660.
- Jette, N., and S.P. Lees-Miller. 2015. The DNA-dependent protein kinase: A multifunctional protein kinase with roles in DNA double strand break repair and mitosis. *Progress in Biophysics and Molecular Biology* 117:194-205.
- Jiang, W., Jennifer, X. Liu, S. Nakajima, Y. Wang, C. Li, Brian, Richard, C. Liu, X. Yu, L. Lan, and S. Zha. 2015. Differential Phosphorylation of DNA-PKcs Regulates the Interplay between End-Processing and End-Ligation during Nonhomologous End-Joining. *Molecular Cell* 58:172-185.
- Jung, S., K. Rajewsky, and A. Radbruch. 1993. Shutdown of class switch recombination by deletion of a switch region control element. 259:984-987.
- Kanehiro, Y., K. Todo, M. Negishi, J. Fukuoka, W. Gan, T. Hikasa, Y. Kaga, M. Takemoto, M. Magari, X. Li, J.L. Manley, H. Ohmori, and N. Kanayama. 2012. Activation-induced cytidine deaminase (AID)-dependent somatic hypermutation requires a splice isoform of the serine/arginine-rich (SR) protein SRSF1. *Proc Natl Acad Sci U S A* 109:1216-1221.
- Kavli, B., S. Andersen, M. Otterlei, N.B. Liabakk, K. Imai, A. Fischer, A. Durandy, H.E. Krokan, and G. Slupphaug. 2005. B cells from hyper-IgM patients carrying UNG mutations lack ability to remove uracil from ssDNA and have elevated genomic uracil. *Journal of Experimental Medicine* 201:2011-2021.
- Kavli, B., G. Slupphaug, C.D. Mol, A.S. Arvai, S.B. Peterson, J.A. Tainer, and H.E. Krokan. 1996. Excision of cytosine and thymine from DNA by mutants of human uracil-DNA glycosylase. *The EMBO Journal* 15:3442-3447.
- Kavli, B., O. Sundheim, M. Akbari, M. Otterlei, H. Nilsen, F. Skorpen, P.A. Aas, L. Hagen, H.E. Krokan, and G. Slupphaug. 2002. hUNG2 Is the Major Repair Enzyme for Removal of Uracil from U:A Matches, U:G Mismatches, and U in Single-stranded DNA, with hSMUG1 as a Broad Specificity Backup. 277:39926-39936.
- Kawaguchi, Y., H. Nariki, N. Kawamoto, Y. Kanehiro, S. Miyazaki, M. Suzuki, M. Magari, H. Tokumitsu, and N. Kanayama. 2017. SRSF1-3 contributes to diversification of the immunoglobulin variable region gene by promoting accumulation of AID in the nucleus. *Biochem Biophys Res Commun* 485:261-266.
- Khair, L., J.E.J. Guikema, E.K. Linehan, A.J. Ucher, N.G.J. Leus, C. Ogilvie, Z. Lou, C.E. Schrader, and J. Stavnezer. 2014. ATM Increases Activation-Induced Cytidine Deaminase Activity at Downstream S Regions during Class-Switch Recombination. 192:4887-4896.
- Khamlichi, A.A., F. Gludet, Z. Oruc, V. Denis, M. Le Bert, and M. Cogne. 2004. Immunoglobulin class-switch recombination in mice devoid of any S mu tandem repeat. *Blood* 103:3828-3836.

- Kieffer-Kwon, K.R., Z. Tang, E. Mathe, J. Qian, M.H. Sung, G. Li, W. Resch, S. Baek, N. Pruett, L. Grontved, L. Vian, S. Nelson, H. Zare, O. Hakim, D. Reyon, A. Yamane, H. Nakahashi, A.L. Kovalchuk, J. Zou, J.K. Joung, V. Sartorelli, C.L. Wei, X. Ruan, G.L. Hager, Y. Ruan, and R. Casellas. 2013. Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation. *Cell* 155:1507-1520.
- Kim, D.I., K.C. Birendra, W. Zhu, K. Motamedchaboki, V. Doye, and K.J. Roux. 2014. Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proc Natl Acad Sci U S A* 111:E2453-2461.
- Kim, J.H., S.R. Lee, L.H. Li, H.J. Park, J.H. Park, K.Y. Lee, M.K. Kim, B.A. Shin, and S.Y. Choi. 2011. High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS One* 6:e18556.
- Kim, Y., and M. Tian. 2010. The recruitment of activation induced cytidine deaminase to the immunoglobulin locus by a regulatory element. *Mol Immunol* 47:1860-1865.
- King, Justin J., Courtney A. Manuel, Crystal V. Barrett, S. Raber, H. Lucas, P. Sutter, and M. Larijani. 2015. Catalytic Pocket Inaccessibility of Activation-Induced Cytidine Deaminase Is a Safeguard against Excessive Mutagenic Activity. *Structure* 23:615-627.
- Klein, F., R. Diskin, J.F. Scheid, C. Gaebler, H. Mouquet, I.S. Georgiev, M. Pancera, T. Zhou, R.B. Incesu, B.Z. Fu, P.N. Gnanapragasam, T.Y. Oliveira, M.S. Seaman, P.D. Kwong, P.J. Bjorkman, and M.C. Nussenzweig. 2013. Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. *Cell* 153:126-138.
- Kluiver, J., E. Haralambieva, D. de Jong, T. Blokzijl, S. Jacobs, B.J. Kroesen, S. Poppema, and A. van den Berg. 2006. Lack of BIC and microRNA miR-155 expression in primary cases of Burkitt lymphoma. *Genes Chromosomes Cancer* 45:147-153.
- Ko, R., and S.E. Bennett. 2005. Physical and functional interaction of human nuclear uracil-DNA glycosylase with proliferating cell nuclear antigen. *DNA Repair (Amst)* 4:1421-1431.
- Kobayashi, J., H. Tauchi, S. Sakamoto, A. Nakamura, K.-I. Morishima, S. Matsuura, T. Kobayashi, K. Tamai, K. Tanimoto, and K. Komatsu. 2002. NBS1 Localizes to  $\gamma$ -H2AX Foci through Interaction with the FHA/BRCT Domain. *Current Biology* 12:1846-1851.
- Kochan, J.A., E.C.B. Desclos, R. Bosch, L. Meister, L.E.M. Vriend, H. van Attikum, and P.M. Krawczyk. 2017. Meta-analysis of DNA double-strand break response kinetics. *Nucleic Acids Res* 45:12625-12637.
- Kohler, K.M., J.J. McDonald, J.L. Duke, H. Arakawa, S. Tan, S.H. Kleinstein, J.M. Buerstedde, and D.G. Schatz. 2012. Identification of core DNA elements that target somatic hypermutation. *J Immunol* 189:5314-5326.
- Kohli, R.M., S.R. Abrams, K.S. Gajula, R.W. Maul, P.J. Gearhart, and J.T. Stivers. 2009. A Portable Hot Spot Recognition Loop Transfers Sequence Preferences from APOBEC Family Members to Activation-induced Cytidine Deaminase. *Journal of Biological Chemistry* 284:22898-22904.
- Kolas, N.K., J.R. Chapman, S. Nakada, J. Ylanko, R. Chahwan, F.D. Sweeney, S. Panier, M. Mendez, J. Wildenhain, T.M. Thomson, L. Pelletier, S.P. Jackson, and D. Durocher. 2007. Orchestration of the DNA-Damage Response by the RNF8 Ubiquitin Ligase. 318:1637-1640.
- Kothapalli, N., D.D. Norton, and S.D. Fugmann. 2008. Cutting edge: a cis-acting DNA element targets AID-mediated sequence diversification to the chicken Ig light chain gene locus. *J Immunol* 180:2019-2023.
- Kou, T., H. Marusawa, K. Kinoshita, Y. Endo, I.-M. Okazaki, Y. Ueda, Y. Kodama, H. Haga, I. Ikai, and T. Chiba. 2007. Expression of activation-induced cytidine deaminase in human hepatocytes during hepatocarcinogenesis. 120:469-476.
- Kracker, S., Y. Bergmann, I. Demuth, P.O. Frappart, G. Hildebrand, R. Christine, Z.Q. Wang, K. Sperling, M. Digweed, and A. Radbruch. 2005. Nibrin functions in Ig class-switch recombination. *Proceedings of the National Academy of Sciences* 102:1584-1589.
- Kracker, S., M. Di Virgilio, J. Schwartzenuber, C. Cuenin, M. Forveille, M.C. Deau, K.M. McBride, J. Majewski, A. Gazumyan, S. Seneviratne, B. Grimbacher, N. Kutukculer, Z. Herceg, M. Cavazzana, N. Jabado, M.C. Nussenzweig, A. Fischer, and A. Durandy. 2015. An inherited



- immunoglobulin class-switch recombination deficiency associated with a defect in the INO80 chromatin remodeling complex. *J Allergy Clin Immunol* 135:998-1007 e1006.
- Krijger, P.H., A. Tsaalbi-Shtylik, N. Wit, P.C.M. Van Den Berk, N. De Wind, and H. Jacobs. 2013. Rev1 is essential in generating G to C transversions downstream of the Ung2 pathway but not the Msh2+Ung2 hybrid pathway. *43:2765-2770*.
- Krijger, P.H.L., P. Langerak, P.C.M. Van Den Berk, and H. Jacobs. 2009. Dependence of nucleotide substitutions on Ung2, Msh2, and PCNA-Ub during somatic hypermutation. *206:2603-2611*.
- Krokan, H.E., F. Drabløs, and G. Slupphaug. 2002. Uracil in DNA – occurrence, consequences and repair. *Oncogene* 21:8935-8948.
- Kruhlak, M., E.E. Crouch, M. Orlov, C. Montañó, S.A. Gorski, A. Nussenzweig, T. Misteli, R.D. Phair, and R. Casellas. 2007. The ATM repair pathway inhibits RNA polymerase I transcription in response to chromosome breaks. *Nature* 447:730-734.
- Kuchen, S., R. Robbins, G.P. Sims, C. Sheng, T.M. Phillips, P.E. Lipsky, and R. Ettinger. 2007. Essential role of IL-21 in B cell activation, expansion, and plasma cell generation during CD4+ T cell-B cell collaboration. *J Immunol* 179:5886-5896.
- Kudo, N., N. Matsumori, H. Taoka, D. Fujiwara, E.P. Schreiner, B. Wolff, M. Yoshida, and S. Horinouchi. 1999. Leptomycin B inactivates CRM1/exportin 1 by covalent modification at a cysteine residue in the central conserved region. *Proc Natl Acad Sci U S A* 96:9112-9117.
- Kumar Singh, A., A. Tamrakar, A. Jaiswal, N. Kanayama, A. Agarwal, P. Tripathi, and P. Kodgire. 2019. Splicing regulator SRSF1-3 that controls somatic hypermutation of IgV genes interacts with topoisomerase 1 and AID. *Mol Immunol* 116:63-72.
- Küppers, R. 2005. Mechanisms of B-cell lymphoma pathogenesis. *Nature Reviews Cancer* 5:251-262.
- Küppers, R., and R. Dalla-Favera. 2001. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene* 20:5580-5594.
- Kutzner, A., S. Pramanik, P.S. Kim, and K. Heese. 2015. All-or-(N)One - an epistemological characterization of the human tumorigenic neuronal paralogous FAM72 gene loci. *Genomics* 106:278-285.
- Kuwahara, K., S. Fujimura, Y. Takahashi, N. Nakagata, T. Takemori, S. Aizawa, and N. Sakaguchi. 2004. Germinal center-associated nuclear protein contributes to affinity maturation of B cell antigen receptor in T cell-dependent responses. *101:1010-1015*.
- LaCava, J., J. Fernandez-Martinez, Z. Hakhverdyan, and M.P. Rout. 2016. Protein Complex Purification by Affinity Capture. *Cold Spring Harb Protoc* 2016:pdb.top077545.
- Lackey, L., E.K. Law, W.L. Brown, and R.S. Harris. 2013. Subcellular localization of the APOBEC3 proteins during mitosis and implications for genomic DNA deamination. *Cell Cycle* 12:762-772.
- Lada, A.G., A. Dhar, R.J. Boissy, M. Hirano, A.A. Rubel, I.B. Rogozin, and Y.I. Pavlov. 2012. AID/APOBEC cytosine deaminase induces genome-wide kataegis. *7:47*.
- Lafaille, J.J., A. DeCloux, M. Bonneville, Y. Takagaki, and S. Tonegawa. 1989. Junctional sequences of T cell receptor gamma delta genes: implications for gamma delta T cell lineages and for a novel intermediate of V-(D)-J joining. *Cell* 59:859-870.
- Lahdesmaki, A., A.M.R. Taylor, K.H. Chrzanowska, and Q. Pan-Hammarstrom. 2004. Delineation of the Role of the Mre11 Complex in Class Switch Recombination. *279:16479-16487*.
- Langerak, P., A.O.H. Nygren, P.H.L. Krijger, P.C.M. Van Den Berk, and H. Jacobs. 2007. A/T mutagenesis in hypermutated immunoglobulin genes strongly depends on PCNAK164 modification. *Journal of Experimental Medicine* 204:1989-1998.
- Larijani, M., and A. Martin. 2012. The biochemistry of activation-induced deaminase and its physiological functions. *Semin Immunol* 24:255-263.
- Larijani, M., A.P. Petrov, O. Kolenchenko, M. Berru, S.N. Krylov, and A. Martin. 2007. AID Associates with Single-Stranded DNA with High Affinity and a Long Complex Half-Life in a Sequence-Independent Manner. *Molecular and Cellular Biology* 27:20-30.

- Le Noir, S., B. Laffleur, C. Carrion, A. Garot, S. Lecardeur, E. Pinaud, Y. Denizot, J. Skok, and M. Cogne. 2017. The IgH locus 3' cis-regulatory super-enhancer co-opts AID for allelic transvection. *Oncotarget* 8:12929-12940.
- Le, Q., and N. Maizels. 2015. Cell Cycle Regulates Nuclear Stability of AID and Determines the Cellular Response to AID. *PLoS Genet* 11:e1005411.
- Le, Q., and N. Maizels. 2019. Activation-induced deaminase (AID) localizes to the nucleus in brief pulses. *PLoS Genet* 15:e1007968.
- Lebecque, S.G., and P.J. Gearhart. 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. *The Journal of Experimental Medicine* 172:1717-1727.
- Lee-Theilen, M., A.J. Matthews, D. Kelly, S. Zheng, and J. Chaudhuri. 2011. CtIP promotes microhomology-mediated alternative end joining during class-switch recombination. *Nature Structural & Molecular Biology* 18:75-79.
- Lenz, G., I. Nagel, R. Siebert, A.V. Roschke, W. Sanger, G.W. Wright, S.S. Dave, B. Tan, H. Zhao, A. Rosenwald, H.K. Muller-Hermelink, R.D. Gascoyne, E. Campo, E.S. Jaffe, E.B. Smeland, R.I. Fisher, W.M. Kuehl, W.C. Chan, and L.M. Staudt. 2007. Aberrant immunoglobulin class switch recombination and switch translocations in activated B cell–like diffuse large B cell lymphoma. *Journal of Experimental Medicine* 204:633-643.
- Lescale, C., and L. Deriano. 2017. The RAG recombinase: Beyond breaking. *Mech Ageing Dev* 165:3-9.
- Li, F., Y. Yan, J. Pieretti, D.A. Feldman, and L.A. Eckhardt. 2010a. Comparison of Identical and Functional Igh Alleles Reveals a Nonessential Role for  $\mu$  in Somatic Hypermutation and Class-Switch Recombination. *The Journal of Immunology* 185:6049-6057.
- Li, L., M.-J. Halaby, A. Hakem, R. Cardoso, S. El Ghamrasni, S. Harding, N. Chan, R. Bristow, O. Sanchez, D. Durocher, and R. Hakem. 2010b. Rnf8 deficiency impairs class switch recombination, spermatogenesis, and genomic integrity and predisposes for cancer. 207:983-997.
- Liang, L., J. Feng, P. Zuo, J. Yang, Y. Lu, and Y. Yin. 2020. Molecular basis for assembly of the shieldin complex and its implications for NHEJ. *Nature Communications* 11:
- Lio, C.J., V. Shukla, D. Samaniego-Castruita, E. Gonzalez-Avalos, A. Chakraborty, X. Yue, D.G. Schatz, F. Ay, and A. Rao. 2019. TET enzymes augment activation-induced deaminase (AID) expression via 5-hydroxymethylcytosine modifications at the Aicda superenhancer. *Sci Immunol* 4:
- Liu, D., H. Xu, C. Shih, Z. Wan, X. Ma, W. Ma, D. Luo, and H. Qi. 2015. T-B-cell entanglement and ICOSL-driven feed-forward regulation of germinal centre reaction. *Nature* 517:214-218.
- Liu, M., J.L. Duke, D.J. Richter, C.G. Vinuesa, C.C. Goodnow, S.H. Kleinstein, and D.G. Schatz. 2008. Two levels of protection for the B cell genome during somatic hypermutation. *Nature* 451:841-845.
- Liu, Y.J., F. Malisan, O. De Bouteiller, C. Guret, S. Lebecque, J. Banchereau, F.C. Mills, E.E. Max, and H. Martinez-Valdez. 1996. Within Germinal Centers, Isotype Switching of Immunoglobulin Genes Occurs after the Onset of Somatic Mutation. 4:241-250.
- Lorenz, M., S. Jung, and A. Radbruch. 1995. Switch transcripts in immunoglobulin class switching. *Science* 267:1825-1828.
- Lou, Z., K. Minter-Dykhouse, S. Franco, M. Gostissa, M.A. Rivera, A. Celeste, J.P. Manis, J. Van Deursen, A. Nussenzweig, T.T. Paull, F.W. Alt, and J. Chen. 2006. MDC1 Maintains Genomic Stability by Participating in the Amplification of ATM-Dependent DNA Damage Signals. *Molecular Cell* 21:187-200.
- Lu, G., J. Duan, S. Shu, X. Wang, L. Gao, J. Guo, and Y. Zhang. 2016. Ligase I and ligase III mediate the DNA double-strand break ligation in alternative end-joining. *Proceedings of the National Academy of Sciences* 113:1256-1260.
- Lu, X., D. Bocangel, B. Nannenga, H. Yamaguchi, E. Appella, and L.A. Donehower. 2004. The p53-Induced Oncogenic Phosphatase PPM1D Interacts with Uracil DNA Glycosylase and Suppresses Base Excision Repair. 15:621-634.

- Luning Prak, E.T., M. Monestier, and R.A. Eisenberg. 2011. B cell receptor editing in tolerance and autoimmunity. *Ann N Y Acad Sci* 1217:96-121.
- Luo, Y., Y. Liu, L. Wu, X. Ma, Q. Liu, F. Huang, X. Zhang, Y. Zhang, J. Zhang, H. Luo, Y. Yang, G. Lu, X. Tang, L. Li, Y. Zeng, T. Pan, and H. Zhang. 2019. CUL7 E3 Ubiquitin Ligase Mediates the Degradation of Activation-Induced Cytidine Deaminase and Regulates the Ig Class Switch Recombination in B Lymphocytes. *J Immunol* 203:269-281.
- Ma, K.W., S.W. Au, and M.M. Waye. 2009. Over-expression of SUMO-1 induces the up-regulation of heterogeneous nuclear ribonucleoprotein A2/B1 isoform B1 (hnRNP A2/B1 isoform B1) and uracil DNA glycosylase (UDG) in hepG2 cells. *Cell Biochem Funct* 27:228-237.
- Ma, Y., U. Pannicke, K. Schwarz, and M.R. Lieber. 2002. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* 108:781-794.
- Macara, I.G. 2001. Transport into and out of the nucleus. *Microbiol Mol Biol Rev* 65:570-594, table of contents.
- Maeda, K., S.K. Singh, K. Eda, M. Kitabatake, P. Pham, M.F. Goodman, and N. Sakaguchi. 2010. GANP-mediated Recruitment of Activation-induced Cytidine Deaminase to Cell Nuclei and to Immunoglobulin Variable Region DNA. 285:23945-23953.
- Mai, T., H. Zan, J. Zhang, J.S. Hawkins, Z. Xu, and P. Casali. 2010. Estrogen Receptors Bind to and Activate theHOXC4/HoxC4Promoter to Potentiate HoxC4-mediated Activation-induced Cytosine Deaminase Induction, Immunoglobulin Class Switch DNA Recombination, and Somatic Hypermutation. *Journal of Biological Chemistry* 285:37797-37810.
- Mailand, N., S. Bekker-Jensen, H. Fastrup, F. Melander, J. Bartek, C. Lukas, and J. Lukas. 2007. RNF8 Ubiquitylates Histones at DNA Double-Strand Breaks and Promotes Assembly of Repair Proteins. 131:887-900.
- Manis, J.P., Y. Gu, R. Lansford, E. Sonoda, R. Ferrini, L. Davidson, K. Rajewsky, and F.W. Alt. 1998. Ku70 is required for late B cell development and immunoglobulin heavy chain class switching. *The Journal of Experimental Medicine* 187:2081-2089.
- Manis, J.P., J.C. Morales, Z. Xia, J.L. Kutok, F.W. Alt, and P.B. Carpenter. 2004a. 53BP1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination. *Nat Immunol* 5:481-487.
- Manis, J.P., J.C. Morales, Z. Xia, J.L. Kutok, F.W. Alt, and P.B. Carpenter. 2004b. 53BP1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination. 5:481-487.
- Martin, A., P.D. Bardwell, C.J. Woo, M. Fan, M.J. Shulman, and M.D. Scharff. 2002. Activation-induced cytidine deaminase turns on somatic hypermutation in hybridomas. 415:802-806.
- Martomo, S.A., W.W. Yang, and P.J. Gearhart. 2004. A Role for Msh6 But Not Msh3 in Somatic Hypermutation and Class Switch Recombination. 200:61-68.
- Masani, S., L. Han, K. Meek, and K. Yu. 2016. Redundant function of DNA ligase 1 and 3 in alternative end-joining during immunoglobulin class switch recombination. 201521630.
- Matsumoto, Y., H. Marusawa, K. Kinoshita, Y. Endo, T. Kou, T. Morisawa, T. Azuma, I.-M. Okazaki, T. Honjo, and T. Chiba. 2007. Helicobacter pylori infection triggers aberrant expression of activation-induced cytidine deaminase in gastric epithelium. 13:470-476.
- Matthews, A.J., S. Husain, and J. Chaudhuri. 2014. Binding of AID to DNA does not correlate with mutator activity. *J Immunol* 193:252-257.
- Maul, R.W., Z. Cao, L. Venkataraman, C.A. Giorgetti, J.L. Press, Y. Denizot, H. Du, R. Sen, and P.J. Gearhart. 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, R.W., H. Chon, K. Sakhuja, S.M. Cerritelli, L.A. Gugliotti, P.J. Gearhart, and R.J. Crouch. 2017. R-Loop Depletion by Over-expressed RNase H1 in Mouse B Cells Increases Activation-Induced Deaminase Access to the Transcribed Strand without Altering Frequency of Isotype Switching. *J Mol Biol* 429:3255-3263.

- McBride, K.M., V. Barreto, A.R. Ramiro, P. Stavropoulos, and M.C. Nussenzweig. 2004. Somatic hypermutation is limited by CRM1-dependent nuclear export of activation-induced deaminase. *J Exp Med* 199:1235-1244.
- McBride, K.M., A. Gazumyan, E.M. Woo, V.M. Barreto, D.F. Robbiani, B.T. Chait, and M.C. Nussenzweig. 2006. Regulation of hypermutation by activation-induced cytidine deaminase phosphorylation. 103:8798-8803.
- McBride, K.M., A. Gazumyan, E.M. Woo, T.A. Schwickert, B.T. Chait, and M.C. Nussenzweig. 2008. Regulation of class switch recombination and somatic mutation by AID phosphorylation. *J Exp Med* 205:2585-2594.
- Meffre, E., and H. Wardemann. 2008. B-cell tolerance checkpoints in health and autoimmunity. *Current Opinion in Immunology* 20:632-638.
- Mehta, V., and L. Trinkle-Mulcahy. 2016. Recent advances in large-scale protein interactome mapping. *F1000Res* 5:782.
- Melchers, F., E.t. Boekel, T. Seid, I.C. Kong, T. Yamagami, K. Onishi, T. Shimizu, A.G. Rolink, and J. Andersson. 2000. Repertoire selection by pre-B-cell receptors and B-cell receptors, and genetic control of B-cell development from immature to mature B cells. *Immunological Reviews* 175:33-46.
- Mesin, L., J. Ersching, and G.D. Victora. 2016. Germinal Center B Cell Dynamics. *Immunity* 45:471-482.
- Methot, S.P., and J.M. Di Noia. 2017a. Molecular Mechanisms of Somatic Hypermutation and Class Switch Recombination. In Elsevier, 37-87.
- Methot, S.P., and J.M. Di Noia. 2017b. Molecular Mechanisms of Somatic Hypermutation and Class Switch Recombination. *Adv Immunol* 133:37-87.
- Methot, S.P., L.C. Litzler, P.G. Subramani, A.K. Eranki, H. Fifield, A.M. Patenaude, J.C. Gilmore, G.E. Santiago, H. Bagci, J.F. Cote, M. Larijani, R.E. Verdun, and J.M. Di Noia. 2018. A licensing step links AID to transcription elongation for mutagenesis in B cells. *Nat Commun* 9:1248.
- Methot, S.P., L.C. Litzler, F. Trajtenberg, A. Zahn, F. Robert, J. Pelletier, A. Buschiazzi, B.G. Magor, and J.M. Di Noia. 2015. Consecutive interactions with HSP90 and eEF1A underlie a functional maturation and storage pathway of AID in the cytoplasm. *J Exp Med* 212:581-596.
- Migliazza, A., S. Martinotti, W. Chen, C. Fusco, B.H. Ye, D.M. Knowles, K. Offit, R.S. Chaganti, and R. Dalla-Favera. 1995. Frequent somatic hypermutation of the 5' noncoding region of the BCL6 gene in B-cell lymphoma. 92:12520-12524.
- Mohni, K.N., S.R. Wessel, R. Zhao, A.C. Wojciechowski, J.W. Luzwick, H. Layden, B.F. Eichman, P.S. Thompson, K.P.M. Mehta, and D. Cortez. 2019. HMCES Maintains Genome Integrity by Shielding Abasic Sites in Single-Strand DNA. *Cell* 176:144-153.e113.
- Mondal, S., N.A. Begum, W. Hu, and T. Honjo. 2016. Functional requirements of AID's higher order structures and their interaction with RNA-binding proteins. *Proc Natl Acad Sci U S A* 113:E1545-1554.
- Montamat-Sicotte, D., L.C. Litzler, C. Abreu, S. Safavi, A. Zahn, A. Orthwein, M. Muschen, P. Oppezio, D.P. Munoz, and J.M. Di Noia. 2015. HSP90 inhibitors decrease AID levels and activity in mice and in human cells. *Eur J Immunol* 45:2365-2376.
- Morgan, H.D., W. Dean, H.A. Coker, W. Reik, and S.K. Petersen-Mahrt. 2004. Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J Biol Chem* 279:52353-52360.
- Muramatsu, M., K. Kinoshita, S. Fagarasan, S. Yamada, Y. Shinkai, and T. Honjo. 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102:553-563.
- Muramatsu, M., V.S. Sankaranand, S. Anant, M. Sugai, K. Kinoshita, N.O. Davidson, and T. Honjo. 1999a. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. *J Biol Chem* 274:18470-18476.
- Muramatsu, M., V.S. Sankaranand, S. Anant, M. Sugai, K. Kinoshita, N.O. Davidson, and T. Honjo. 1999b. Specific Expression of Activation-induced Cytidine Deaminase (AID), a Novel Member

- of the RNA-editing Deaminase Family in Germinal Center B Cells. *Journal of Biological Chemistry* 274:18470-18476.
- Muschen, M., D. Re, A. Brauninger, J. Wolf, M.L. Hansmann, V. Diehl, R. Kuppers, and K. Rajewsky. 2000. Somatic mutations of the CD95 gene in Hodgkin and Reed-Sternberg cells. *Cancer Res* 60:5640-5643.
- Muto, T., I.M. Okazaki, S. Yamada, Y. Tanaka, K. Kinoshita, M. Muramatsu, H. Nagaoka, and T. Honjo. 2006. Negative regulation of activation-induced cytidine deaminase in B cells. *Proceedings of the National Academy of Sciences* 103:2752-2757.
- Nagelhus, T.A., T. Haug, K.K. Singh, K.F. Keshav, F. Skorpen, M. Otterlei, S. Bharati, T. Lindmo, S. Benichou, R. Benarous, and H.E. Krokan. 1997. A Sequence in the N-terminal Region of Human Uracil-DNA Glycosylase with Homology to XPA Interacts with the C-terminal Part of the 34-kDa Subunit of Replication Protein A. *Journal of Biological Chemistry* 272:6561-6566.
- Nakamura, J., and J.A. Swenberg. 1999. Endogenous apurinic/apyrimidinic sites in genomic DNA of mammalian tissues. *Cancer Res* 59:2522-2526.
- Nakamura, M., S. Kondo, M. Sugai, M. Nazarea, S. Imamura, and T. Honjo. 1996. High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. *Int Immunol* 8:193-201.
- Nambu, Y., M. Sugai, H. Gonda, C.G. Lee, T. Katakai, Y. Agata, Y. Yokota, and A. Shimizu. 2003. Transcription-coupled events associating with immunoglobulin switch region chromatin. *Science* 302:2137-2140.
- Nehar, S., M. Mishra, and K. Heese. 2009. Identification and characterisation of the novel amyloid-beta peptide-induced protein p17. *FEBS Lett* 583:3247-3253.
- Nemazee, D. 2017. Mechanisms of central tolerance for B cells. *Nature Reviews Immunology* 17:281-294.
- Nicolas, L., M. Cols, J.E. Choi, J. Chaudhuri, and B. Vuong. 2018. Generating and repairing genetically programmed DNA breaks during immunoglobulin class switch recombination. *F1000Res* 7:458.
- Nik-Zainal, S., L.B. Alexandrov, D.C. Wedge, P. Van Loo, C.D. Greenman, K. Raine, D. Jones, J. Hinton, J. Marshall, L.A. Stebbings, A. Menzies, S. Martin, K. Leung, L. Chen, C. Leroy, M. Ramakrishna, R. Rance, K.W. Lau, L.J. Mudie, I. Varela, D.J. McBride, G.R. Bignell, S.L. Cooke, A. Shlien, J. Gamble, I. Whitmore, M. Maddison, P.S. Tarpey, H.R. Davies, E. Papaemmanuil, P.J. Stephens, S. McLaren, A.P. Butler, J.W. Teague, G. Jonsson, J.E. Garber, D. Silver, P. Miron, A. Fatima, S. Boyault, A. Langerod, A. Tutt, J.W. Martens, S.A. Aparicio, A. Borg, A.V. Salomon, G. Thomas, A.L. Borresen-Dale, A.L. Richardson, M.S. Neuberger, P.A. Futreal, P.J. Campbell, M.R. Stratton, and C. Breast Cancer Working Group of the International Cancer Genome. 2012. Mutational processes molding the genomes of 21 breast cancers. *Cell* 149:979-993.
- Nilsen, H., M. Otterlei, T. Haug, K. Solum, T.A. Nagelhus, F. Skorpen, and H.E. Krokan. 1997. Nuclear and mitochondrial uracil-DNA glycosylases are generated by alternative splicing and transcription from different positions in the UNG gene. *Nucleic Acids Res* 25:750-755.
- Nilsen, H., G. Stamp, S. Andersen, G. Hrivnak, H.E. Krokan, T. Lindahl, and D.E. Barnes. 2003. Gene-targeted mice lacking the Ung uracil-DNA glycosylase develop B-cell lymphomas. *Oncogene* 22:5381-5386.
- Nishimura, K., T. Fukagawa, H. Takisawa, T. Kakimoto, and M. Kanemaki. 2009. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat Methods* 6:917-922.
- Nonaka, T., T. Doi, T. Toyoshima, M. Muramatsu, T. Honjo, and K. Kinoshita. 2009. Carboxy-terminal domain of AID required for its mRNA complex formation in vivo. 106:2747-2751.
- Noordermeer, S.M., S. Adam, D. Setiaputra, M. Barazas, S.J. Pettitt, A.K. Ling, M. Olivieri, A. Álvarez-Quilón, N. Moatti, M. Zimmermann, S. Annunziato, D.B. Krastev, F. Song, I. Brandsma, J. Frankum, R. Brough, A. Sherker, S. Landry, R.K. Szilard, M.M. Munro, A. McEwan, T. Gouillet de Ruyg, Z.-Y. Lin, T. Hart, J. Moffat, A.-C. Gingras, A. Martin, H. van Attikum, J. Jonkers, C.J. Lord, S. Rottenberg, and D. Durocher. 2018. The shieldin complex mediates 53BP1-dependent DNA repair. *Nature*

- Nowak, U., A.J. Matthews, S. Zheng, and J. Chaudhuri. 2011. The splicing regulator PTBP2 interacts with the cytidine deaminase AID and promotes binding of AID to switch-region DNA. *Nat Immunol* 12:160-166.
- Nussenzweig, A., and M.C. Nussenzweig. 2010. Origin of Chromosomal Translocations in Lymphoid Cancer. *Cell* 141:27-38.
- Nutt, S.L., P.D. Hodgkin, D.M. Tarlinton, and L.M. Corcoran. 2015. The generation of antibody-secreting plasma cells. *Nat Rev Immunol* 15:160-171.
- Oettgen, H.C. 2000. Regulation of the IgE isotype switch: new insights on cytokine signals and the functions of  $\epsilon$  germline transcripts. *Curr opin in Immunology* 12:618-623.
- Oettinger, M.A., D.G. Schatz, C. Gorka, and D. Baltimore. 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 248:1517-1523.
- Okazaki, I.-M., H. Hiai, N. Kakazu, S. Yamada, M. Muramatsu, K. Kinoshita, and T. Honjo. 2003a. Constitutive Expression of AID Leads to Tumorigenesis. *The Journal of Experimental Medicine* 197:1173-1181.
- Okazaki, I.-M., K. Kinoshita, M. Muramatsu, K. Yoshikawa, and T. Honjo. 2002. The AID enzyme induces class switch recombination in fibroblasts. *416:340-345*.
- Okazaki, I.M., H. Hiai, N. Kakazu, S. Yamada, M. Muramatsu, K. Kinoshita, and T. Honjo. 2003b. Constitutive expression of AID leads to tumorigenesis. *J Exp Med* 197:1173-1181.
- Okazaki, I.M., K. Okawa, M. Kobayashi, K. Yoshikawa, S. Kawamoto, H. Nagaoka, R. Shinkura, Y. Kitawaki, H. Taniguchi, T. Natsume, S.I. Iemura, and T. Honjo. 2011. Histone chaperone Spt6 is required for class switch recombination but not somatic hypermutation. *108:7920-7925*.
- Orthwein, A., A.M. Patenaude, B. Affar el, A. Lamarre, J.C. Young, and J.M. Di Noia. 2010. Regulation of activation-induced deaminase stability and antibody gene diversification by Hsp90. *J Exp Med* 207:2751-2765.
- Orthwein, A., A. Zahn, S.P. Methot, D. Godin, S.G. Conticello, K. Terada, and J.M. Di Noia. 2012. Optimal functional levels of activation-induced deaminase specifically require the Hsp40 DnaJa1. *EMBO J* 31:679-691.
- Oshiumi, H., T. Kouwaki, and T. Seya. 2016. Accessory Factors of Cytoplasmic Viral RNA Sensors Required for Antiviral Innate Immune Response. *Front Immunol* 7:200.
- Otterlei, M., T. Haug, T.A. Nagelhus, G. Slupphaug, T. Lindmo, and H.E. Krokan. 1998. Nuclear and mitochondrial splice forms of human uracil-DNA glycosylase contain a complex nuclear localisation signal and a strong classical mitochondrial localisation signal, respectively. *Nucleic Acids Research* 26:4611-4617.
- Otterlei, M., E. Warbrick, T.A. Nagelhus, T. Haug, G. Slupphaug, M. Akbari, P.A. Aas, K. Steinsbekk, O. Bakke, and H.E. Krokan. 1999. Post-replicative base excision repair in replication foci. *EMBO J* 18:3834-3844.
- Pan-Hammarstrom, Q., S. Dai, Y. Zhao, I.F. Van Dijk-Hard, R.A. Gatti, A.L. Borresen-Dale, and L. Hammarstrom. 2003. ATM Is Not Required in Somatic Hypermutation of VH, but Is Involved in the Introduction of Mutations in the Switch Region. *170:3707-3716*.
- Pan-Hammarström, Q., A.-M. Jones, A. Lähdesmäki, W. Zhou, R.A. Gatti, L. Hammarström, A.R. Gennery, and M.R. Ehrenstein. 2005. Impact of DNA ligase IV on nonhomologous end joining pathways during class switch recombination in human cells. *201:189-194*.
- Pan, Q., C. Petit-Frère, A. Lähdesmäki, H. Gregorek, K.H. Chrzanowska, and L. Hammarström. 2002. Alternative end joining during switch recombination in patients with Ataxia-Telangiectasia. *Eur J of Immunology* 32:1300-1308.
- Pannunzio, N.R., G. Watanabe, and M.R. Lieber. 2018. Nonhomologous DNA end-joining for repair of DNA double-strand breaks. *Journal of Biological Chemistry* 293:10512-10523.
- Pape, K.A., V. Kouskoff, D. Nemazee, H.L. Tang, J.G. Cyster, L.E. Tze, K.L. Hippen, T.W. Behrens, and M.K. Jenkins. 2003. Visualization of the Genesis and Fate of Isotype-switched B Cells during a Primary Immune Response. *197:1677-1687*.

- Park, S.R., P.H. Kim, K.S. Lee, S.H. Lee, G.Y. Seo, Y.C. Yoo, J. Lee, and P. Casali. 2013. APRIL stimulates NF-kappaB-mediated HoxC4 induction for AID expression in mouse B cells. *Cytokine* 61:608-613.
- Park, S.R., H. Zan, Z. Pal, J. Zhang, A. Al-Qahtani, E.J. Pone, Z. Xu, T. Mai, and P. Casali. 2009. HoxC4 binds to the promoter of the cytidine deaminase AID gene to induce AID expression, class-switch DNA recombination and somatic hypermutation. *Nat Immunol* 10:540-550.
- Parker, D.C. 1993a. The functions of antigen recognition in T cell-dependent B cell activation. *Semin Immunol* 5:413-420.
- Parker, D.C. 1993b. T cell-dependent B cell activation. *Annu Rev Immunol* 11:331-360.
- Parsa, J.Y., S. Ramachandran, A. Zaheen, R.M. Nepal, A. Kapelnikov, A. Belcheva, M. Berru, D. Ronai, and A. Martin. 2012. Negative supercoiling creates single-stranded patches of DNA that are substrates for AID-mediated mutagenesis. *PLoS Genet* 8:e1002518.
- Parsyan, A., D. Shahbazian, Y. Martineau, E. Petroulakis, T. Alain, O. Larsson, G. Mathonnet, G. Tettweiler, C.U. Hellen, T.V. Pestova, Y.V. Svitkin, and N. Sonenberg. 2009. The helicase protein DHX29 promotes translation initiation, cell proliferation, and tumorigenesis. *PLoS Biol* 106:22217-22222.
- Pasqualucci, L., G. Bhagat, M. Jankovic, M. Compagno, P. Smith, M. Muramatsu, T. Honjo, H.C. Morse, 3rd, M.C. Nussenzweig, and R. Dalla-Favera. 2008. AID is required for germinal center-derived lymphomagenesis. *Nat Genet* 40:108-112.
- Pasqualucci, L., R. Guglielmino, J. Houldsworth, J. Mohr, S. Aoufouchi, R. Polakiewicz, R.S. Chaganti, and R. Dalla-Favera. 2004. Expression of the AID protein in normal and neoplastic B cells. *Blood* 104:3318-3325.
- Pasqualucci, L., Y. Kitaura, H. Gu, and R. Dalla-Favera. 2006. PKA-mediated phosphorylation regulates the function of activation-induced deaminase (AID) in B cells. *PLoS Biol* 103:395-400.
- Pasqualucci, L., A. Migliazza, N. Fracchiolla, C. William, A. Neri, L. Baldini, R.S.K. Chaganti, U. Klein, R. Küppers, K. Rajewsky, and R. Dalla-Favera. 1998. BCL-6 mutations in normal germinal center B cells: Evidence of somatic hypermutation acting outside Ig loci. *Proceedings of the National Academy of Sciences* 95:11816-11821.
- Pasqualucci, L., P. Neumeister, T. Goossens, G. Nanjangud, R.S. Chaganti, R. Küppers, and R. Dalla-Favera. 2001. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* 412:341-346.
- Patenaude, A.M., and J.M. Di Noia. 2010. The mechanisms regulating the subcellular localization of AID. *Nucleus* 1:325-331.
- Patenaude, A.M., A. Orthwein, Y. Hu, V.A. Campo, B. Kavli, A. Buschiazzi, and J.M. Di Noia. 2009. Active nuclear import and cytoplasmic retention of activation-induced deaminase. *Nat Struct Mol Biol* 16:517-527.
- Pauklin, S., and S.K. Petersen-Mahrt. 2009. Progesterone Inhibits Activation-Induced Deaminase by Binding to the Promoter. *PLoS Biol* 183:1238-1244.
- Pauklin, S., I.V. Sernandez, G. Bachmann, A.R. Ramiro, and S.K. Petersen-Mahrt. 2009. Estrogen directly activates AID transcription and function. *J Exp Med* 206:99-111.
- Paul, K., M. Wang, E. Mladenov, A. Bencsik-Theilen, T. Bednar, W. Wu, H. Arakawa, and G. Iliakis. 2013. DNA Ligases I and III Cooperate in Alternative Non-Homologous End-Joining in Vertebrates. *PLoS ONE* 8:e59505.
- Pavri, R. 2017. R Loops in the Regulation of Antibody Gene Diversification. *Genes (Basel)* 8:154.
- Pavri, R., A. Gazumyan, M. Jankovic, M. Di Virgilio, I. Klein, C. Ansarah-Sobrinho, W. Resch, A. Yamane, B. Reina San-Martin, V. Barreto, T.J. Nieland, D.E. Root, R. Casellas, and M.C. Nussenzweig. 2010. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* 143:122-133.
- Pefanis, E., J. Wang, G. Rothschild, J. Lim, J. Chao, R. Rabadan, A.N. Economides, and U. Basu. 2014. Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature* 514:389-393.

- Peña-Díaz, J., S. Bregenhorn, M. Ghodgaonkar, C. Follonier, M. Artola-Borán, D. Castor, M. Lopes, Alessandro, and J. Jiricny. 2012. Noncanonical Mismatch Repair as a Source of Genomic Instability in Human Cells. *Molecular Cell* 47:669-680.
- Pérez-Durán, P., L. Belver, V.G. De Yébenes, P. Delgado, D.G. Pisano, and A.R. Ramiro. 2012. UNG shapes the specificity of AID-induced somatic hypermutation. 209:1379-1389.
- Péron, S., A. Metin, P. Gardès, M.-A. Alyanakian, E. Sheridan, C.P. Kratz, A. Fischer, and A. Durandy. 2008. Human PMS2 deficiency is associated with impaired immunoglobulin class switch recombination. 205:2465-2472.
- Peters, A., and U. Storb. 1996. Somatic Hypermutation of Immunoglobulin Genes Is Linked to Transcription Initiation. *Immunity* 4:57-65.
- Petersen-Mahrt, S.K., R.S. Harris, and M.S. Neuberger. 2002a. AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature* 418:99-103.
- Petersen-Mahrt, S.K., R.S. Harris, and M.S. Neuberger. 2002b. AID mutates *E. coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature* 418:99-104.
- Petersen, S., R. Casellas, B. Reina-San-Martin, H.T. Chen, M.J. Difilippantonio, P.C. Wilson, L. Hanitsch, A. Celeste, M. Muramatsuk, D.R. Pilch, C. Redon, T. Ried, W.M. Bonner, T. Honjo, M.C. Nussenzweig, and A. Nussenzweig. 2001. AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. *Nature* 414:660-665.
- Pham, P., S.A. Afif, M. Shimoda, K. Maeda, N. Sakaguchi, L.C. Pedersen, and M.F. Goodman. 2016. Structural analysis of the activation-induced deoxycytidine deaminase required in immunoglobulin diversification. *DNA Repair (Amst)* 43:48-56.
- Pham, P., R. Bransteitter, J. Petruska, and M.F. Goodman. 2003. Processive AID-catalysed cytosine deamination on single-stranded DNA simulates somatic hypermutation. *Nature* 424:103-107.
- Pham, P., S. Malik, C. Mak, P.C. Calabrese, R.G. Roeder, and M.F. Goodman. 2019. AID-RNA polymerase II transcription-dependent deamination of IgV DNA. *Nucleic Acids Res* 47:10815-10829.
- Pilzecker, B., and H. Jacobs. 2019. Mutating for Good: DNA Damage Responses During Somatic Hypermutation. *Frontiers in Immunology* 10:
- Pinaud, E., M. Marquet, R. Fiancette, S. Peron, C. Vincent-Fabert, Y. Denizot, and M. Cogne. 2011. The IgH locus 3' regulatory region: pulling the strings from behind. *Adv Immunol* 110:27-70.
- Pone, E.J., J. Zhang, T. Mai, C.A. White, G. Li, J.K. Sakakura, P.J. Patel, A. Al-Qahtani, H. Zan, Z. Xu, and P. Casali. 2012. BCR-signalling synergizes with TLR-signalling for induction of AID and immunoglobulin class-switching through the non-canonical NF-kappaB pathway. *Nat Commun* 3:767.
- Pramanik, S., A. Kutzner, and K. Heese. 2015. Lead discovery and in silico 3D structure modeling of tumorigenic FAM72A (p17). *Tumour Biol* 36:239-249.
- Qiao, Q., L. Wang, F.-L. Meng, J.K. Hwang, F.W. Alt, and H. Wu. 2017a. AID Recognizes Structured DNA for Class Switch Recombination. *Molecular Cell* 67:361-373.e364.
- Qiao, Q., L. Wang, F.L. Meng, J.K. Hwang, F.W. Alt, and H. Wu. 2017b. AID Recognizes Structured DNA for Class Switch Recombination. *Mol Cell* 67:361-373 e364.
- Quy H. Phung, David B. Winter, R. Alrefai, and P.J. Gearhart. 1999. Hypermutation in Ig V Genes from Mice Deficient in the MLH1 Mismatch Repair Protein. *the journal of Immunology* 162:3121-3124.
- R.Ehrenstein, M., and M. S.Neuberger. 1999. Deficiency in Msh2 affects the efficiency and local sequence specificity of immunoglobulin class-switch recombination: parallels with somatic hypermutation. *The EMBO Journal* 18:3484-3490.
- Rada, C., J.M. Di Noia, and M.S. Neuberger. 2004a. Mismatch recognition and uracil excision provide complementary paths to both Ig switching and the A/T-focused phase of somatic mutation. *Mol Cell* 16:163-171.
- Rada, C., J.M. Di Noia, and M.S. Neuberger. 2004b. Mismatch Recognition and Uracil Excision Provide Complementary Paths to Both Ig Switching and the A/T-Focused Phase of Somatic Mutation. *Molecular Cell* 16:163-171.



- Rada, C., J.M. Jarvis, and C. Milstein. 2002a. AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proc Natl Acad Sci U S A* 99:7003-7008.
- Rada, C., and C. Milstein. 2001. The intrinsic hypermutability of antibody heavy and light chain genes decays exponentially. *EMBO J* 20:4570-4576.
- Rada, C., G.T. Williams, H. Nilsen, D.E. Barnes, T. Lindahl, and M.S. Neuberger. 2002b. Immunoglobulin Isotype Switching Is Inhibited and Somatic Hypermutation Perturbed in UNG-Deficient Mice. *Current Biology* 12:1748-1755.
- Rada, C., G.T. Williams, H. Nilsen, D.E. Barnes, T. Lindahl, and M.S. Neuberger. 2002c. Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in UNG-deficient mice. *Curr Biol* 12:1748-1755.
- Rahane, C.S., A. Kutzner, and K. Heese. 2019. A cancer tissue-specific FAM72 expression profile defines a novel glioblastoma multiform (GBM) gene-mutation signature. *J Neurooncol* 141:57-70.
- Rajagopal, D., R.W. Maul, A. Ghosh, T. Chakraborty, A.A. Khamlichi, R. Sen, and P.J. Gearhart. 2009. Immunoglobulin switch mu sequence causes RNA polymerase II accumulation and reduces dA hypermutation. *J Exp Med* 206:1237-1244.
- Rajewsky, K. 1996. Clonal selection and learning in the antibody system. *Nature* 381:751.
- Rajewsky, K., I. Forster, and A. Cumano. 1987. Evolutionary and somatic selection of the antibody repertoire in the mouse. *Science* 238:1088-1094.
- Ramachandran, S., R. Chahwan, R.M. Nepal, D. Frieder, S. Panier, S. Roa, A. Zaheen, D. Durocher, M.D. Scharff, and A. Martin. 2010. The RNF8/RNF168 ubiquitin ligase cascade facilitates class switch recombination. *107:809-814*.
- Ramiro, A.R., M. Jankovic, E. Callen, S. Difilippantonio, H.-T. Chen, K.M. McBride, T.R. Eisenreich, J. Chen, R.A. Dickins, S.W. Lowe, A. Nussenzweig, and M.C. Nussenzweig. 2006. Role of genomic instability and p53 in AID-induced c-myc-IgH translocations. *Nature* 440:105-109.
- Ramiro, A.R., M. Jankovic, T. Eisenreich, S. Difilippantonio, S. Chen-Kiang, M. Muramatsu, T. Honjo, A. Nussenzweig, and M.C. Nussenzweig. 2004. AID Is Required for c-myc/IgH Chromosome Translocations In Vivo. *Cell* 118:431-438.
- Ramiro, A.R., P. Stavropoulos, M. Jankovic, and M.C. Nussenzweig. 2003. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nat Immunol* 4:452-456.
- Ranjit, S., L. Khair, E.K. Linehan, A.J. Ucher, M. Chakrabarti, C.E. Schrader, and J. Stavnezer. 2011. AID Binds Cooperatively with UNG and Msh2-Msh6 to Ig Switch Regions Dependent upon the AID C Terminus. *The Journal of Immunology* 187:2464-2475.
- Rass, E., A. Grabarz, I. Plo, J. Gautier, P. Bertrand, and B.S. Lopez. 2009. Role of Mre11 in chromosomal nonhomologous end joining in mammalian cells. *16:819-824*.
- Ratmeyer, L., R. Vinayak, Y.Y. Zhong, G. Zon, and W.D. Wilson. 1994. Sequence specific thermodynamic and structural properties for DNA:RNA duplexes. *Biochemistry* 33:5298-5304.
- Reaban, M.E., and J.A. Griffin. 1990. Induction of RNA-stabilized DNA conformers by transcription of an immunoglobulin switch region. *Nature* 348:342-344.
- Recaldin, T., and D.J. Fear. 2016. Transcription factors regulating B cell fate in the germinal centre. *183:65-75*.
- Recaldin, T., P.S. Hobson, E.H. Mann, F. Ramadani, D.J. Cousins, P. Lavender, and D.J. Fear. 2018. miR-29b directly targets activation-induced cytidine deaminase in human B cells and can limit its inappropriate expression in naive B cells. *Mol Immunol* 101:419-428.
- Reina-San-Martin, B., H.T. Chen, A. Nussenzweig, and M.C. Nussenzweig. 2004. ATM Is Required for Efficient Recombination between Immunoglobulin Switch Regions. *200:1103-1110*.
- Reina-San-Martin, B., J. Chen, A. Nussenzweig, and M.C. Nussenzweig. 2007. Enhanced intra-switch region recombination during immunoglobulin class switch recombination in 53BP1-/- B cells. *European Journal of Immunology* 37:235-239.

- Reina-San-Martin, B., S. Difilippantonio, L. Hanitsch, R.F. Masilamani, A. Nussenzweig, and M.C. Nussenzweig. 2003. H2AX Is Required for Recombination Between Immunoglobulin Switch Regions but Not for Intra-Switch Region Recombination or Somatic Hypermutation. *Journal of Experimental Medicine* 197:1767-1778.
- Reina-San-Martin, B., M.C. Nussenzweig, A. Nussenzweig, and S. Difilippantonio. 2005. Genomic instability, endoreduplication, and diminished Ig class-switch recombination in B cells lacking Nbs1. *102:1590-1595*.
- Revilla-I-Domingo, R., I. Bilic, B. Vilagos, H. Tagoh, A. Ebert, I.M. Tamir, L. Smeenk, J. Trupke, A. Sommer, M. Jaritz, and M. Busslinger. 2012. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *The EMBO Journal* 31:3130-3146.
- Revy, P., T. Muto, Y. Levy, F. Geissmann, A. Plebani, O. Sanal, N. Catalan, M. Forveille, R. Dufourcq-Labelouse, A. Gennery, I. Tezcan, F. Ersoy, H. Kayserili, A.G. Ugazio, N. Brousse, M. Muramatsu, L.D. Notarangelo, K. Kinoshita, T. Honjo, A. Fischer, and A. Durandy. 2000. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell* 102:565-575.
- Reynaud, C. 1985. A single rearrangement event generates most of the chicken immunoglobulin light chain diversity. *Cell* 40:283-291.
- Reynaud, C.A., V. Anquez, H. Grimal, and J.C. Weill. 1987. A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell* 48:379-388.
- Ribeiro De Almeida, C., S. Dhir, A. Dhir, A.E. Moghaddam, Q. Sattentau, A. Meinhart, and N.J. Proudfoot. 2018a. RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination. *Molecular Cell* 70:650-662.e658.
- Ribeiro de Almeida, C., S. Dhir, A. Dhir, A.E. Moghaddam, Q. Sattentau, A. Meinhart, and N.J. Proudfoot. 2018b. RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination. *Mol Cell* 70:650-662 e658.
- Roa, S., E. Avdievich, J.U. Peled, T. Maccarthy, U. Werling, F.L. Kuang, R. Kan, C. Zhao, A. Bergman, P.E. Cohen, W. Edelmann, and M.D. Scharff. 2008. Ubiquitylated PCNA plays a role in somatic hypermutation and class-switch recombination and is required for meiotic progression. *105:16248-16253*.
- Robbiani, D.F., A. Bothmer, E. Callen, B. Reina-San-Martin, Y. Dorsett, S. Difilippantonio, D.J. Bolland, H.T. Chen, A.E. Corcoran, A. Nussenzweig, and M.C. Nussenzweig. 2008. AID Is Required for the Chromosomal Breaks in c-myc that Lead to c-myc/IgH Translocations. *Cell* 135:1028-1038.
- Robbiani, D.F., S. Bunting, N. Feldhahn, A. Bothmer, J. Camps, S. Deroubaix, K.M. McBride, I.A. Klein, G. Stone, T.R. Eisenreich, T. Ried, A. Nussenzweig, and M.C. Nussenzweig. 2009. AID produces DNA double-strand breaks in non-Ig genes and mature B cell lymphomas with reciprocal chromosome translocations. *Mol Cell* 36:631-641.
- Robert, I., F. Dantzer, and B. Reina-San-Martin. 2009. Parp1 facilitates alternative NHEJ, whereas Parp2 suppresses IgH/c-myc translocations during immunoglobulin class switch recombination. *206:1047-1056*.
- Roberts, R.W., and D.M. Crothers. 1992. Stability and properties of double and triple helices: dramatic effects of RNA or DNA backbone composition. *Science* 258:1463-1466.
- Rocha, P.P., and J.A. Skok. 2013. The origin of recurrent translocations in recombining lymphocytes: a balance between break frequency and nuclear proximity. *Curr Opin Cell Biol* 25:365-371.
- Roco, J.A., L. Mesin, S.C. Binder, C. Nefzger, P. Gonzalez-Figueroa, P.F. Canete, J. Ellyard, Q. Shen, P.A. Robert, J. Cappello, H. Vohra, Y. Zhang, C.R. Nowosad, A. Schiepers, L.M. Corcoran, K.M. Toellner, J.M. Polo, M. Meyer-Hermann, G.D. Victoria, and C.G. Vinuesa. 2019. Class-Switch Recombination Occurs Infrequently in Germinal Centers. *Immunity* 51:337-350 e337.
- Rodriguez, G., A. Esadze, B.P. Weiser, J.D. Schonhoft, P.A. Cole, and J.T. Stivers. 2017. Disordered N-Terminal Domain of Human Uracil DNA Glycosylase (hUNG2) Enhances DNA Translocation. *ACS Chem Biol* 12:2260-2263.

- Rogozin, I.B., and M. Diaz. 2004. Cutting Edge: DGYW/WRCH Is a Better Predictor of Mutability at G:C Bases in Ig Hypermutation Than the Widely Accepted RGYW/WRCY Motif and Probably Reflects a Two-Step Activation-Induced Cytidine Deaminase-Triggered Process. *The Journal of Immunology* 172:3382-3384.
- Rogozin, I.B., and N.A. Kolchanov. 1992a. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochim Biophys Acta* 1171:11-18.
- Rogozin, I.B., and N.A. Kolchanov. 1992b. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochimica et Biophysica Acta* 1171:11-18.
- Rogozin, I.B., Y.I. Pavlov, K. Bebenek, T. Matsuda, and T.A. Kunkel. 2001. Somatic mutation hotspots correlate with DNA polymerase  $\eta$  error spectrum. *Nature Immunology* 2:530-536.
- Romanello, M., D. Schiavone, A. Frey, and J.E. Sale. 2016. Histone H3.3 promotes IgV gene diversification by enhancing formation of AID-accessible single-stranded DNA. *EMBO J* 35:1452-1464.
- Ronai, D., M.D. Iglesias-Ussel, M. Fan, Z. Li, A. Martin, and M.D. Scharff. 2007. Detection of chromatin-associated single-stranded DNA in regions targeted for somatic hypermutation. *J Exp Med* 204:181-190.
- Rouaud, P., C. Vincent-Fabert, A. Saintamand, R. Fiancette, M. Marquet, I. Robert, B. Reina-San-Martin, E. Pinaud, M. Cogne, and Y. Denizot. 2013. The IgH 3' regulatory region controls somatic hypermutation in germinal center B cells. *J Exp Med* 210:1501-1507.
- Roux, K.J., D.I. Kim, M. Raida, and B. Burke. 2012. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* 196:801-810.
- Roy, D., and M.R. Lieber. 2009. G Clustering Is Important for the Initiation of Transcription-Induced R-Loops In Vitro, whereas High G Density without Clustering Is Sufficient Thereafter. 29:3124-3133.
- Roy, D., Z. Zhang, Z. Lu, C.L. Hsieh, and M.R. Lieber. 2010. Competition between the RNA Transcript and the Nontemplate DNA Strand during R-Loop Formation In Vitro: a Nick Can Serve as a Strong R-Loop Initiation Site. 30:146-159.
- Safavi, S., A. Larouche, A. Zahn, A.-M. Patenaude, D. Domanska, K. Dionne, T. Rognes, F. Dingler, S.-K. Kang, Y. Liu, N. Johnson, J. Hébert, R.E. Verdun, C.A. Rada, F. Vega, H. Nilsen, and M. Javier. 2020. The uracil-DNA glycosylase UNG protects the fitness of normal and cancer B cells expressing AID. *NAR Cancer* 2:
- Saintamand, A., P. Rouaud, F. Saad, G. Rios, M. Cogne, and Y. Denizot. 2015. Elucidation of IgH 3' region regulatory role during class switch recombination via germline deletion. *Nat Commun* 6:7084.
- Sakano, H., Y. Kurosawa, M. Weigert, and S. Tonegawa. 1981. Identification and nucleotide sequence of a diversity DNA segment (D) of immunoglobulin heavy-chain genes. *Nature* 290:562-565.
- Santos, M.A., M.S.Y. Huen, M. Jankovic, H.-T. Chen, A.J. López-Contreras, I.A. Klein, N. Wong, J.L.R. Barbancho, O. Fernandez-Capetillo, M.C. Nussenzweig, J. Chen, and A. Nussenzweig. 2010. Class switching and meiotic defects in mice lacking the E3 ubiquitin ligase RNF8. 207:973-981.
- Sarno, A., M. Lundbæk, N.B. Liabakk, P.A. Aas, R. Mjelle, L. Hagen, M.M.L. Sousa, H.E. Krokan, and B. Kavli. 2019a. Uracil-DNA glycosylase UNG1 isoform variant supports class switch recombination and repairs nuclear genomic uracil. *Nucleic Acids Research* 47:4569-4585.
- Sarno, A., M. Lundbæk, N.B. Liabakk, P.A. Aas, R. Mjelle, L. Hagen, M.M.L. Sousa, H.E. Krokan, and B. Kavli. 2019b. Uracil-DNA glycosylase UNG1 isoform variant supports class switch recombination and repairs nuclear genomic uracil. *Nucleic Acids Res* 47:4569-4585.
- Saunders, A., L.J. Core, and J.T. Lis. 2006. Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol* 7:557-567.
- Sawai, Y., Y. Kodama, T. Shimizu, Y. Ota, T. Maruno, Y. Eso, A. Kurita, M. Shiokawa, Y. Tsuji, N. Uza, Y. Matsumoto, T. Masui, S. Uemoto, H. Marusawa, and T. Chiba. 2015. Activation-Induced

- Cytidine Deaminase Contributes to Pancreatic Tumorigenesis by Inducing Tumor-Related Gene Mutations. *75*:3292-3301.
- Sayegh, C.E., M.W. Quong, Y. Agata, and C. Murre. 2003. E-proteins directly regulate expression of activation-induced deaminase in mature B cells. *Nat Immunol* 4:586-593.
- Schanz, S., D. Castor, F. Fischer, and J. Jiricny. 2009. Interference of mismatch and base excision repair during the processing of adjacent U/G mispairs may play a key role in somatic hypermutation. *106*:5593-5598.
- Schatz, D.G., M.A. Oettinger, and D. Baltimore. 1989. The V(D)J recombination activating gene, RAG-1. *Cell* 59:1035-1048.
- Schrader, C.E., W. Edelmann, R. Kucherlapati, and J. Stavnezer. 1999. Reduced Isotype Switching in Splenic B Cells from Mice Deficient in Mismatch Repair Enzymes. *190*:323-330.
- Schrader, C.E., J.E. Guikema, E.K. Linehan, E. Selsing, and J. Stavnezer. 2007a. Activation-induced cytidine deaminase-dependent DNA breaks in class switch recombination occur during G1 phase of the cell cycle and depend upon mismatch repair. *J Immunol* 179:6064-6071.
- Schrader, C.E., J.E.J. Guikema, E.K. Linehan, E. Selsing, and J. Stavnezer. 2007b. Activation-Induced Cytidine Deaminase-Dependent DNA Breaks in Class Switch Recombination Occur during G1 Phase of the Cell Cycle and Depend upon Mismatch Repair. *The Journal of Immunology* 179:6064-6071.
- Schrader, C.E., E.K. Linehan, S.N. Mochevova, R.T. Woodland, and J. Stavnezer. 2005a. Inducible DNA breaks in Ig S regions are dependent on AID and UNG. *J Exp Med* 202:561-568.
- Schrader, C.E., E.K. Linehan, S.N. Mochevova, R.T. Woodland, and J. Stavnezer. 2005b. Inducible DNA breaks in Ig S regions are dependent on AID and UNG. *Journal of Experimental Medicine* 202:561-568.
- Sen, D., and W. Gilbert. 1988. Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature* 334:364-366.
- Senigl, F., Y. Maman, R.K. Dinesh, J. Alinikula, R.B. Seth, L. Pecnova, A.D. Omer, S.S.P. Rao, D. Weisz, J.-M. Buerstedde, E.L. Aiden, R. Casellas, J. Hejnar, and D.G. Schatz. 2019. Topologically Associated Domains Delineate Susceptibility to Somatic Hypermutation. *Cell Reports* 29:3902-3915.e3908.
- Serebrenik, A.A., G.J. Starrett, S. Leenen, M.C. Jarvis, N.M. Shaban, D.J. Salamango, H. Nilsen, W.L. Brown, and R.S. Harris. 2019. The deaminase APOBEC3B triggers the death of cells lacking uracil DNA glycosylase. *Proc Natl Acad Sci U S A* 116:22158-22163.
- Sernandez, I.V., V.G. de Yebenes, Y. Dorsett, and A.R. Ramiro. 2008. Haploinsufficiency of activation-induced deaminase for antibody diversification and chromosome translocations both in vitro and in vivo. *PLoS One* 3:e3927.
- Shalhout, S., D. Haddad, A. Sosin, T.C. Holland, A. Al-Katib, A. Martin, and A.S. Bhagwat. 2014. Genomic Uracil Homeostasis during Normal B Cell Maturation and Loss of This Balance during B Cell Cancer Development. *Molecular and Cellular Biology* 34:4019-4032.
- Shanbhag, N.M., I.U. Rafalska-Metcalf, C. Balane-Bolivar, S.M. Janicki, and R.A. Greenberg. 2010. ATM-Dependent Chromatin Changes Silence Transcription In cis to DNA Double-Strand Breaks. *Cell* 141:970-981.
- Sharbeen, G., C.W. Yee, A.L. Smith, and C.J. Jolly. 2012. Ectopic restriction of DNA repair reveals that UNG2 excises AID-induced uracils predominantly or exclusively during G1 phase. *J Exp Med* 209:965-974.
- Sharpe, M.J., C. Milstein, J.M. Jarvis, and M.S. Neuberger. 1991. Somatic hypermutation of immunoglobulin kappa may depend on sequences 3' of C kappa and occurs on passenger transgenes. *EMBO J* 10:2139-2145.
- Shen, H.M., A. Peters, B. Baron, X. Zhu, and U. Storb. 1998. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science* 280:1750-1752.
- Shinkura, R., S. Ito, N.A. Begum, H. Nagaoka, M. Muramatsu, K. Kinoshita, Y. Sakakibara, H. Hijikata, and T. Honjo. 2004. Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nature Immunology* 5:707-712.

- Shinkura, R., F. Matsuda, T. Sakiyama, T. Tsubata, H. Hial, M. Paumen, S. Miyawaki, and T. Honjo. 1996. Defects of somatic hypermutation and class switching in alymphoplasia ( aly ) mutant mice. *8:1067-1075*.
- Shinkura, R., M. Tian, M. Smith, K. Chua, Y. Fujiwara, and F.W. Alt. 2003. The influence of transcriptional orientation on endogenous switch region function. *Nat Immunol* 4:435-441.
- Shukla, V., L. Halabelian, S. Balagere, D. Samaniego-Castruita, D.E. Feldman, C.H. Arrowsmith, A. Rao, and L. Aravind. 2020. HMCES Functions in the Alternative End-Joining Pathway of the DNA DSB Repair during Class Switch Recombination in B Cells. *Mol Cell* 77:384-394 e384.
- Simsek, D., E. Brunet, S.Y.-W. Wong, S. Katyal, Y. Gao, P.J. McKinnon, J. Lou, L. Zhang, J. Li, E.J. Rebar, P.D. Gregory, M.C. Holmes, and M. Jasin. 2011. DNA Ligase III Promotes Alternative Nonhomologous End-Joining during Chromosomal Translocation Formation. *PLoS Genetics* 7:e1002080.
- Singh, A.K., A. Tamrakar, A. Jaiswal, N. Kanayama, and P. Kodgire. 2020. SRSF1-3, a splicing and somatic hypermutation regulator, controls transcription of IgV genes via chromatin regulators SATB2, UBN1 and histone variant H3.3. *Mol Immunol* 119:69-82.
- Singh, S.K., K. Maeda, M.M. Eid, S.A. Almofty, M. Ono, P. Pham, M.F. Goodman, and N. Sakaguchi. 2013. GANP regulates recruitment of AID to immunoglobulin variable regions by modulating transcription and nucleosome occupancy. *Nat Commun* 4:1830.
- Soulas-Sprauel, P., G.L. Le Guyader, P. Rivera-Munoz, V. Abramowski, C. Olivier-Martin, C.C. Goujet-Zalc, P. Charneau, and J.-P. De Villartay. 2007. Role for DNA repair factor XRCC4 in immunoglobulin class switch recombination. *Journal of Experimental Medicine* 204:1717-1727.
- Stanlie, A., M. Aida, M. Muramatsu, T. Honjo, and N.A. Begum. 2010. Histone3 lysine4 trimethylation regulated by the facilitates chromatin transcription complex is critical for DNA cleavage in class switch recombination. *Proc Natl Acad Sci U S A* 107:22190-22195.
- Stavnezer, J., J.E. Guikema, and C.E. Schrader. 2008. Mechanism and regulation of class switch recombination. *Annu Rev Immunol* 26:261-292.
- Stephens, P.J., C.D. Greenman, B. Fu, F. Yang, G.R. Bignell, L.J. Mudie, E.D. Pleasance, K.W. Lau, D. Beare, L.A. Stebbings, S. McLaren, M.-L. Lin, D.J. McBride, I. Varela, S. Nik-Zainal, C. Leroy, M. Jia, A. Menzies, A.P. Butler, J.W. Teague, M.A. Quail, J. Burton, H. Swerdlow, N.P. Carter, L.A. Morsberger, C. Iacobuzio-Donahue, G.A. Follows, A.R. Green, A.M. Flanagan, M.R. Stratton, P.A. Futreal, and P.J. Campbell. 2011. Massive Genomic Rearrangement Acquired in a Single Catastrophic Event during Cancer Development. *Cell* 144:27-40.
- Stratigopoulou, M., T.P. Van Dam, and J.E.J. Guikema. 2020. Base Excision Repair in the Immune System: Small DNA Lesions With Big Consequences. *Frontiers in Immunology* 11:
- Stucki, M., J.A. Clapperton, D. Mohammad, M.B. Yaffe, S.J. Smerdon, and S.P. Jackson. 2005. MDC1 Directly Binds Phosphorylated Histone H2AX to Regulate Cellular Responses to DNA Double-Strand Breaks. *Cell* 123:1213-1226.
- Sugimoto, N., H. Mitoma, T. Kim, S. Hanabuchi, and Y.J. Liu. 2014. Helicase proteins DHX29 and RIG-I cosense cytosolic nucleic acids in the human airway system. *Proceedings of the National Academy of Sciences* 111:7747-7752.
- Sun, J., C.D. Keim, J. Wang, D. Kazadi, P.M. Oliver, R. Rabadan, and U. Basu. 2013a. E3-ubiquitin ligase Nedd4 determines the fate of AID-associated RNA polymerase II in B cells. *Genes & Development* 27:1821-1833.
- Sun, J., G. Rothschild, E. Pefanis, and U. Basu. 2013b. Transcriptional stalling in B-lymphocytes: a mechanism for antibody diversification and maintenance of genomic integrity. *Transcription* 4:127-135.
- Svilar, D., E.M. Goellner, K.H. Almeida, and R.W. Sobol. 2011. Base excision repair and lesion-dependent subpathways for repair of oxidative DNA damage. *Antioxid Redox Signal* 14:2491-2507.
- Swaminathan, S., L. Klemm, E. Park, E. Papaemmanuil, A. Ford, S.M. Kweon, D. Trageser, B. Hasselfeld, N. Henke, J. Mooster, H. Geng, K. Schwarz, S.C. Kogan, R. Casellas, D.G. Schatz,

- M.R. Lieber, M.F. Greaves, and M. Muschen. 2015. Mechanisms of clonal evolution in childhood acute lymphoblastic leukemia. *Nat Immunol* 16:766-774.
- Szymczak, A.L., and D.A. Vignali. 2005. Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin Biol Ther* 5:627-638.
- Ta, V.-T., H. Nagaoka, N. Catalan, A. Durandy, A. Fischer, K. Imai, S. Nonoyama, J. Tashiro, M. Ikegawa, S. Ito, K. Kinoshita, M. Muramatsu, and T. Honjo. 2003a. AID mutant analyses indicate requirement for class-switch-specific cofactors. 4:843-848.
- Ta, V.T., H. Nagaoka, N. Catalan, A. Durandy, A. Fischer, K. Imai, S. Nonoyama, J. Tashiro, M. Ikegawa, S. Ito, K. Kinoshita, M. Muramatsu, and T. Honjo. 2003b. AID mutant analyses indicate requirement for class-switch-specific cofactors. *Nat Immunol* 4:843-848.
- Takizawa, M., H. Tolarova, Z. Li, W. Dubois, S. Lim, E. Callen, S. Franco, M. Mosaico, L. Feigenbaum, F.W. Alt, A. Nussenzweig, M. Potter, and R. Casellas. 2008. AID expression levels determine the extent of cMyc oncogenic translocations and the incidence of B cell tumor development. *J Exp Med* 205:1949-1957.
- Tang, E.S., and A. Martin. 2007. Immunoglobulin gene conversion: Synthesizing antibody diversification and DNA repair. *DNA Repair* 6:1557-1571.
- Taylor, B.J., S. Nik-Zainal, Y.L. Wu, L.A. Stebbings, K. Raine, P.J. Campbell, C. Rada, M.R. Stratton, and M.S. Neuberger. 2013. DNA deaminases induce break-associated mutation showers with implication of APOBEC3B and 3A in breast cancer kataegis. *eLife* 2:
- Teng, G., P. Hakimpour, P. Landgraf, A. Rice, T. Tuschl, R. Casellas, and F.N. Papavasiliou. 2008. MicroRNA-155 is a negative regulator of activation-induced cytidine deaminase. *Immunity* 28:621-629.
- Thientosapol, E.S., G. Sharbeen, K.K.E. Lau, D. Bosnjak, T. Durack, I. Stevanovski, W. Weninger, and C.J. Jolly. 2017. Proximity to AGCT sequences dictates MMR-independent versus MMR-dependent mechanisms for AID-induced mutation via UNG2. *Nucleic Acids Res* 45:3146-3157.
- Thomas-Claudepierre, A.-S., I. Robert, P.P. Rocha, R. Raviram, E. Schiavo, V. Heyer, R. Bonneau, V.M. Luo, J.K. Reddy, T. Borggreffe, J.A. Skok, and B. Reina-San-Martin. 2016. Mediator facilitates transcriptional activation and dynamic long-range contacts at the IgH locus during class switch recombination. *jem*.20141967.
- Thomas-Claudepierre, A.S., E. Schiavo, V. Heyer, M. Fournier, A. Page, I. Robert, and B. Reina-San-Martin. 2013. The cohesin complex regulates immunoglobulin class switch recombination. *J Exp Med* 210:2495-2502.
- Toellner, K.-M., S.A. Luther, D.M.Y. Sze, R.K.W. Choy, D.R. Taylor, I.C.M. Maclennan, and H. Acha-Orbea. 1998. T Helper 1 (Th1) and Th2 Characteristics Start to Develop During T Cell Priming and Are Associated with an Immediate Ability to Induce Immunoglobulin Class Switching. *Journal of Experimental Medicine* 187:1193-1204.
- Toellner, K.M., A. Gulbranson-Judge, D.R. Taylor, D.M. Sze, and I.C. Maclennan. 1996. Immunoglobulin switch transcript production in vivo related to the site and time of antigen-specific B cell activation. *The Journal of Experimental Medicine* 183:2303-2312.
- Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature* 302:575-581.
- Torseth, K., B. Doseth, L. Hagen, C. Olaisen, N.B. Liabakk, H. Graesmann, A. Durandy, M. Otterlei, H.E. Krokan, B. Kavli, and G. Slupphaug. 2012. The UNG2 Arg88Cys variant abrogates RPA-mediated recruitment of UNG2 to single-stranded DNA. *DNA Repair (Amst)* 11:559-569.
- Tran, T.H., M. Nakata, K. Suzuki, N.A. Begum, R. Shinkura, S. Fagarasan, T. Honjo, and H. Nagaoka. 2010. B cell-specific and stimulation-responsive enhancers derepress Aicda by overcoming the effects of silencers. *Nat Immunol* 11:148-154.
- Trepel, F. 1974. Number and distribution of lymphocytes in man. A critical analysis. *Klin Wochenschr* 52:511-515.
- Tyanova, S., and J. Cox. 2018. Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. *Methods Mol Biol* 1711:133-148.
- Tyanova, S., T. Temu, and J. Cox. 2016a. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat Protoc* 11:2301-2319.

- Tyanova, S., T. Temu, P. Sinitcyn, A. Carlson, M.Y. Hein, T. Geiger, M. Mann, and J. Cox. 2016b. The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* 13:731-740.
- Tytgat, H.L., G. Schoofs, M. Driesen, P. Proost, E.J. Van Damme, J. Vanderleyden, and S. Lebeer. 2015. Endogenous biotin-binding proteins: an overlooked factor causing false positives in streptavidin-based protein detection. *Microb Biotechnol* 8:164-168.
- Uchimura, Y., L.F. Barton, C. Rada, and M.S. Neuberger. 2011. REG-gamma associates with and modulates the abundance of nuclear activation-induced deaminase. *J Exp Med* 208:2385-2391.
- Van Oers, J.M.M., S. Roa, U. Werling, Y. Liu, J. Genschel, H. Hou, R.S. Sellers, P. Modrich, M.D. Scharff, and W. Edelmann. 2010. PMS2 endonuclease activity has distinct biological functions and is essential for genome maintenance. 107:13384-13389.
- Varnaite, R., and S.A. MacNeill. 2016. Meet the neighbors: Mapping local protein interactomes by proximity-dependent labeling with BioID. *Proteomics* 16:2503-2518.
- Victoria, G.D., and M.C. Nussenzweig. 2012. Germinal Centers. *Annual Review of Immunology* 30:429-457.
- Victoria, G.D., T.A. Schwickert, D.R. Fooksman, A.O. Kamphorst, M. Meyer-Hermann, M.L. Dustin, and M.C. Nussenzweig. 2010. Germinal Center Dynamics Revealed by Multiphoton Microscopy with a Photoactivatable Fluorescent Reporter. *Cell* 143:592-605.
- Vincent-Fabert, C., R. Fiancette, E. Pinaud, V. Truffinet, N. Cogne, M. Cogne, and Y. Denizot. 2010. Genomic deletion of the whole IgH 3' regulatory region (hs3a, hs1,2, hs3b, and hs4) dramatically affects class switch recombination and Ig secretion to all isotypes. *Blood* 116:1895-1898.
- Vuong, B.Q., K. Herrick-Reynolds, B. Vaidyanathan, J.N. Pucella, A.J. Ucher, N.M. Donghia, X. Gu, L. Nicolas, U. Nowak, N. Rahman, M.P. Strout, K.D. Mills, J. Stavnezer, and J. Chaudhuri. 2013. A DNA break- and phosphorylation-dependent positive feedback loop promotes immunoglobulin class-switch recombination. *Nature Immunology* 14:1183-1189.
- Vuong, B.Q., M. Lee, S. Kabir, C. Irimia, S. Macchiarulo, G.S. McKnight, and J. Chaudhuri. 2009. Specific recruitment of protein kinase A to the immunoglobulin locus regulates class-switch recombination. 10:420-426.
- Wang, L., R. Wuerffel, S. Feldman, A.A. Khamlichi, and A.L. Kenter. 2009. S region sequence, RNA polymerase II, and histone modifications create chromatin accessibility during class switch recombination. *J Exp Med* 206:1817-1830.
- Wang, M., C. Rada, and M.S. Neuberger. 2010. Altering the spectrum of immunoglobulin V gene somatic hypermutation by modifying the active site of AID. 207:141-153.
- Wang, M., W. Wu, W. Wu, B. Rosidi, L. Zhang, H. Wang, and G. Iliakis. 2006. PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. 34:6170-6182.
- Wang, Q., K.R. Kieffer-Kwon, T.Y. Oliveira, C.T. Mayer, K. Yao, J. Pai, Z. Cao, M. Dose, R. Casellas, M. Jankovic, M.C. Nussenzweig, and D.F. Robbiani. 2017. The cell cycle restricts activation-induced cytidine deaminase activity to early G1. *J Exp Med* 214:49-58.
- Wang, X., M. Fan, S. Kalis, L. Wei, and M.D. Scharff. 2014. A source of the single-stranded DNA substrate for activation-induced deaminase during somatic hypermutation. *Nat Commun* 5:4137.
- Ward, I.M., B. Reina-San-Martin, A. Olaru, K. Minn, K. Tamada, J.S. Lau, M. Cascalho, L. Chen, A. Nussenzweig, F. Livak, M.C. Nussenzweig, and J. Chen. 2004. 53BP1 is required for class switch recombination. *Journal of Cell Biology* 165:459-464.
- Watanabe, S., K. Watanabe, V. Akimov, J. Bartkova, B. Blagoev, J. Lukas, and J. Bartek. 2013. JMJD1C demethylates MDC1 to regulate the RNF8 and BRCA1-mediated chromatin response to DNA breaks. *Nat Struct Mol Biol* 20:1425-1433.
- Weiser, B.P. 2020. Analysis of uracil DNA glycosylase (UNG2) stimulation by replication protein A (RPA) at ssDNA-dsDNA junctions. *Biochim Biophys Acta Proteins Proteom* 1868:140347.

- Weiser, B.P., J.T. Stivers, and P.A. Cole. 2017. Investigation of N-Terminal Phospho-Regulation of Uracil DNA Glycosylase Using Protein Semisynthesis. *Biophys J* 113:393-401.
- Wiedemann, E.M., M. Peycheva, and R. Pavri. 2016. DNA Replication Origins in Immunoglobulin Switch Regions Regulate Class Switch Recombination in an R-Loop-Dependent Manner. *Cell Rep* 17:2927-2942.
- Willmann, K.L., S. Milosevic, S. Pauklin, K.M. Schmitz, G. Rangam, M.T. Simon, S. Maslen, M. Skehel, I. Robert, V. Heyer, E. Schiavo, B. Reina-San-Martin, and S.K. Petersen-Mahrt. 2012. A role for the RNA pol II-associated PAF complex in AID-induced immune diversification. *J Exp Med* 209:2099-2111.
- Wu, L.C., and A.A. Zarrin. 2014. The production and regulation of IgE by the immune system. *Nature Reviews Immunology* 14:247-259.
- Xie, A., A. Kwok, and R. Scully. 2009. Role of mammalian Mre11 in classical and alternative nonhomologous end joining. *Nature Structural & Molecular Biology* 16:814-818.
- Xu, X., L. Wang, L. Hu, W.G. Dirks, Y. Zhao, Z. Wei, D. Chen, Z. Li, Z. Wang, Y. Han, L. Wei, H.G. Drexler, and Z. Hu. 2020. Small molecular modulators of JMJD1C preferentially inhibit growth of leukemia cells. *Int J Cancer* 146:400-412.
- Xu, Z., Z. Fulop, G. Wu, E.J. Pone, J. Zhang, T. Mai, L.M. Thomas, A. Al-Qahtani, C.A. White, S.-R. Park, P. Steinacker, Z. Li, J. Yates, B. Herron, M. Otto, H. Zan, H. Fu, and P. Casali. 2010. 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. *J Exp Med* 207:1124-1135.
- Xue, K., C. Rada, and M.S. Neuberger. 2006. The in vivo pattern of AID targeting to immunoglobulin switch regions deduced from mutation spectra in *msh2*<sup>-/-</sup> *ung*<sup>-/-</sup> mice. *J Exp Med* 203:2085-2094.
- Yadav, A., A. Olaru, M. Saltis, A. Setren, J. Cerny, and F. Livak. 2006. Identification of a ubiquitously active promoter of the murine activation-induced cytidine deaminase (AICDA) gene. *Mol Immunol* 43:529-541.
- Yagi, T., S. Nada, N. Watanabe, H. Tamemoto, N. Kohmura, Y. Ikawa, and S. Aizawa. 1993. A novel negative selection for homologous recombinants using diphtheria toxin A fragment gene. *Anal Biochem* 214:77-86.
- Yamane, A., W. Resch, N. Kuo, S. Kuchen, Z. Li, H.-W. Sun, D.F. Robbiani, K. McBride, M.C. Nussenzweig, and R. Casellas. 2011a. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *J Exp Med* 207:62-69.
- Yamane, A., W. Resch, N. Kuo, S. Kuchen, Z. Li, H.W. Sun, D.F. Robbiani, K. McBride, M.C. Nussenzweig, and R. Casellas. 2011b. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nat Immunol* 12:62-69.
- Yamane, A., D.F. Robbiani, W. Resch, A. Bothmer, H. Nakahashi, T. Oliveira, P.C. Rommel, E.J. Brown, A. Nussenzweig, M.C. Nussenzweig, and R. Casellas. 2013. RPA accumulation during class switch recombination represents 5'-3' DNA-end resection during the S-G2/M phase of the cell cycle. *Cell Rep* 3:138-147.
- Yan, C.T., C. Boboila, E.K. Souza, S. Franco, T.R. Hickernell, M. Murphy, S. Gumaste, M. Geyer, A.A. Zarrin, J.P. Manis, K. Rajewsky, and F.W. Alt. 2007. IgH class switching and translocations use a robust non-classical end-joining pathway. *J Exp Med* 204:478-482.
- Yeap, L.S., J.K. Hwang, Z. Du, R.M. Meyers, F.L. Meng, A. Jakubauskaite, M. Liu, V. Mani, D. Neuberger, T.B. Kepler, J.H. Wang, and F.W. Alt. 2015. Sequence-Intrinsic Mechanisms that Target AID Mutational Outcomes on Antibody Genes. *Cell* 163:1124-1137.
- Yen, W.-F., A. Chaudhry, B. Vaidyanathan, W.T. Yewdell, J.N. Pucella, R. Sharma, Y. Liang, K. Li, A.Y. Rudensky, and J. Chaudhuri. 2017. BRCT-domain protein BRIT1 influences class switch recombination. *Proceedings of the National Academy of Sciences* 114:8354-8359.
- Yoshikawa, K., I.-m. Okazaki, T. Eto, K. Kinoshita, M. Muramatsu, H. Nagaoka, and T. Honjo. 2002. AID Enzyme-Induced Hypermutation in an Actively Transcribed Gene in Fibroblasts. *Science* 296:2033-2036.



- Yousif, A.S., A. Stanlie, S. Mondal, T. Honjo, and N.A. Begum. 2014. Differential regulation of S-region hypermutation and class-switch recombination by noncanonical functions of uracil DNA glycosylase. *Proc Natl Acad Sci U S A* 111:E1016-1024.
- Yu, K., F. Chedin, C.L. Hsieh, T.E. Wilson, and M.R. Lieber. 2003. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat Immunol* 4:442-451.
- Yu, K., F.T. Huang, and M.R. Lieber. 2004. DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine. *J Biol Chem* 279:6496-6500.
- Zan, H., C. Tat, Z. Qiu, J.R. Taylor, J.A. Guerrero, T. Shen, and P. Casali. 2017. Rad52 competes with Ku70/Ku86 for binding to S-region DSB ends to modulate antibody class-switch DNA recombination. *8:14244*.
- Zan, H., Clayton A. White, Lisa M. Thomas, T. Mai, G. Li, Z. Xu, J. Zhang, and P. Casali. 2012. Rev1 Recruits Ung to Switch Regions and Enhances dU Glycosylation for Immunoglobulin Class Switch DNA Recombination. *Cell Reports* 2:1220-1232.
- Zanotti, K.J., and P.J. Gearhart. 2016. Antibody diversification caused by disrupted mismatch repair and promiscuous DNA polymerases. *38:110-116*.
- Zaprazna, K., and M.L. Atchison. 2012. YY1 Controls Immunoglobulin Class Switch Recombination and Nuclear Activation-Induced Deaminase Levels. *Molecular and Cellular Biology* 32:1542-1554.
- Zaprazna, K., A. Basu, N. Tom, V. Jha, S. Hodawadekar, L. Radova, J. Malcikova, B. Tichy, S. Pospisilova, and M.L. Atchison. 2018. Transcription factor YY1 can control AID-mediated mutagenesis in mice. *Eur J Immunol* 48:273-282.
- Zeng, X., D.B. Winter, C. Kasmer, K.H. Kraemer, A.R. Lehmann, and P.J. Gearhart. 2001. DNA polymerase  $\eta$  is an A-T mutator in somatic hypermutation of immunoglobulin variable genes. *Nature Immunology* 2:537-541.
- Zhang, J., A. Bottaro, S. Li, V. Stewart, and F.W. Alt. 1993a. A selective defect in IgG2b switching as a result of targeted mutation of the  $\gamma$ 2b promoter and exon. *The EMBO Journal* 12:3529-3537.
- Zhang, J., A. Bottaro, S. Li, V. Stewart, and F.W. Alt. 1993b. A selective defect in IgG2b switching as a result of targeted mutation of the  $\gamma$ 2b promoter and exon. *The EMBO Journal* 12:3529-3537.
- Zhang, Z.Z., N.R. Pannunzio, L. Han, C.L. Hsieh, K. Yu, and M.R. Lieber. 2014a. The strength of an Ig switch region is determined by its ability to drive R loop formation and its number of WGCV sites. *Cell Rep* 8:557-569.
- Zhang, Z.Z., N.R. Pannunzio, C.L. Hsieh, K. Yu, and M.R. Lieber. 2014b. The role of G-density in switch region repeats for immunoglobulin class switch recombination. *42:13186-13193*.
- Zheng, S., Bao, B. Vaidyanathan, J.-Y. Lin, F.-T. Huang, and J. Chaudhuri. 2015a. Non-coding RNA Generated following Lariat Debranching Mediates Targeting of AID to DNA. *Cell* 161:762-773.
- Zheng, S., A. Kusnadi, J.E. Choi, B.Q. Vuong, D. Rhodes, and J. Chaudhuri. 2019. NME proteins regulate class switch recombination. *FEBS Letters* 593:80-87.
- Zheng, S., B.Q. Vuong, B. Vaidyanathan, J.Y. Lin, F.T. Huang, and J. Chaudhuri. 2015b. Non-coding RNA Generated following Lariat Debranching Mediates Targeting of AID to DNA. *Cell* 161:762-773.
- Zhu, Q., P. Tan, Y. Li, M. Lin, C. Li, J. Mao, J. Cui, W. Zhao, H.Y. Wang, and R.F. Wang. 2018. DHX29 functions as an RNA co-sensor for MDA5-mediated EMCV-specific antiviral immunity. *PLoS Pathog* 14:e1006886.
- Zlatanou, A., E. Despras, T. Braz-Petta, I. Boubakour-Azzouz, C. Pouvelle, Grant, S. Nakajima, A. Yasui, Alexander, and Patricia. 2011. The hMsh2-hMsh6 Complex Acts in Concert with Monoubiquitinated PCNA and Pol  $\eta$  in Response to Oxidative DNA Damage in Human Cells. *Molecular Cell* 43:649-662.
- Zotos, D., and D.M. Tarlinton. 2012. Determining germinal centre B cell fate. *Trends Immunol* 33:281-288.

## Mécanismes impliqués dans la régulation de l'enzyme AID et de la réparation des dommages qu'elle induit à l'ADN

### Résumé

Au cours des réponses immunes, le répertoire des lymphocytes B est diversifié par les mécanismes d'hypermutation somatique (HMS) et de commutation isotypique (CI). Ces processus sont strictement dépendants de l'expression, du recrutement et de l'activité de l'enzyme Activation Induced cytidine Deaminase (AID) aux loci des immunoglobulines (Ig), et impliquent la génération consécutive de mutations et de cassures double brins dans l'ADN. Cependant, si AID induit des dommages à des loci non-Ig et que ceux-ci sont mal réparés, cela peut aboutir à une instabilité génomique à l'origine de lymphomes. AID possède ainsi un fort potentiel oncogénique et les mécanismes qui limitent celui-ci restent mal compris. Notre hypothèse est qu'il existe des protéines qui contrôlent et déterminent l'activité d'AID au niveau des loci Ig et des cibles non-Ig. Afin d'identifier ces facteurs, j'ai mis au point la technologie d'identification de protéines par biotinylation à proximité (BioID). Grâce à elle, j'ai identifié de nouveaux partenaires potentiels d'AID, dont la caractérisation est en cours. J'ai également participé à la caractérisation d'un facteur mal connu appelée Fam72a et qui interagit avec Ung2, l'enzyme principale capable d'exciser les uraciles de l'ADN introduits par AID. Nous avons pu observer que les cellules Fam72a<sup>-/-</sup> présentent un défaut de CI mais également une fréquence de mutations réduites dans l'HMS. Nos données indiquent que Fam72a permet de faire pencher la balance vers une réparation mutagène des uraciles aux loci Ig, en régulant la quantité cellulaire d'Ung2. Ces travaux devraient nous permettre de mieux comprendre comment se fait la régulation de l'activité d'AID et de la réparation des dommages qu'elle induit à l'ADN.

Mots clés : AID, commutation isotypique, hypermutation somatique, Fam72a, Ung, crible protéomique, BioID.

### Summary

During immune responses, B cell repertoire is diversified through somatic hypermutation (SHM) and class switch recombination (CSR). These mechanisms are strictly dependent on the expression, the recruitment and the activity of the enzyme Activation-induced cytidine deaminase (AID) to immunoglobulin (Ig) loci, and imply the consecutive generation of mutations and double strand breaks in DNA. However, if AID induces damage at non-Ig loci and if these are not faithfully repaired, this can lead to genomic instability that give rise to lymphomas. AID thus has a strong oncogenic potential and the mechanisms that limit this potential remain poorly understood. Our hypothesis is that there are proteins that control and determine AID activity at Ig loci and non-Ig targets. In order to identify these factors, I used the proximity-dependant biotinylation (BioID) technology. With it, I have identified potential new AID partners, which are currently being characterized. I also helped study a poorly known factor called Fam72a that interacts with Ung2, the main enzyme capable of excising DNA uracils introduced by AID. We were able to observe that Fam72a<sup>-/-</sup> cells show a defect in CSR but also a reduced frequency of mutations in SHM. Our data indicate that Fam72a<sup>-/-</sup> tilts the balance towards mutagenic repair of uracils at Ig loci by regulating the cellular amount of Ung2. Overall, this work should allow us to better understand how AID activity is regulated and how the DNA damage it induces is repaired.

Keywords : AID, Class switch recombination, Somatic hypermutation, Fam72a, Ung, Proteomic screen, BioID.