



HAL
open science

Mécanismes moléculaires impliqués lors de la commutation isotypique

Mélanie Rogier

► **To cite this version:**

Mélanie Rogier. Mécanismes moléculaires impliqués lors de la commutation isotypique. Immunologie. Université de Strasbourg, 2020. Français. NNT : 2020STRAJ085 . tel-04416348

HAL Id: tel-04416348

<https://theses.hal.science/tel-04416348>

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.

UNIVERSITÉ DE STRASBOURG

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

THÈSE présentée par :

Mélanie ROGIER

soutenue le : **14 décembre 2020**

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline/ Spécialité : Aspects moléculaires et cellulaires de la biologie

Immunologie

**Mécanismes moléculaires impliqués
lors de la commutation isotypique**

THÈSE dirigée par :

Dr. Bernardo REINA-SAN-MARTIN

DR2, IGBMC, Université de Strasbourg

RAPPORTEURS :

Dr. Amine Ahmed KHAMLI

DR2, IPBS Toulouse

Dr. Sandrine ROULLAND

DR, CIML, Marseille

AUTRES MEMBRES DU JURY :

Pr. Pauline SOULAS-SPRAUEL

Professeur, CRBS, Strasbourg

Remerciements

Je tiens à remercier tous les membres de mon jury de thèse, Sandrine Roulland, Amine Khamlichi et Pauline Soulas Sprauel pour avoir accepté d'évaluer ce travail.

Je remercie mon directeur de thèse, Bernardo, pour m'avoir ouvert les portes de ton équipe il y a maintenant plus de 8 ans. Merci pour ta confiance, tes encouragements et ton soutien durant toutes ces années.

Un énorme merci à mon binôme et mon partenaire de labo et de vie, Jacques. Merci pour ton aide, ton soutien et ta disponibilité durant ces 4 années. Merci d'avoir tenté ou réussi à trouver les mots justes pour m'apaiser lors des grands moments de stress (présentations orales, écriture de thèse). Ta présence à mes côtés m'a été très précieuse.

Un énorme merci également à Isa, pour ton aide, ton soutien sans faille et ta disponibilité. Merci d'avoir rendu nos journées au labo plus magiques ainsi que nos nocturnes, avec une bonne Jossélina, à sous cloner les librairies en écoutant du Ed Sheeran. Je pense que tu sais ce que je ressens ou à quel point tu comptes pour moi. Ton surnom "la maman du labo" né bien avant mon arrivée au labo veut tout dire. Tu es tout simplement une personne extraordinaire sur qui on peut toujours compter et avoir une confiance aveugle. Aujourd'hui, tu es bien plus qu'une collègue ou une simple amie. Ton départ du labo a été très difficile, mais heureusement nos retrouvailles chaque mois autour d'une table de l'un de nos 2 restaurants préférés nous réchauffent le cœur. J'ai hâte de te retrouver lorsque la situation sera meilleure. Encore merci pour tout Isa.

Arthur, je te remercie pour toute ton aide apportée ces derniers mois et surtout durant l'écriture de cette thèse ! Un grand merci pour ta relecture de l'introduction et pour tes conseils ! C'est triste que tu nous quittes déjà ! Bonne route à toi !

Merci à toi Vincent, mon premier mentor quand je suis arrivée au labo. Merci pour tous les chocolats gratuits et les tablettes gagnées suite à nos paris !

Un grand merci à mes deux adorables stagiaires, Aurore et Dilara. Merci à toutes les deux pour votre aide sur les différents projets de cette thèse ainsi que pour les bons moments passés ensemble. J'espère bientôt vous revoir.

Merci à toi Rocio, pour ta disponibilité, ton aide et ton soutien apportés durant ces 4 années.

Un grand merci à toi Léa, pour toute ton aide, ton soutien et ta gentillesse au labo mais également encore aujourd'hui. Merci pour tous les ateliers DIY réalisés pour chaque anniversaire ou thèse ainsi que tous les bons moments passés ensemble au labo, en rando ou lors de nos retrouvailles. J'ai hâte de te retrouver lorsque la situation sera meilleure autour d'une Jossélina avec Isa et Jacques !

Merci à toi Ebe, pour ta gentillesse sincère et pour ton soutien par message durant l'écriture de cette thèse ! J'ai hâte de te retrouver autour d'une table à Florence !

Merci à toi Anne-So pour tous les bons moments partagés au labo ainsi qu'à l'extérieur.

Merci à Evi Soutoglou, Mickael Weber et Laurent Brino, pour avoir fait partie de mon comité de mi-thèse et pour tous les conseils que vous m'avez apportés.

Merci à toute l'équipe d'Evi Soutoglou avec qui nous partageons notre pièce culture et de très bons moments. Je remercie plus particulièrement Duygu, merci pour ton soutien et pour ton aide. Ton amitié m'a beaucoup apporté ces derniers mois ! Je remercie également Sylvain, Ophélie, Lucile, Audrey et Céline pour tous les bons moments passés ensemble au labo ainsi qu'à l'extérieur. J'espère que l'on pourra bientôt venir à Brighton pour vous revoir !

Merci à toute l'équipe Tomasetto, avec qui je partage le box 3061. Merci de m'avoir accueilli à bras ouvert et de me considérer comme l'un des membres de votre équipe. Merci à Laetitia, Amélie et Mehdi avec qui j'ai partagé ce box 3061, pour tous les moments passés ensemble. Merci à Catherine et Fabien pour toutes les discussions scientifiques et d'avenir ainsi que pour vos conseils. Merci à toute l'équipe pour toutes les discussions, scientifiques ou non, et tous les conseils apportés.

Merci à l'équipe Ricci et plus particulièrement à Kévin et Eric pour votre aide et vos conseils mais aussi pour les bons moments passés ensemble.

Un énorme merci à Claudine et Muriel pour votre aide précieuse durant les analyses FACS ou les tris mais aussi durant toutes ces années. Vos petits gâteaux et nos discussions vont me manquer ! Merci à Laurent Brino, pour tous les conseils apportés sur le crible CRISPR. Merci à Bernard, Romain et Doulaye pour leur disponibilité et les discussions concernant l'analyse et le séquençage à haut débit. Merci également à France pour l'organisation administrative de ma soutenance de thèse.

Merci à La Ligue contre le Cancer pour avoir financé mes 3 premières années.

Enfin, merci à tous mes amis de l'IGBMC avec qui j'ai partagé de très bons moments : Duygu, Chiara, Sarah, Séverine, Laetitia, Amélie, Thomas, Kévin, Arthur, Medhi, Sylvain, Ophélie et Lucile. Un énorme merci pour votre soutien !

Un merci à l'infini à mes 4 murs porteurs, mes parents, Marcelle et Jacky et mes frères, Samuel et Florian. Maman et papa, je vous remercie pour toutes les valeurs de la vie que vous m'avez enseignées et qui ont fait la personne que je suis aujourd'hui. Merci de m'avoir fait confiance et d'avoir respecté mes choix. Merci pour votre bienveillance et votre soutien (moral, physique et bien évidemment financier) qui m'ont permis de traverser toutes les épreuves rencontrées.

Sami et Flo, je vous remercie pour votre soutien et votre bienveillance depuis toujours. Je suis très certainement votre première et plus grande admiratrice ! Vous êtes des exemples pour moi ! Vos réussites et vos encouragements m'ont toujours rendue plus forte pour continuer et réussir. Un grand merci aussi à mes belles sœurs Jehane et Fanny, pour votre soutien ainsi qu'à tous mes neveux et nièce, Ilhan, Olivia, Noham et Adam.

Un immense merci et hommage à mes grands-parents maternels, Marcel et Lulu, qui ont gardé presque tous leurs petits-enfants jusqu'au collège. Merci pour tous ces moments passés ensemble qui m'ont apporté les valeurs de partage, de loyauté, de persévérance et de force. Merci à toi, grand père, qui a toujours cru en moi, merci de m'avoir ouvert ton jardin extraordinaire et de m'avoir montré l'importance de la terre et des choses simples. Merci à toi, ma plus grande confidente, ma petite Lulu pour tous ces moments partagés toutes les deux qui resteront à jamais gravés dans ma mémoire.

Un immense merci aussi et hommage à mon parrain Francis et un immense merci à ma tante Josette sur qui j'ai toujours pu compter. Merci pour votre confiance, votre soutien et votre aide depuis toujours. Merci également à Emilie et Jérôme d'avoir toujours été là pour moi.

Un énorme merci à ma Tic, ma meilleure amie et bien plus encore. Je te remercie pour tout le soutien que tu m'as apporté durant ces 15 dernières années ainsi que ces 4 années de thèse et plus particulièrement pendant l'écriture ! Merci pour tes appels et tes messages de force. Un énorme merci aussi à ton mari Cyril qui a corrigé, je l'espère toutes mes fautes ! Merci beaucoup d'avoir pris le temps de lire cette thèse et d'y avoir ajouté des corrections ou des commentaires.

Enfin, merci à tous les deux pour tous ces moments passés ensemble et que nous pourrons, très prochainement je l'espère, partager à nouveau.

Merci à ma marraine, Claudine, ainsi que tout le reste de ma famille pour leur soutien et leur bienveillance.

Je remercie énormément ma belle-famille, en particulier mes beaux-parents, Didier et Isabelle pour votre soutien, votre écoute et votre aide durant ces 3 dernières années. Merci également de nous avoir donné l'opportunité d'habiter dans un cadre exceptionnel juste avant le confinement. Merci aussi à ma belle-soeur Hélène et à son mari Thomas, pour votre soutien et votre aide ainsi qu'à Armand et Simone.

Merci à mes deux meilleurs thérapeutes à 4 pattes, mes M&M's qui partagent ma vie depuis 5 et 6 ans. Merci pour leurs câlins et leur présence rassurante. 🐾 🐾

Table des matières

Liste des figures.....	1
Liste des abréviations.....	3
Préambule.....	6

Introduction

I- La diversification du récepteur des cellules B (BCR).....	9
1. Le récepteur des cellules B (BCR).....	9
2. L'assemblage du BCR.....	11
2.1. L'organisation des loci Ig murins.....	11
2.2. La recombinaison V(D)J.....	12
3. La diversification dépendante de l'antigène.....	16
3.1. Les centres germinatifs.....	16
3.2. L'hypermutation somatique.....	17
3.3. La commutation isotypique.....	19
3.4. La conversion génique.....	21
II- L'enzyme Activation-induced Cytidine Deaminase (AID).....	24
1. La découverte d'AID.....	24
2. La structure et la fonction d'AID.....	26
3. Le modèle de désamination de l'ADN.....	28
4. Le potentiel oncogénique d'AID et la régulation de l'activité d'AID.....	29
4.1. Le potentiel oncogénique d'AID.....	29
4.1.1 Les mutations hors-cible.....	29
4.1.2. Les translocations.....	30
4.1.3. L'expression ectopique d'AID.....	31
4.1.4. <i>Kataegis</i>	32
4.2. La régulation d'AID.....	33
4.2.1. La régulation transcriptionnelle.....	33
4.2.2. La régulation de la localisation cellulaire.....	35
4.2.3. La régulation de l'activité d'AID.....	36
5. Le ciblage d'AID.....	38
5.1. Les motifs.....	38
5.2. Les éléments en <i>cis</i>	39
5.3. La transcription.....	40
5.4. Le complexe de l'exosome à ARN.....	40
5.5. L'épissage des ARN.....	41

III-	La génération de mutations induites par AID durant l'HMS.....	45
IV-	Le rôle d'AID dans la commutation isotypique.....	48
1.	La transcription des régions de switch.....	48
2.	La séquence et la spécificité structurelle du locus IgH.....	54
3.	La formation de DSBs par le BER et le MMR.....	55
4.	La DDR.....	60
5.	La réparation des DSBs.....	65
5.1.	La réparation des DSBs par le C-NHEJ.....	64
5.2.	La réparation des DSBs par le A-EJ.....	65
V-	Les immunodéficiences de la commutation isotypique.....	68
1.	Les CSR-ID dues aux défauts de l'activation des cellules B.....	68
2.	Les déficiences de la CI dues aux défauts intrinsèques des cellules B.....	69
2.1.	La déficience de la CI causée par des mutations dans le gène Aicda.....	70
2.2.	La déficience de la CI causée par un défaut de la réparation de l'ADN.....	70
	Hypothèse de travail.....	72

Résultats

Chapitre 1 : Vers l'identification de gènes impliqués dans la commutation isotypique74

I-	Identification des gènes impliqués dans la commutation isotypique par criblage CRISPR/Cas9 à l'échelle du génome murin.....	77
1.1.	Stratégie du criblage fonctionnel CRISPR/Cas9 sur la CI.....	79
1.2.	D'une librairie lentivirale vers une librairie rétrovirale.....	81
1.2.1.	La préparation du vecteur LentiCRISPR-v6.....	81
1.2.2.	Le test de fonctionnalité du vecteur LentiCRISPR-v6.....	83
1.2.3.	Une nouvelle stratégie : l'utilisation de vecteurs rétroviraux.....	83
1.2.3.1.	La génération et la caractérisation d'une lignée cellulaire CH12 stable exprimant Cas9.....	85
1.2.3.2.	Preuve de principe : criblage fonctionnel sur la CI à petite échelle.....	88
1.3.	Le crible CRISPR/Cas9 pour identifier les gènes impliqués dans la CI avec la librairie mGeCKO-v2.....	92
1.3.1.	La génération de la librairie Rétro-mGeCKO-v2.....	92
1.3.2.	Le crible fonctionnel sur la CI avec la librairie Rétro-mGeCKO-v2.....	96
1.3.2.1.	La réalisation du crible.....	96
1.3.2.2.	L'analyse du crible fonctionnel sur la CI.....	97
1.4.	Un crible CRISPR/Cas9 d'une plus grande précision pour identifier les cofacteurs de la CI avec la librairie Rétro-Brie.....	101
1.4.1.	La génération de la librairie Rétro-Brie.....	101
1.4.2.	Le crible fonctionnel sur la CI à l'aide de la librairie Rétro-Brie.....	104
1.4.2.1.	La réalisation du crible.....	104
1.4.2.2.	L'analyse du crible fonctionnel sur la CI.....	108

1.5.	Analyse globale des deux cribles.....	111
II-	Discussion.....	115
Chapitre 2 : Etude du rôle de Fam72a dans la diversification des Ig.....		119
I-	Fam72a contrôle la balance entre la réparation de l'ADN mutagène et fidèle durant la diversification des anticorps.....	120
	Ebauche du papier	
II-	Résultats supplémentaires.....	145
	1. Reconstitution de l'expression de mFam72a et mFam72a ^{W125R} dans les cellules Fam72a ^{-/-}	145
	2. Reconstitution de l'expression des versions humaines hFAM72A-D dans les cellules Fam72a ^{-/-}	147
	3. L'activité d'Ung est-elle modifiée en l'absence de Fam72a ^{-/-} ?.....	149
III-	Discussion.....	150
	1. Ung2 et la CI : une histoire de quantité ?.....	150
	1.1. Une trop grande quantité d'Ung2 porte-t-elle atteinte à la CI ?.....	150
	1.2. Et une quantité trop faible ?.....	151
	2. Fam72a et Ung2 dans la CI : d'autres régulations possibles.....	152
	3. Et Ung1 dans tout ça ?.....	153
	4. Fam72a et Ung2 dans l'HMS.....	154
	5. Perspectives.....	155
	5.1. Ung2 est-elle bien surexprimée en absence de Fam72a ?.....	155
	5.2. Fam72a a-t-il un autre rôle que celui de dégrader Ung2 ?.....	155
	5.3. Est-ce que le recrutement d'AID à S μ est altéré en absence de Fam72a ?.....	156
	5.4. Est-ce que l'activité d'Ung est altérée en absence de Fam72a ?.....	156
	5.5. Y-a-t-il moins d'uracile dans l'ADN en absence de Fam72a ?.....	156
	5.6. La dégradation d'Ung2 est-elle dépendante du protéasome ?.....	157
	5.7. Y-a-t-il plus ou moins de DSBs non réparées induites par AID en l'absence de Fam72a ?.....	157
	5.8. Est-ce que Ung2 est supprimée dans les lignées de cellules cancéreuses humaines surexprimant Fam72a ?.....	158
	5.9. Est-ce que la transcription des régions S est affectées dans les cellules B primaires Fam72a ^{-/-} ?.....	158
	6. Conclusion.....	158
Chapitre 3 : Etude de l'impact de Zbtb24 dans la CI.....		159
I-	La perte de ZBTB24 compromet le NHEJ et la CI chez les patients atteints d'ICF.....	160
	Publication	
II-	Discussion.....	193
Chapitre 4 : Caractérisation fonctionnelle de l'enhancer γ1E lors de la CI.....		195
I-	Une nouvelle région régulatrice contrôle la transcription du locus IgH et la CI à un sous-ensemble d'isotypes.....	196
	Publication	
II-	Discussion.....	205

Conclusion générale.....	208
Matériels et Méthodes.....	209
Bibliographie.....	213

Liste des figures

Figure 1: Structure du récepteur des cellules B.....	8
Figure 2: Mécanismes de diversification des anticorps.....	8
Figure 3: Organisation des loci Ig murins.	10
Figure 4: Les étapes de la recombinaison V(D)J.....	10
Figure 5: Le développement des cellules B.....	14
Figure 6: L'hypermutation somatique.....	18
Figure 7: La commutation isotypique.....	18
Figure 8: La conversion génique.....	22
Figure 9: Les domaines fonctionnels de la protéine Activation-Induced cytidine Deaminase (AID).....	25
Figure 10: Le modèle de désamination de l'ADN par AID.....	27
Figure 11: Les mécanismes régulateurs d'AID.....	43
Figure 12: La réaction multi-étapes de la commutation isotypique.....	49
Figure 13: Les domaines fonctionnels des isoformes de l' Uracil-DNA-Glycosylase (UNG).....	57
Figure 14: Formation et réparation des cassures double brin (DSBs) durant la CI.....	61
Figure 15: Stratégie du crible d'inactivation par CRISPR/Cas9 sur la CI à large échelle.....	78
Figure 16: Le nouveau LentiCRISPR-v6 est fonctionnel mais n'infecte pas la lignée cellulaire CH12.....	80
Figure 17: Nouvelle stratégie basée sur l'utilisation de vecteurs rétroviraux.....	82
Figure 18: La lignée cellulaire CH12 exprimant la nucléase Cas9 (CH12 ^{Cas9}) est fonctionnelle.....	84
Figure 19: Un crible à petite échelle permet d'identifier les gènes requis lors la CI.....	86
Figure 20: Le clonage de la librairie Rétro-mGeCKO-v2.....	90
Figure 21: Criblage de la librairie Rétro-mGeCKO-v2.....	94

Figure 22: Le crible CRISPR/Cas9 avec la librairie Rétro-mGeCKO-v2 permet d'identifier des gènes connus pour être impliqués dans la CI et de nouveaux gènes candidats.....	98
Figure 23: Le clonage de la librairie Rétro-Brie.....	102
Figure 24: Criblage de la librairie Rétro-Brie.....	106
Figure 25: Le crible CRISPR/Cas9 avec la librairie Rétro-Brie permet d'identifier des gènes connus pour être impliqués dans la CI et de nouveaux gènes candidats.....	109
Figure 26: Analyse des deux cribles CRISPR/Cas9 réunis.....	113
Figure 27: Reconstitution des cellules Fam72a ^{-/-} avec les formes souris, mFam72a et mFam72a ^{W125R} et les formes humaines hFAM72A-D.....	146
Figure 28: Test d'activité d'Ung en absence de Fam72a.....	148

Liste des abbreviations

3'RR: region régulatrice en 3', 3' regulatory region
53BP1: p53 binding protein
Ac: anticorps
ADNss : ADN simple brin
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
APRIL: A proliferation-inducing ligand
ARNg: ARN guide
ARNPII: ARN Polymérase II
ATM: ataxia telangiectasia mutated
BATF: basic leucine zipper transcriptional factor ATF-like)
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
CLL : leucémie lymphocytaire chronique
CMHII : complexe majeur d'histocompatibilité de classe II
C-NHEJ : voie Classique du NHEJ
CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats
CRM1: chromosomal maintenance1, exportin 1
CSR-ID: Class switch recombination-immundeficiency
C-ter : extrémité carboxy-terminale
CtIP : CtBP-interacting protein
CTNBL1 : catenin Beta-like protein1
D : segments géniques de diversité
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
EDA-ID: ectodermal dysplasia associated with immunodeficiency
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
FOXO1: Forkhead box protein O1
GANP: Germinal center associated nuclear Protein
GC: centre germinatifs, germinal centers
GeCKO: genome-scale CRISPR knock-out
GLT: germline transcript
HIGM: Hyper-IgM syndrome
HMCEs: 5-Hydroxymethylcytosine binding, ES-cell-specific
HMS: hypermutation somatique
HMW: high molecular weight
hnRNPs: heterogeneous nuclear ribonucleoproteins
HP1: heterochromatin protein 1
HR: homologous recombination
HSP90: heat shock protein-90
ICF: immunodeficiency, centromeric region instability, facial anomalies
Id2: inhibitor of differentiation 2
Ig: immunoglobuline
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
J : segments géniques de jonction
KAP1 : KRAB domain-associated protein 1
LMW: light molecular weight
LPS: lipopolysaccharide
LZ : zone claire, light zone
MAGeCKO: model-based analysis of genome-wide CRISPR-Cas9 knockout
MDC1: mediator of DNA damage checkpoint protein 1
MLH1 : MutL homolog 1
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
NEMO: NF-kB essential modulator
NES: signal d'export nucléaire nuclear export signal
NF-κB: 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 NH₂-terminale
PAF: RNAP II-associated factor complex
PAM: protospacer adjacent motif
Parg: poly (ADP-ribose) glycohydrolase
Parp: poly (ADP-ribose) polymerase
PCNA: proliferating cell nuclear antigen
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
Regγ: 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
Stat6: signal transducer and activator of transcription 6
TAD: topologically associating domain
TCR: T cell receptor
TdT : terminal deoxynucleotidyl transferase
Tfh : cellules T auxiliaires folliculaire, T follicular helper
TGF-β: transforming growth factor β
TLR: Toll-like receptor
TSS : site d'initiation de la transcription, transcription start site
UNG: uracil-DNA glycosylase
V : région variable
WT: wild-type
XLF: XRCC4-like factor
XRCC4: X-ray repair cross-complementing protein
YY1: Yin Yang 1

Préambule

Notre système immunitaire assure la défense de notre corps face à la grande diversité de microorganismes que nous rencontrons chaque jour (virus, bactéries, champignons ou parasites). Chez les mammifères, le système immunitaire regroupe une grande variété de cellules et de molécules et est dissocié en deux lignes de défense : le système immunitaire inné et adaptatif. Le système immunitaire inné fournit une protection immédiate mais ne reconnaît que des structures invariantes des pathogènes. Il permet également l'activation du système immunitaire adaptatif qui est constitué par les lymphocytes T et B. La grande spécificité de ces 2 types cellulaires réside dans la capacité quasi-infinie de reconnaissance des pathogènes de par la gigantesque diversité des récepteurs qu'ils expriment. De plus, les lymphocytes B, producteurs d'anticorps, ont la capacité de diversifier une seconde fois les récepteurs qu'ils expriment afin de s'adapter et de mieux reconnaître le pathogène auxquels ils font face. Les mécanismes qui permettent cette seconde diversification sont au cœur de mon sujet de recherche. Pour que ces mécanismes puissent avoir lieu, l'induction de dommages à l'ADN est nécessaire. Ces lésions peuvent être très dangereuses pour l'intégrité génomique des cellules B et peuvent être à l'origine d'une transformation cancéreuse. Ainsi, au cours de l'évolution, des mécanismes moléculaires sont apparus afin de permettre la diversification efficace des anticorps tout en protégeant les cellules B des dommages collatéraux induits. Cependant, ces mécanismes restent aujourd'hui encore mal connus, et mon objectif a ainsi été de les étudier.

Introduction

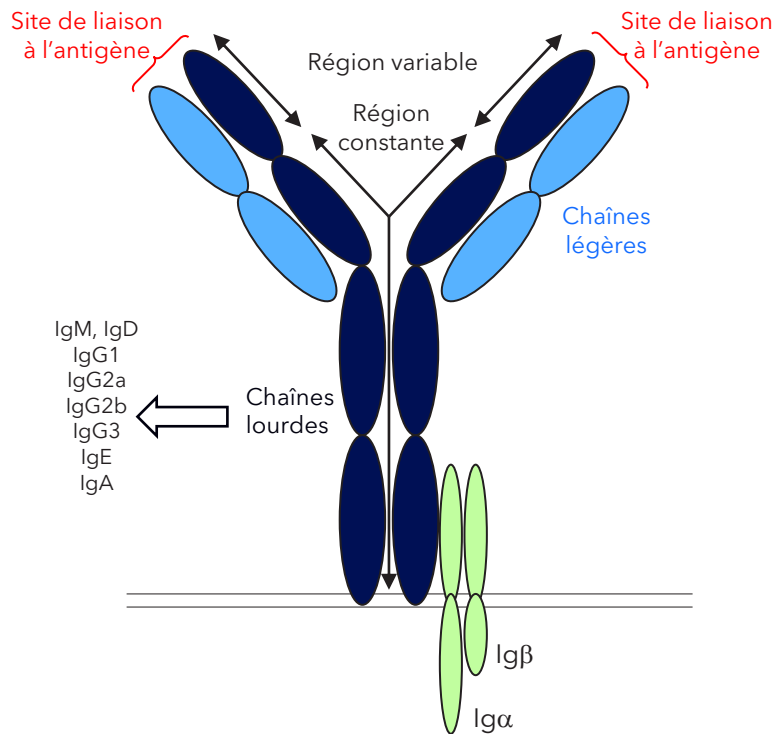


Figure 1: Structure du récepteur des cellules B. Le récepteur des cellules B (BCR) est composé de deux chaînes lourdes et de deux chaînes légères, chacune d'elle contient une région variable et une région constante. Les régions variables des chaînes lourdes et légères forment le site de liaison de l'antigène, tandis que les régions constantes des chaînes lourdes déterminent l'isotype et la fonction effectrice de l'anticorps. Le BCR est exprimé à la surface des cellules B et s'associe à l'hétérodimère Ig α /Ig β pour permettre la transduction du signal suite à la reconnaissance d'un antigène.

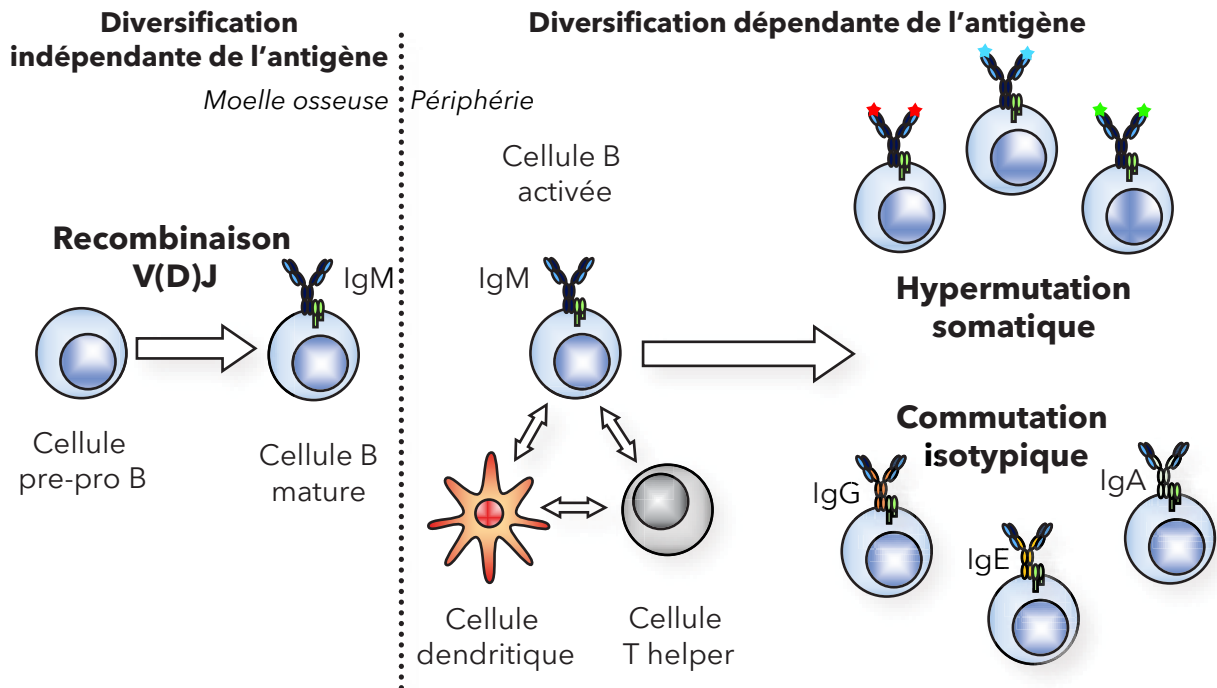


Figure 2: Mécanismes de diversification des anticorps. La génération du répertoire primaire des anticorps se produit dans la moelle osseuse, lors du développement des cellules B, par la recombinaison V(D)J. Les cellules B matures migrent ensuite en périphérie. Suite à la reconnaissance de l'antigène apparesenté et à l'activation médiée par les cellules T, elles vont se diriger dans les organes lymphoïdes secondaires, où elles vont diversifier leur répertoire d'anticorps par les mécanismes d'hypermuation somatique et de commutation isotypique.

I- La diversification du récepteur des cellules B (BCR)

La génération de l'immunité humorale spécifique et adaptative permettant l'élimination des agents pathogènes est dépendante de la diversité du répertoire des cellules B. Celle-ci se produit lors du développement des cellules B dans la moelle osseuse ainsi qu'au cours des réponses immunes.

1. Le récepteur des cellules B (BCR)

Afin de combattre efficacement les infections, les cellules B expriment à leur surface un récepteur unique pour l'antigène (Ag), appelé récepteur des cellules B (BCR, de l'anglais B cell receptor), capable de reconnaître spécifiquement une grande variété d'antigènes de bactéries, virus et d'autres pathogènes. Le BCR est une immunoglobuline (Ig) membranaire associée à deux éléments de signalisation qui forment l'hétérodimère $Ig\alpha/Ig\beta$ (CD79a/CD79b). L'Ig membranaire, aussi appelée anticorps (Ac) membranaire, est constituée de deux chaînes légères ($Ig\kappa$ ou $Ig\lambda$, collectivement appelées IgL) et de deux chaînes lourdes (IgH) identiques entre elles, reliées par des ponts disulfures (**Figure 1**). Chaque chaîne comprend une région variable (V) et une région constante (C). D'une part, l'association des régions variables des chaînes IgH et IgL comprend 6 régions déterminant la complémentarité de l'Ac pour l'Ag (CDRs, de l'anglais complementarity-determining region) qui constituent le site de reconnaissance de l'Ag. D'autre part, la région constante de la chaîne IgH détermine l'isotype exprimé (IgM , IgD , IgG , IgA ou IgE) et la fonction effectrice des anticorps. Le BCR permet donc de reconnaître l'antigène et de transduire le signal de reconnaissance de l'Ag par l'intermédiaire de l'hétérodimère $Ig\alpha/Ig\beta$ afin d'activer les cellules B et d'initier des réponses immunitaires humorales (Chen and Wang, 2019).

Le répertoire des Ac produit par les lymphocytes B est estimé à plus de 10^{11} et est généré par des réarrangements chromosomiques (recombinaison et mutation) (Hozumi and Tonegawa, 1976) qui ont lieu au sein des loci IgH et IgL . Notamment, quatre mécanismes ont été identifiés pour contribuer à la diversification des Ig : la recombinaison V(D)J qui se produit durant les premiers stades du développement

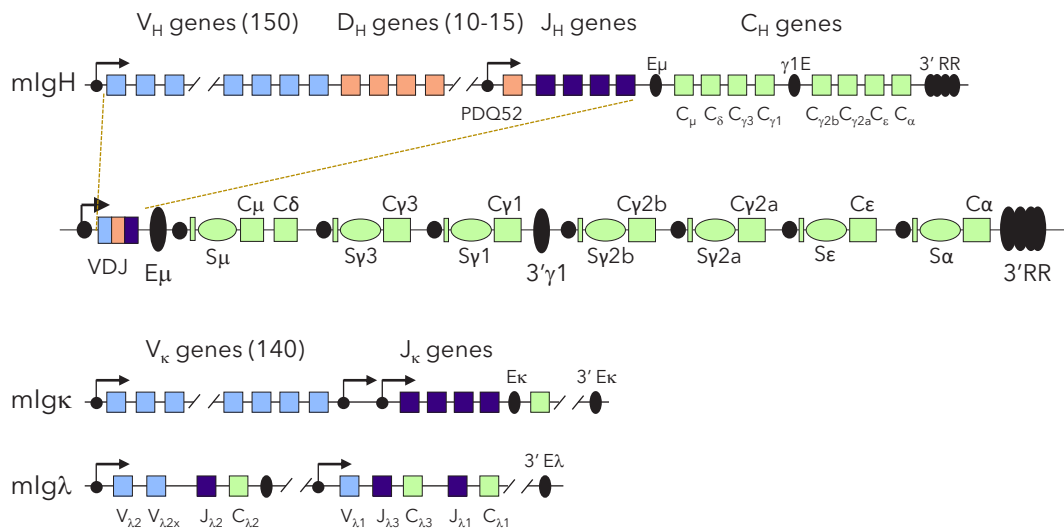


Figure 3: Organisation des loci Ig murins. Représentation schématique des loci IgH avant et après recombinaison V(D)J, Igκ et Igλ murins. Le locus IgH murin est composé de segments de gènes Variables (V; en bleu), de Diversité (D; orange), et de Jonction (J; violet) qui codent pour la région variable de l'Ig, suivis par les unités transcriptionnelles codant la région constante (C). Les chaînes IgL (Igκ et Igλ) contiennent les gènes V, J et C qui sont organisés par famille pour la chaîne κ et en cassettes distinctes pour la chaîne λ. Les nombres de segments géniques sont indiqués entre parenthèses. Les promoteurs sont représentés par des points noirs et les enhancers en formes ovales noires. Adapté de (Cobb et al. 2006 ; Schatz et al., 2011).

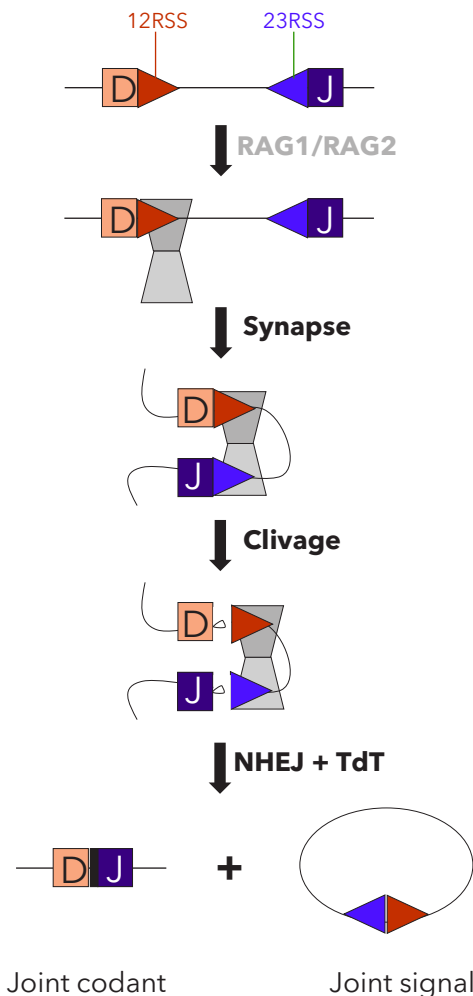


Figure 4: Les étapes de la recombinaison V(D)J.

Les segments de gènes D et J sont bordés par des séquences signal de recombinaison (RSS) 12RSS et 23RSS. Le complexe RAG (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 de gène et les RSS. Ensuite, RAG participe à la réparation de ces cassures double 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 ligé et va être éliminé. Adapté de (Ru, Zhang et al. 2018).

des cellules B avant la rencontre avec l'antigène ; et les trois mécanismes dépendants de l'antigène tels que l'hypermutation somatique (HMS), la commutation isotypique (CI) et la conversion génique (**Figure 2**).

2. L'assemblage du BCR :

2.1. L'organisation des loci Ig murin

Les gènes codant pour les Ig sont organisés sous forme de segments géniques non fonctionnels et sont regroupés en 3 loci, IgH, Igκ et Igλ (**Figure 3**).

Le locus IgH murin est localisé sur le chromosome 12 et s'étend sur 3 Mb (Chevallard et al., 2002). De 5' en 3', il est composé des segments de gènes variables (V), de diversité (D) et de jonction (J) qui forment la région V de l'IgH suivie par des exons codant pour la région C de l'IgH (**Figure 3**). Les gènes V_H sont approximativement 150 -selon la souche de souris- et classifiés en 16 familles de gènes définies par des similitudes de séquence, puis sont accompagnés par 10 à 15 segments géniques D_H et enfin 4 segments J. Les gènes codant pour la région V_H sont ensuite suivis par 8 exons de la région constante codant pour les différents isotypes (C_μ , C_δ , $C_{\gamma 3}$, $C_{\gamma 1}$, $C_{\gamma 2b}$, $C_{\gamma 2a}$, C_ϵ et C_α). Les gènes de la région constante de l'IgH sont organisés en unité transcriptionnelle individuelle qui est constituée d'un promoteur sensible aux cytokines en amont d'un exon intronique (I), d'une région de « switch », aussi appelée région S, à l'exception de C_δ , et d'un exon constant (C). Les régions S sont des régions répétitives mais non-homologues avec des séquences riches en GC.

Le locus IgH contient trois principaux éléments régulateurs en *cis*: le promoteur/enhancer PDQ52 qui est positionné en 3' du dernier segment D_H , l'enhancer E_μ qui est localisé dans l'intron entre l'exon J_{H4} et C_μ , et le super-enhancer 3'RR se trouvant à l'extrémité 3' du locus IgH, aussi appelé région régulatrice en 3', qui contient 7 sites hypersensibles à la DNaseI (hs3a, 3b, 1-2, 4, 5, 6, et 7), constituant des sites de liaisons de facteurs de transcription.

Les loci codants pour les chaînes Igκ et Igλ affichent une organisation légèrement différente en présentant seulement les segments géniques V, J et C. Le locus Igκ est situé sur le chromosome 6 et est composé d'environ 140 segments géniques V_k et

4 exons J_k fonctionnels, suivis par un unique exon C_k tandis que le locus $Ig\lambda$ est situé sur le chromosome 16 et est constitué de trois différentes cassettes de segments V_λ/J_λ suivis d'exons C_λ (Cobb et al., 2006).

Comme chez la souris, les gènes Ig humains sont regroupés en 3 loci, IgH , $Ig\kappa$ et $Ig\lambda$, localisés sur trois chromosomes. Le locus IgH humain, localisé sur le chromosome 14, est semblable au locus IgH murin mais possède des isotypes différents et en plus grand nombre dû à une duplication génomique des gènes constants γ , ϵ et α et de la région 3'RR (D'Addabbo et al., 2011). De 5'en 3', il est donc composé des segments de gènes variables (V), de diversité (D) et de jonction (J) suivis par les exons de la région constante C_μ , C_δ , $C\gamma3$, $C\gamma1$, $C\phi e$, $C\alpha1$, $C\phi g$, $C\gamma2$, $C\gamma4$, $C\epsilon$ et $C\alpha2$. Il possède également l'enhancer E_μ ainsi que deux régions 3'RR situées à chaque extrémité 3' de l'exon constant $C\alpha$.

La région variable des Ig doit être assemblée par la recombinaison V(D)J, mécanisme indépendant de la rencontre avec l'Ag se produisant durant le développement des cellules B dans la moelle osseuse afin de permettre l'expression des chaînes IgH et IgL .

2.2. La recombinaison V(D)J

La recombinaison V(D)J est une recombinaison qui assemble les exons de la région variable des récepteurs des cellules B mais aussi des cellules T (TCR, en anglais T cell receptor), générant un répertoire diversifié de récepteurs d'Ag. Cette recombinaison permet l'assemblage des segments de gènes V, D et J codant pour la région variable des Ig (IgH , $Ig\kappa$ et $Ig\lambda$) et des loci du TCR afin de générer un récepteur fonctionnel.

La recombinaison V(D)J est dépendante de l'expression des recombinases RAG1 et 2 (RAG, de l'anglais recombination-activating gene) (Mombaerts et al., 1992; Shinkai et al., 1992), dans les cellules pré-pro B. L'ensemble des deux recombinases est appelé le complexe RAG : l'endonucléase RAG1 est essentielle pour la liaison et le clivage de l'ADN ainsi que la formation de l'épingle nécessaire pour la réaction de clivage (Landree et al., 1999; Schatz and Swanson, 2011) et la recombinase RAG2 renforce les interactions entre RAG1 et l'élément heptamère, et est un cofacteur

essentiel pour le clivage de l'ADN (Rommel et al., 2017). La recombinaison V(D)J se déroule au niveau de sites spécifiques appelés séquences de signal de recombinaison (RSS, de l'anglais recombination signal sequences) qui entourent les segments de gènes codant V, D et J. Ces courtes séquences sont constituées d'éléments heptamère et nonamère conservés, séparés par une région relativement non conservée de 12 ou 23 paires de bases, respectivement appelées 12RSS et 23RSS (Akira et al., 1987; McBlane et al., 1995). Le complexe RAG initie la recombinaison V(D)J durant la phase G1 en se liant à une RSS et en introduisant des cassures double brin (DSBs, de l'anglais Double strand breaks) entre les segments de gènes V, D et J et les RSSs. La formation efficace de DSBs requiert l'assemblage de paires 12/23 RSS, appelé règle 12/23 (Alt et al., 2013; Schatz and Swanson, 2011). C'est-à-dire qu'au sein du locus IgH, les segments de gènes V_H et J_H sont adjacents à des séquences 23RSS alors que les segments de gènes D_H sont adjacents à des séquences 12RSS. Cette règle permet la recombinaison en 2 étapes avec en premier lieu la recombinaison entre un segment D_H et J_H puis entre V_H et D_H-J_H. Cet assemblage a lieu en complexe synaptique avec les protéines RAG (**Figure 4**).

Les cassures double brin induites sont rapidement reconnues par la voie de réponse aux dommages de l'ADN (DDR, de l'anglais DNA damage response) formant un complexe de post-clivage. Celui-ci regroupe les protéines RAG1/2 qui sont restées associées aux extrémités, codante et « signal », et les protéines de la DDR dont la protéine kinase ATM (ataxia telangiectasia mutated). Celles-ci vont permettre de stabiliser les extrémités de l'ADN au sein du complexe de post-clivage. Les protéines RAG participent à la religation des extrémités via la voie de réparation des extrémités non-homologues ou NHEJ (non-homologous end-joining). Ce qui aboutit à un joint codant et un joint « signal » qui va être éliminé sous forme d'épisome circulaire (Alt et al., 2013).

La combinaison aléatoire des différents gènes V, D et J génère un répertoire d'anticorps primaires varié. Celui-ci subit une diversification additionnelle induite par l'ADN polymérase Terminale Transférerase (TdT) qui introduit des nucléotides sans l'aide d'ADN matrice, au niveau des jonctions codantes D-J et V-DJ. Des

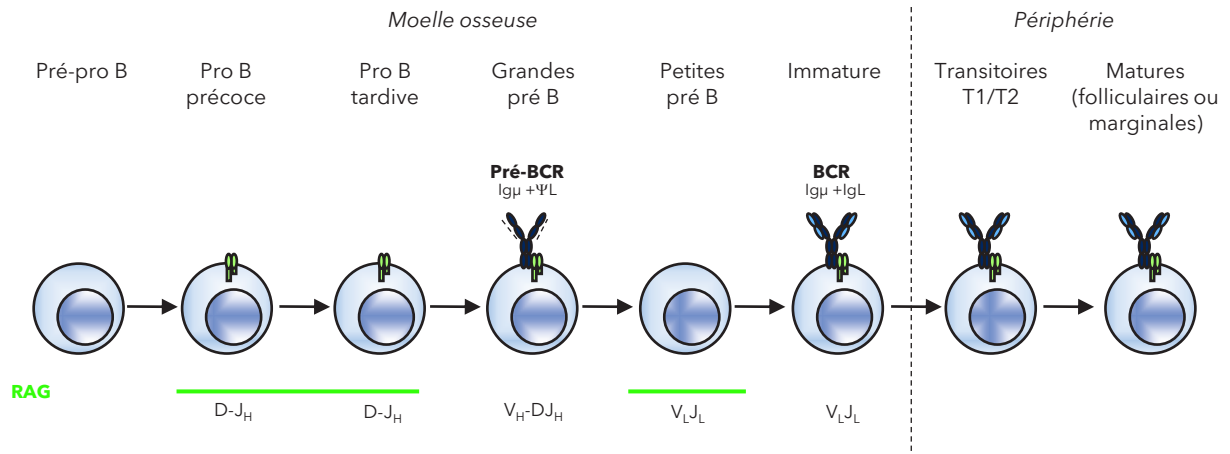


Figure 5: Le développement des cellules B. Les premiers précurseurs de cellules B engagés sont des cellules pré-pro-B. RAG est exprimé dans les cellules pro-B et provoque la recombinaison de la chaîne lourde IgH, dans l'ordre $D-J_H$ et V_H-DJ_H . Au stade pré-B, le pré-BCR, composé d'une chaîne lourde qui a eu un réarrangement productif $Ig\mu$ et d'une chaîne légère de substitution ΨL , est exprimé à la surface de la cellule. Ces cellules pré-B engagent donc la recombinaison du locus de la chaîne légère V_LJ_L pour exprimer un BCR fonctionnel au stade immature. Les cellules B vont ensuite migrer en périphérie où elles poursuivent leur maturation en cellules B folliculaires ou marginales. Adapté de (Meffre et al. 2000).

nucléotides sont également éliminés des séquences jonctionnelles ce qui constitue avec l'introduction de nucléotides, une source majeure de diversification du répertoire primaire (Bassing et al., 2002).

La recombinaison V(D)J est régulée par l'expression des protéines RAG qui sont hautement exprimées dans les stades précoces du développement des lymphocytes, par l'accessibilité aux RSS dans la chromatine mais aussi par la position et l'architecture tridimensionnelle des loci des récepteurs d'Ag dans le noyau, permettant ensemble une recombinaison entre des segments de gènes très espacés.

Les différentes étapes de la recombinaison V(D)J correspondent aux étapes précoces du développement des lymphocytes (**Figure 5**). En effet, la recombinaison V(D)J est un mécanisme très ordonné, dans lequel l'évènement de recombinaison initial a lieu au sein du locus IgH où les réarrangements D_H-J_H se produisent en premier, suivis par les réarrangements V_H-DJ_H (Alt et al., 1984). Un réarrangement VDJ productif sur un allèle va générer une chaîne Ig μ . L'association d'une chaîne Ig μ avec une chaîne IgL non réarrangée forme un récepteur pré-BCR à la surface de la cellule (Melchers et al., 2000; Rajewsky, 1996) La présence du pré-BCR à la surface de la cellule va donner un signal pour empêcher le réarrangement de l'autre allèle s'il est en configuration DJ_H et favoriser le début de la recombinaison V(D)J des IgL (Cobb et al., 2006). Ce processus permettant l'arrêt de la recombinaison V(D)J des IgH et assurant l'expression d'une seule chaîne Ig μ est appelé l'exclusion allélique et compte parmi de nombreux points de contrôles établis durant le développement des cellules B. La recombinaison V(D)J efficace des loci IgH et IgL ainsi que le recrutement de l'hétérodimère Ig α /Ig β permettent aux cellules B d'exprimer un BCR fonctionnel, devenues cellules B naïves.

La spécificité du BCR généré est ensuite testée pour sa reconnaissance envers des molécules du soi (auto-Ag liés à la membrane). Les cellules B auto-réactives vont subir des réarrangements secondaires sur la chaîne des IgL, mécanisme appelé édition du récepteur, afin de générer un nouveau récepteur de l'antigène avec une spécificité inoffensive envers le soi (Halverson et al., 2004). Les cellules B auto-réactives, qui ont échappé au mécanisme de l'édition du récepteur, subissent un

second mécanisme, appelé la délétion clonale, qui va éliminer ces cellules par apoptose (Meffre and Wardemann, 2008; Nemazee, 2017; Rajewsky, 1996). Ces deux mécanismes de tolérance centrale permettent aux cellules B dans la moelle osseuse de maintenir une spécificité unique en exprimant un récepteur ayant une reconnaissance robuste des Ag et une haute spécificité, et de contribuer à l'élaboration du répertoire des cellules B matures. Suite à ces mécanismes, les cellules B tolérantes aux auto-Ag vont migrer vers les organes lymphoïdes secondaires afin de devenir matures et d'exprimer un BCR de spécificité unique de type IgM et/ou IgD. Des défauts dans ces mécanismes de tolérance peuvent être impliqués dans la pathogénèse de maladies auto-immunes et certains désordres d'immunodéficiência.

3. La diversification dépendante de l'antigène

Les cellules B qui ont réalisées avec succès la recombinaison V(D)J et qui expriment un BCR fonctionnel migrent comme cellules B naïves de la moelle osseuse vers la périphérie afin de peupler les organes lymphoïdes secondaires.

Lors de la rencontre avec un antigène étranger, les cellules B naïves sont activées par l'interaction avec les cellules T CD4+ dans la zone riche en cellules T des tissus lymphoïdes. Puis, elles prolifèrent de manière massive et diversifient leur BCR par les mécanismes d'hypermutation somatique, de commutation isotypique et de conversion génique qui se produisent majoritairement dans les centres germinatifs.

3.1. Les centres germinatifs

Les centres germinatifs (GC, de l'anglais germinal centers) sont des structures micro-anatomiques dans le centre des follicules des cellules B dans lesquels les cellules B activées par l'Ag prolifèrent et diversifient leur BCR. Ces structures ont été initialement décrites comme des compartiments organisés contenant des cellules en cours de division (Nieuwenhuis and Opstelten, 1984). Les centres germinatifs se forment uniquement en réponse à l'interaction spécifique à l'antigène entre le récepteur des cellules T et le peptide du CMHII présent sur la cellule B. Les cellules B des GCs ont la particularité d'exprimer fortement le facteur de transcription Bcl6. Un centre germinatif est fonctionnellement polarisé en deux zones : une zone

sombre (DZ, de l'anglais dark zone) qui contient exclusivement des cellules B en prolifération massive, appelée centroblastes, qui subissent la réaction d'HMS et d'une zone claire (LZ, de l'anglais light zone) dans laquelle les cellules B, appelées centrocytes, sont sélectionnées sur la base de l'affinité de leur BCR pour l'antigène (Victoria et al., 2010). Le passage de la zone sombre à la zone claire et inversement est contrôlé par un réseau complexe de signaux cellulaires qui affectent les réponses des cellules B du GC en activant et en réprimant des programmes de transcription spécifique.

Dans le GC, les cellules B transitent de la zone sombre, où elles prolifèrent et subissent l'HMS vers la zone claire où elles sont entremêlées entre des cellules dendritiques folliculaires (FDCs), présentatrices d'Ag, et des cellules T auxiliaires folliculaires (Tfh, de l'anglais T follicular helper) (De Silva and Klein, 2015; Victoria and Nussenzweig, 2012).

Les cellules B possédant un BCR d'une affinité supérieure subissent des cycles répétés de mutations et de sélection clonale par un processus de ré-entrée cyclique de l'état LZ à DZ. Ces microenvironnements permettent la génération de cellules B mémoires et de cellules plasmablastes qui se différencient en plasmocytes, cellules productrices d'Ac de haute affinité, qui sont nécessaires pour se protéger face aux micro-organismes (Basso and Dalla-Favera, 2015). Durant le processus de sélection clonale dans la LZ, les FDCs retiennent l'antigène et coopèrent avec les Tfh pour soutenir les cellules B compétitives du GC ayant un BCR d'une affinité supérieure (Bannard and Cyster, 2017; Ramezani-Rad and Rickert, 2017). L'aide apportée par les cellules Tfh durant ce processus est requise pour la survie des centrocytes et initie la génération des cellules plasma (Victoria et al., 2010; Zotos and Tarlinton, 2012).

Cependant, les répétitions de cycles de mutations/sélection peuvent accroître la susceptibilité à la transformation maligne. En effet, la majorité des lymphomes B sont originaires des GCs (Küppers, 2005; Küppers et al., 1999).

3.2. L'hypermutation somatique

L'hypermutation somatique introduit des mutations ponctuelles et occasionnellement des petites insertions et délétions dans les régions variables des

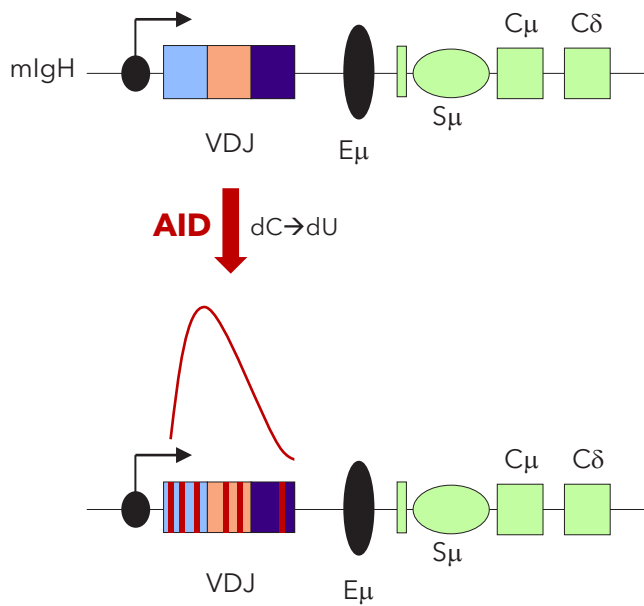


Figure 6: L'hypermutation somatique. L'hypermutation somatique modifie la partie variable réarrangée (V-D-J) des chaînes IgH et IgL en y introduisant des mutations ponctuelles (traits rouges). La fréquence de mutations, qui est plus élevée à proximité du promoteur, est représentée par la courbe rouge.

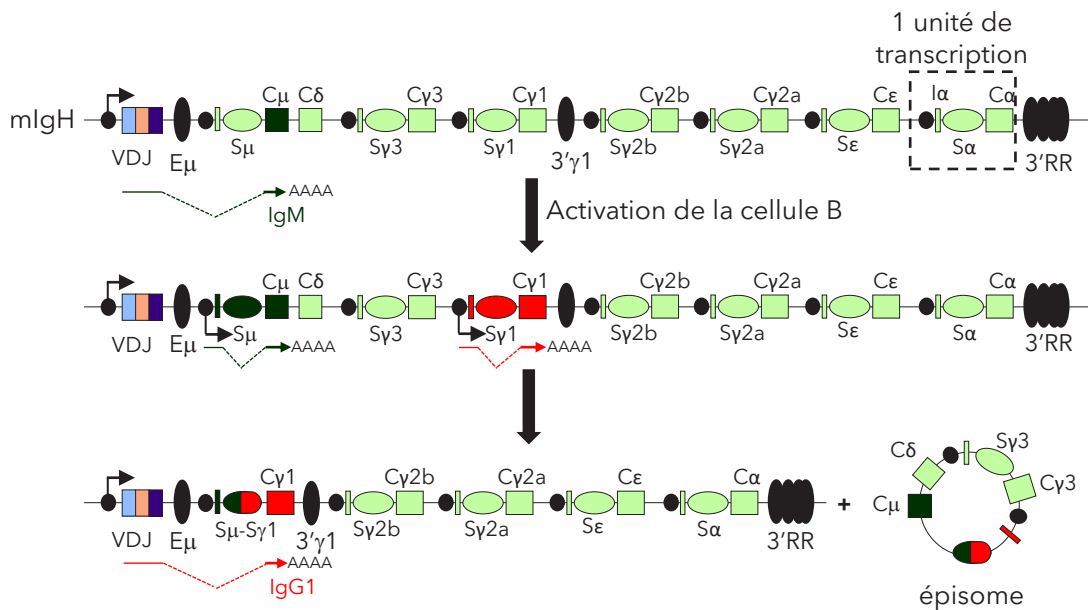


Figure 7: La commutation isotypique. La commutation isotypique est une réaction de recombinaison région-spécifique qui a lieu au sein du locus IgH. Les gènes codants pour la région constante sont organisés en unité transcriptionnelle, constituée, d'un promoteur sensible au cytokines (point noir), d'un exon intronique (Ix), d'une région de switch (Sx) et d'un exon constant (Cx). La recombinaison implique les régions S qui sont des séquences répétitives et non-homologues. Ici, la recombinaison implique la région donneuse S_{μ} et acceptrice $S_{\gamma 1}$ qui va entraîner l'excision de la région C_{μ} et son remplacement par la région $C_{\gamma 1}$ pour aboutir à l'expression d'un nouvel isotype IgG1. Les séquences intermédiaires sont éliminées sous forme d'épisome circulaire.

chaînes lourdes et légères des Ig (IgV) (**Figure 6**). Ces mutations ont lieu à une fréquence de 10^{-3} par paire de bases par génération (Di Noia and Neuberger, 2007) dans un domaine qui commence à quelque 150 pb en aval du promoteur IgV et s'étend sur environ 2kb (Lebecque and Gearhart, 1990). Ce mécanisme qui modifie la région déterminant la complémentarité de l'Ac pour l'Ag (CDRs) (Betz et al., 1993; Rogozin and Kolchanov, 1992; Sharpe et al., 1991) des IgV donne naissance à un répertoire d'anticorps de très haute affinité pour l'Ag suite à de multiples séries de mutations et sélections dans les centres germinatifs, processus appelé maturation d'affinité. Le mécanisme d'HMS sera décrit plus en détail dans la partie III.

3.3. La commutation isotypique

La commutation isotypique (CI) modifie les fonctions effectrices des Ac en induisant le changement de l'isotype exprimé (de IgM vers IgG, IgE ou IgA), tout en maintenant son affinité pour l'antigène. La CI est une réaction de recombinaison de longue distance région-spécifique qui se produit au sein du locus IgH entre deux régions de « switch » où des cassures double brin sont générées (**Figure 7**). Ces régions de « switch », aussi appelées régions S sont des régions répétées non-homologues de 3 à 12 kb, localisées en amont de chaque région C codant pour les différents isotypes (excepté IgD). La recombinaison entre la région S dite « donneuse » S_{μ} avec une région S « acceptrice » S_x mène à l'excision de la région C_{μ} et son remplacement par une région constante C_x . La CI est une recombinaison d'ADN par délétion car les séquences intermédiaires sont éliminées en épisode circulaire. L'étude de ce mécanisme étant le sujet principal de mes travaux de thèse, la CI sera donc décrite plus en détail dans la partie IV.

Chaque isotype détermine la manière dont les antigènes capturés sont éliminés ou neutralisés et la localisation où l'Ac est délivré ou accumulé. Les premiers anticorps produits au début des réponses immunes sont les IgM. Chez les mammifères, les IgM sont sécrétées sous forme de monomères puis vont former principalement des pentamères qui possèdent 10 sites de liaison à l'Ag, ce qui leur confère une avidité globale plus élevée. Cette avidité plus élevée de la forme pentamère permet de compenser l'affinité relativement faible des monomères IgM. Cette structure pentamérique rend les IgM capables d'activer efficacement le complément mais

restreint leur présence principalement dans le sang. A l'inverse, les IgG sont sous forme monomérique et diffusent facilement du sang vers les tissus. L'IgG est l'isotype le plus abondant dans le sérum mais se trouve également dans les espaces extracellulaires des tissus. Les IgG passent la barrière du placenta pour fournir une protection maternelle au fœtus. Cet isotype est impliqué dans la neutralisation et la destruction de pathogènes par l'opsonisation de ceux-ci, ce qui favorise la phagocytose, et l'activation du complément (Nimmerjahn and Ravetch, 2008). Les IgA sont abondantes dans les muqueuses et les sécrétions, comme les larmes et la salive et peuvent être sous forme de dimères. Cet isotype permet d'enrober les pathogènes pour empêcher leur prolifération et est capable de neutraliser les toxines et les virus. Les IgA ont aussi la capacité de traverser les membranes épithéliales (Macpherson et al., 2008; Woof and Kerr, 2006). L'IgE est l'isotype le moins présent dans le sérum mais se trouve au niveau des épithéliums et des muqueuses. Cet isotype permet la lutte contre les nématodes parasites. Il est aussi à l'origine des réactions d'hypersensibilités immédiates par sa liaison avec des basophiles et mastocytes qui activent les relâchements d'histamines (Oettgen, 2000; Wu and Zarrin, 2014).

La commutation isotypique a longtemps été décrite comme un mécanisme se produisant principalement dans les centres germinatifs (Liu et al., 1996; Shinkura et al., 1996), puis pouvant aussi avoir lieu en dehors des GCs durant les réponses dépendantes et indépendantes des cellules T (Fagarasan et al., 2001; Jacob et al., 1991; Pape et al., 2003; Toellner et al., 1998; Toellner et al., 1996). Mais, une étude récente (Roco et al., 2019) remet en question la croyance générale selon laquelle la CI est essentiellement un processus du GC et révèle que la CI se produit avant la formation des GCs, et plus précisément, au centre pré-germinal ou dans les premiers centres germinaux, et diminue lors de la création du GC. Cependant, l'idée selon laquelle la CI se produit au sein des GCs est encore très présente et cela, pour plusieurs raisons. Premièrement, la CI ainsi que l'HMS sont deux mécanismes dépendants de l'activité de l'enzyme AID (Activation cytidine deaminase) (Muramatsu et al., 2000), qui est exprimée en quantité maximale dans les GCs. Deuxièmement, la démonstration de l'association entre la CI et les centres germinatifs provient d'une étude menée sur des souris dépourvues de GCs qui a

renforcé l'idée que la CI était un mécanisme de GC (Shinkura et al., 1996) tandis que ces souris présentaient également des altérations au niveau des rencontres précoces entre les cellules T et les cellules B nécessaires à la CI (Toellner et al., 1996). Enfin, une autre étude a suggéré que la CI se produit majoritairement dans les GCs en observant une expression stérile sélective $I\gamma$, $I\alpha$ et $I\epsilon$ dans des centrocytes (identifiés CD77-) sur la base du marqueur CD77 (Liu et al., 1996). Ce marqueur CD77 a finalement été décrit quelques années plus tard comme un marqueur ne pouvant discriminer les centroblastes et les centrocytes et que la population CD77- représente un sous-ensemble hétérogène de cellules, comprenant des centroblastes, des centrocytes et des plasmablastes (Hogerkorp and Borrebaeck, 2006).

Cette étude récente (Roco et al., 2019) montre que l'expression de transcrits stériles $\gamma 1$ est détectée 1.5 jours après l'immunisation, dans des cellules localisées à la frontière cellules T : B (Bergqvist et al., 2010; Toellner et al., 1998; Toellner et al., 1996), ce qui révèle que la CI a lieu avant la formation des GCs et qu'elle est donc peu fréquente dans les GCs. Ils montrent aussi par des analyses phylogénétiques que la plupart des cellules B ont subi un changement de classe avant d'accumuler les mutations de la région V et confirme donc que la CI précède le mécanisme d'HMS.

3.4. La conversion génique

Chez certaines espèces, particulièrement chez les animaux de la ferme, la formation du répertoire secondaire d'anticorps se réalise par conversion génique des gènes IgV, mécanisme découvert chez le poulet (Reynaud et al., 1985). Ce mécanisme de diversification a lieu dans les organes lymphoïdes primaires et plus précisément dans la bourse de Fabricius chez le poulet (Reynaud et al., 1987; Reynaud et al., 1989). Le locus IgL du poulet affiche un unique exon $V\lambda$ qui est recombinaison avec un seul segment génique $J\lambda$. Également, le locus IgH compte un unique exon V et J ainsi que 15 gènes D fonctionnels en supplément (Reynaud et al., 1989). La conversion génique permet d'apporter de la diversité dans le répertoire primaire limité obtenu par recombinaison V(D)J par l'intermédiaire d'une recombinaison avec un ensemble d'éléments pseudogènes se trouvant en amont de la région IgV

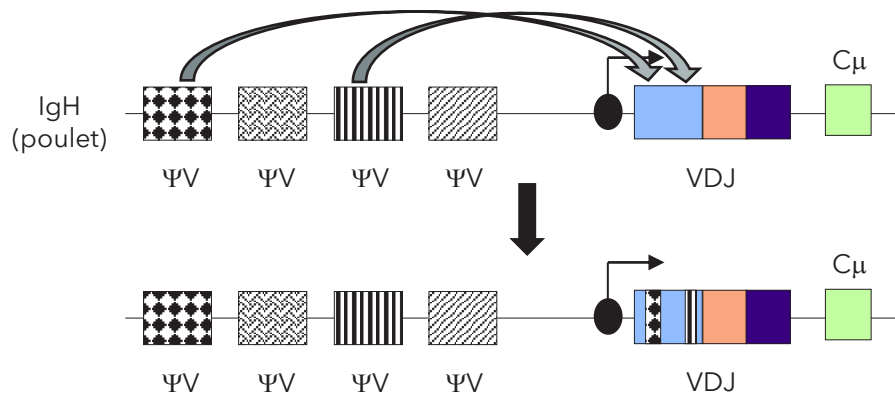


Figure 8: La conversion génique. La conversion génique est un mécanisme de diversification des anticorps qui a seulement lieu chez certaines espèces comme le poulet ou bien le lapin. La conversion génique repose sur l'utilisation de pseudogènes ΨV pour diversifier la région IgV réarrangée V-D-J. Adapté de (Maizels et al., 2005; Tang et al., 2007)

réarrangée (Reynaud et al., 1985; Reynaud et al., 1987). Au cours de cette conversion génique, des blocs de séquences aléatoires de ces pseudogènes sont introduits dans la région IgV réarrangée causant des lésions réparées par recombinaison homologue en utilisant, comme matrice, les pseudogènes situés en amont (**Figure 8**). Le répertoire d'anticorps peut être à nouveau diversifié lors des réponses immunes par les réactions d'HMS et de conversion génique dans les centres germinatifs (Arakawa et al., 1996).

Ces trois mécanismes sont dépendants de l'enzyme AID (Activation-Induced cytidine Deaminase), une enzyme identifiée il y a une vingtaine d'année par Tasuku Honjo (Muramatsu et al., 2000; Muramatsu et al., 1999) chez la souris et Anne Durandy (Revy et al., 2000) chez l'homme. Cette identification a représenté une percée dans la compréhension des mécanismes d'hypermutation somatique, de commutation isotypique (Muramatsu et al., 2000; Muramatsu et al., 1999; Revy et al., 2000), et de conversion génique (Arakawa, 2002).

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

1. La découverte d'AID

Afin de comprendre le mécanisme moléculaire et les acteurs principaux de la réaction de commutation isotypique, une lignée clonale de lymphome B murin réalisant la CI de IgM vers IgA, appelée lignée cellulaire CH12, a été établie (Bishop and Haughton, 1986; Nakamura et al., 1996). Celle-ci peut réaliser efficacement la commutation isotypique de IgM vers IgA en étant stimulée avec du TGF β , IL-4 et CD40L.

L'enzyme Activation Induced cytidine Deaminase (AID) a été identifiée (Muramatsu et al., 1999) au début des années 2000 par l'équipe de Tasuko Honjo via un criblage par hybridation soustractive d'ADNc réalisé dans la lignée CH12 en comparant des cellules non stimulées et stimulées. De plus, AID est hautement exprimée *in vivo* dans les cellules B des centres germinatifs où les mécanismes de diversification ont lieu (C. Berek, 1991; Liu et al., 1989). Puis, la caractérisation de patients affectés du syndrome hyper-IgM de type 2 (HIGM2), immunodéficiência due à l'impossibilité de réaliser les mécanismes de diversification, montre que ce sont des mutations dans le gène codant pour AID (Revy et al., 2000) qui sont à l'origine de l'abolition de la réaction de la CI et d'HMS et donc de l'immunodéficiência. Ensuite, la caractérisation de souris déficientes pour AID montre l'absence totale de réactions de CI et d'HMS chez ces animaux (Muramatsu et al., 2000). Ensemble et appuyées par d'autres travaux, ces études dévoilent qu'AID est l'unique facteur spécifique aux cellules B, requis pour initier les réactions d'hypermutation somatique (Martin et al., 2002; Yoshikawa et al., 2002) et de commutation isotypique (Okazaki et al., 2002). Enfin, AID est décrite comme un facteur requis pour la conversion génique dans la lignée cellulaire B de poulet, appelée DT40 (Arakawa and Buerstedde, 2004; Harris et al., 2002). Elle est donc l'enzyme requise pour initier les trois mécanismes de diversification des Ac.

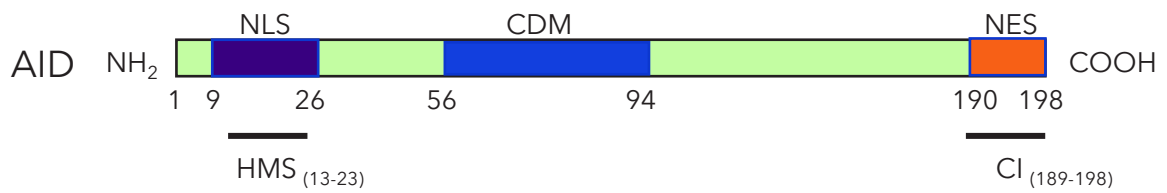


Figure 9: Les domaines fonctionnels de la protéine Activation-Induced cytidine Deaminase (AID). AID possède à son extrémité N-terminale un signal de localisation nucléaire (NLS), représenté par le rectangle violet, ainsi qu'un domaine requis pour l'hypermutation somatique (HMS). AID contient également 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) est compris entre les acides aminés 56 et 94). Les chiffres indiqués représentent les positions des acides aminés délimitant un domaine.

2. La structure et la fonction d'AID

AID est une petite protéine de 198 acides aminés qui abrite un signal de localisation nucléaire (NLS, de l'anglais nuclear localization signal) en N-ter, un signal d'export nucléaire (NES, de l'anglais nuclear export signal) en C-ter et un site catalytique (résidus 56-94) lui attribuant l'activité enzymatique de désamination des cytosines en uraciles (**Figure 9**).

De nombreuses études visant à clarifier les domaines fonctionnels d'AID durant les mécanismes de diversification des Ac ont montré que les signaux NLS et NES ne déterminent pas seulement sa localisation sous-cellulaire (Ito et al., 2004; McBride et al., 2004; Patenaude et al., 2009). En effet, la génération de cinq mutations ponctuelles produisant des changements dans la région N-ter d'AID ont détruit l'activité d'HMS, sans affecter l'activité de commutation isotypique (Shinkura et al., 2004). Tandis que la délétion de 10 acides aminés du domaine C-ter d'AID (AID^{Δ189-198}) abolit la réaction de CI, sans affecter l'HMS (Barreto et al., 2003; Ta et al., 2003). Ces résultats révèlent, tout d'abord, le découplage des réactions d'HMS et de CI mais aussi que la fonction d'AID dans l'HMS et la CI ne se limite pas à la désamination des cytosines. Le nombre et la distribution normale de mutations S_μ observés avec la version tronquée AID^{Δ189-198} exposent néanmoins, que la région C-ter n'est pas essentielle pour le ciblage d'AID aux régions S (Barreto et al., 2003). Cependant, la région C-ter d'AID pourrait jouer un rôle de plateforme permettant de recruter des cofacteurs spécifiques à la CI (Okazaki et al., 2002; Ta et al., 2003; Yoshikawa et al., 2002). Également, le domaine N-ter d'AID a un rôle de liaison à des cofacteurs spécifiques à l'HMS (Shinkura et al., 2004; Ta et al., 2003).

Ayant une structure et une activité enzymatique similaire, AID appartient à la famille des désaminases APOBEC (apolipoprotein B mRNA editing enzyme catalytic polypeptide-like) qui inclut APOBEC1, APOBEC2 et APOBEC3. Les enzymes APOBECs désaminent les cytosines dans l'ARN et/ou les résidus cytosines dans l'ADN et régulent divers mécanismes. AID a été originellement proposée comme une enzyme d'édition de l'ARN (Muramatsu et al., 1999) due à l'homologie de sa séquence avec son paralogue, APOBEC1, qui est une enzyme d'édition de l'ARN. En revanche, de nombreuses études génétiques et biochimiques ont prouvé qu'AID

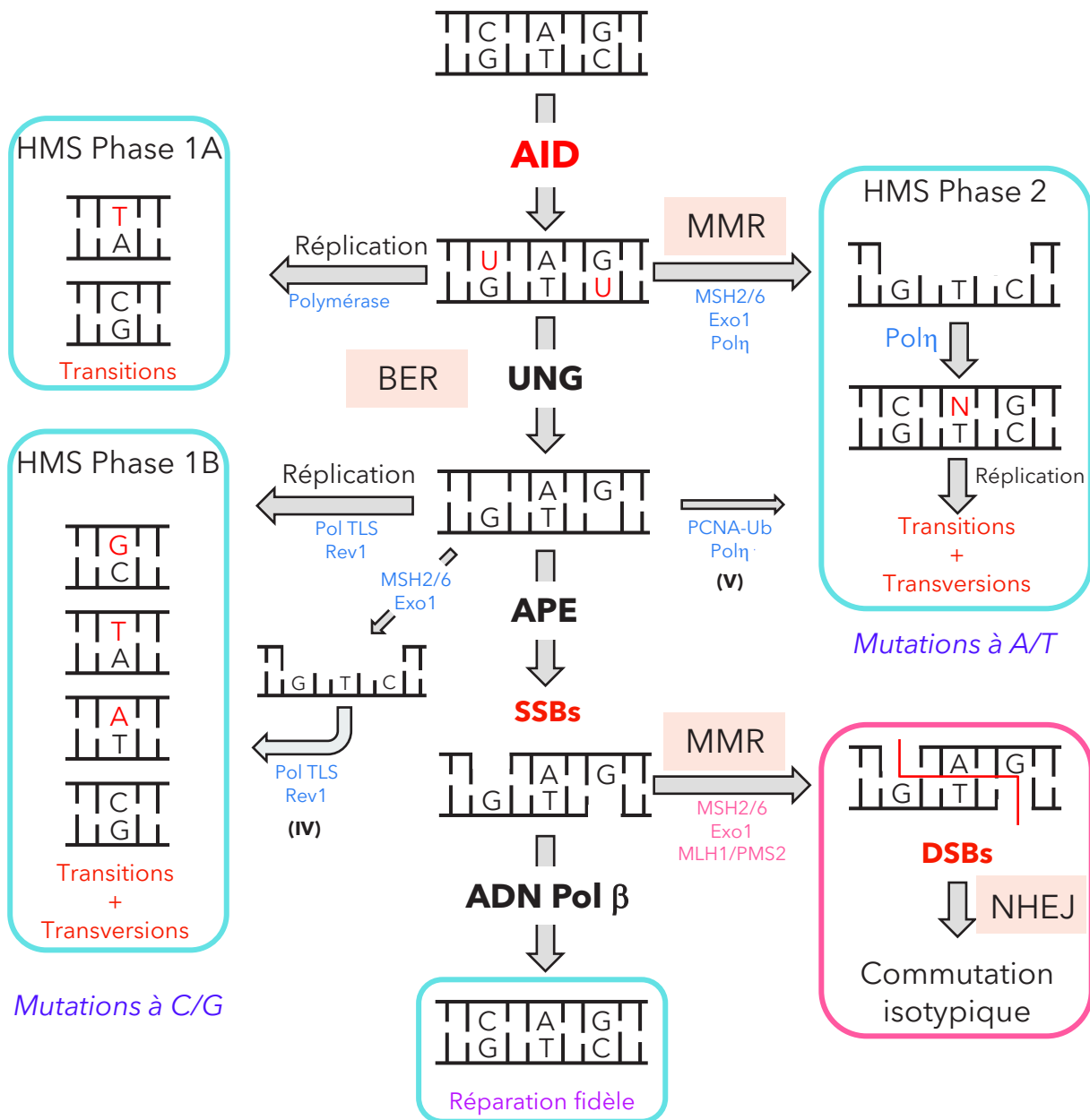


Figure 10: Le modèle de désamination de l'ADN par AID. AID convertit des cytosines (C) en uraciles (U) dans l'ADN des régions V et des régions S des loci Ig. La réplication sans réparation préalable des U générés peut entraîner l'apparition de mutations de type transitions (HMS phase 1A). Les U générés 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. Celle-ci 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 opposé est également présente. Cette DSB est prise en charge par le NHEJ pour réaliser la recombinaison nécessaire à la commutation isotypique.

agit comme une désaminase désoxycytidine spécifique à l'ADN (Bransteitter et al., 2003; Dickerson et al., 2003; Fritz et al., 2013; Petersen-Mahrt et al., 2002).

3. Le modèle de désamination de l'ADN

Le modèle de désamination de l'ADN suggéré par le groupe de M. Neuberger révèle pour la première fois qu'AID désamine les cytosines en uraciles dans l'ADN (Petersen-Mahrt et al., 2002) (**Figure 10**). En effet, ils ont montré que la surexpression d'AID chez *E. coli* induit des mutations de transition à dC : dG dans l'ADN et que les mutations provoquées par AID sont augmentées par la déficience de l'enzyme UNG (Uracil DNA Glycosylase). 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 aux DSBs et aux mutations pour la CI et l'HMS respectivement (Petersen-Mahrt et al., 2002; Rada et al., 2002b). Puis, de nombreuses études *in vitro* ont révélé qu'AID désamine les cytosines en uraciles spécifiquement dans l'ADN simple brin, dévoilant que l'activité d'AID est étroitement couplée à la transcription afin de faire apparaître l'ADN simple brin (Bransteitter et al., 2003; Chaudhuri et al., 2003; Dickerson et al., 2003; Pham et al., 2003; Ramiro et al., 2003; Sohail, 2003; Yu et al., 2004). Enfin, des expériences d'immunoprécipitation de la chromatine ont démontré qu'AID s'associe aux régions S durant la CI (Nambu et al., 2003) via son interaction avec de nombreuses cibles génomiques, telles que RPA et Spt5 (Pavri et al., 2010; Yamane et al., 2011).

La désamination médiée par AID convertit les cytosines en uraciles, générant ainsi des mésappariements dU : dG dans l'ADN. Ces mésappariements sont reconnus et traités par différentes voies de réparation de l'ADN incluant la voie de la réparation de l'excision de base (BER, de l'anglais base excision repair) et la voie de la réparation des mésappariements (MMR, de l'anglais mismatch repair) qui sont détournées afin de générer des mutations dans les régions V des Ig durant l'HMS et/ou des cassures double brin dans les régions S du locus IgH durant la CI.

4. Le potentiel oncogénique d'AID et la régulation de l'activité d'AID

4.1. Le potentiel oncogénique d'AID

Bien que les régions S et les régions V des Ig soient des cibles physiologiques d'AID durant les mécanismes de diversification des anticorps, AID peut générer des DSBs et des 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). Malgré un taux nettement plus faible de formation de DSBs et de mutations au niveau de ces gènes non-Ig (Yamane et al., 2011), les translocations ou mutations dans ces gènes hors-cibles contribuent au développement de lymphomes de cellules B matures (Nussenzweig and Nussenzweig, 2010; Pasqualucci et al., 2008; Rocha and Skok, 2013). Ainsi, des mécanismes permettant le ciblage d'AID spécifiquement aux loci Ig pour promouvoir la CI et l'HMS, tout en limitant l'accès à AID au reste du génome des cellules B sont nécessaires. La dérégulation de ces mécanismes, de l'expression d'AID et la réparation aberrante de ses lésions favorisent son potentiel oncogénique.

Chez l'homme, 95% des lymphomes proviennent des cellules B. De nombreuses études visant à définir l'origine cellulaire des lymphomes B humains ont montré que la majorité de ceux-ci sont dérivés de cellules B du centre germinatif (lymphome de Burkitt, lymphome hodgkinien et non hodgkinien, lymphome diffus à grandes cellules (DLBCL)) ou post-GC (DLBCL, myélome multiple, leucémie lymphocytaire chronique (CLL)) (Seifert et al., 2013). Ces lymphomes se caractérisent par des translocations chromosomiques impliquant les loci Ig et des proto-oncogènes, comme c-myc, Bcl2 ou Bcl6. Ces translocations sont les sous-produits des processus de diversification des anticorps : la recombinaison V(D)J, l'HMS et la CI.

4.1.1. Les mutations hors-cibles

AID peut participer à la lymphomagenèse par l'introduction de mutations somatiques au niveau des gènes non-Ig. Des lymphomes de cellules B matures présentent fréquemment des mutations dans de multiples proto-oncogènes

(Pasqualucci et al., 2001). Celles-ci ont été montrées dépendantes de la transcription, centrées sur les motifs WRC et hautement biaisées en paires G/C, ce qui correspond à la signature exacte des mutations induites par AID durant l'HMS. Ces mutations ont été reportées au niveau de proto-oncogènes comme Bcl6 (Gaidano et al., 1997; Migliazza et al., 1995; Pasqualucci et al., 1998; Shen et al., 1998), Fas (Muschen et al., 2000), et les constituants du BCR, CD79a et CD79b (Gordon et al., 2003) dans des cellules B tumorales (Gaidano et al., 1997; Migliazza et al., 1995) et normales (Pasqualucci et al., 1998; Shen et al., 1998). Cependant, il a aussi été montré que des gènes comme Pim1, c-myc, RhoH et Pax5 sont seulement mutés dans des cellules B tumorales issues de DLBCLs (Pasqualucci et al., 2001), indiquant qu'il existe deux niveaux de protection pour le génome des cellules B durant l'HMS (Liu et al., 2008). Le spectre des cibles génomiques d'AID a été révélé par ChIP-seq, confirmant qu'AID peut être recrutée à de nombreux gènes non-Ig dans des cellules B activées non tumorales (Yamane et al., 2011). Ceci montre que la capacité d'induction de mutations sur l'ensemble du génome d'AID, durant les mécanismes de diversification des Ig, peut menacer l'intégrité du génome. Une étude récente a révélé une collection de 275 gènes pouvant être ciblés par AID dont 21 sont déjà annotés comme mutés dans les DLBCLs, tels que BCL6, RHOH, PIM1 et PAX5 (Alvarez-Prado et al., 2018). Alvarez-Prado et al. ont ainsi identifié 9 nouveaux gènes mutés dans des DLBCLs humains qui accumulent les mutations induites par AID tels que MEF2B, LYN, TNFAIP3, GNA13 ET IRF8.

4.1.2. Les translocations

Bien avant la découverte d'AID, des études génétiques révélaient des translocations chromosomiques impliquant des gènes Ig dans des lymphomes B humains (Küppers and Dalla-Favera, 2001). Une réparation aberrante des DSBs induites par AID peut favoriser des translocations chromosomiques réciproques entre un proto-oncogène et les régions S ou V du locus IgH qui sont la marque de nombreux types de lymphomes à cellules B. En effet, par exemple, les lymphomes de Burkitt sont caractérisés par la présence de translocations IgH/c-myc tandis que les lymphomes folliculaires sont caractérisés par des translocations IgH/Bcl2 (Küppers, 2005). Il a été montré que dans les translocations IgH/c-myc, les cassures au niveau de c-myc

sont initiées par AID (Ramiro et al., 2004; Robbiani et al., 2008) et requièrent, comme dans la CI, le traitement des mésappariements U:G par UNG (Ramiro et al., 2006) pour la génération de la translocation. D'autre part, il a été montré que la fréquence d'HMS est augmentée en l'absence combinée d'Ung et de Msh2, indiquant que le BER et le MMR se soutiennent mutuellement pour réparer fidèlement la plupart des lésions induites par AID dans les cellules B du GC (Alvarez-Prado et al., 2018; Liu et al., 2008; Methot and Di Noia, 2017; Pérez-Durán et al., 2012). De plus, des études plus récentes montrent un rôle protecteur d'UNG face à l'évolution des cellules B à la transformation maligne (Delgado et al., 2020; Safavi et al., 2020) et que l'activité d'HMS d'AID peut, par elle-même, conduire ou contribuer à la transformation des cellules B (Delgado et al., 2020), déjà démontrée par de nombreuses études (Loeffler et al., 2015; Pasqualucci et al., 2014; Swaminathan et al., 2015). Enfin, il semble que les mécanismes conduisant à la transformation maligne des cellules B ne soient pas entièrement compris et que de nombreuses questions restent encore à élucider.

4.1.3. L'expression ectopique d'AID

La dérégulation de l'expression d'AID a un lien étroit avec la tumorigenèse. En effet, son expression ubiquitaire et constitutive dans un modèle murin mène au développement de lymphomes à cellules T, à des micro-adénomes pulmonaires et des adénocarcinomes mais ne développe pas de lymphomes B (Okazaki et al., 2003). La dérégulation de l'expression d'AID seule n'est pas suffisante pour promouvoir des lymphomes à cellules B mais conduit à d'importantes mutations somatiques dans les gènes non-Ig (Delgado et al., 2020; Robbiani et al., 2009). De plus, la surexpression d'AID spécifiquement dans les cellules B ne produit pas de lymphomes à cellules B (Muto et al., 2006). Cependant, l'expression aberrante d'AID peut aussi être associée à de nombreux cancers tels que les cancers de l'estomac (Matsumoto et al., 2007), du pancréas (Sawai et al., 2015), du foie (Endo et al., 2007; Kou et al., 2007; Matsumoto et al., 2015) ou du colon (Endo et al., 2008) ainsi qu'à des infections bactériennes de *H. pylori* par exemple (Marusawa, 2008; Matsumoto et al., 2007) ou virales (Endo et al., 2007). Il semble que l'inflammation chronique liée à ces infections soit la cause de l'induction de l'expression d'AID dans des

cellules non B. Enfin, AID est indirectement impliquée dans la pathogenèse des lymphomes car elle est exprimée dans certaines formes de lymphomes, notamment le DLBCL (Lenz et al., 2007).

4.1.4. *Kataegis*

Une nouvelle signature de mutations a été révélée par le séquençage génomique d'un large éventail de tumeurs humaines (Alexandrov et al., 2013; Nik-Zainal et al., 2012), soulignant ainsi une nouvelle activité hors cible pour les enzymes AID dans les loci Ig (Alexandrov et al., 2013; Lada et al., 2012) et APOBEC3 dans les loci non-Ig (Alexandrov et al., 2013; Burns et al., 2013; Lada et al., 2012; Taylor et al., 2013). Ces nombreux groupes de mutations coordonnées se trouvant dans des petites régions génomiques <10kb, appelés *kataegis* autrement dit « tempêtes de mutations », présentent plusieurs caractéristiques qui les distinguent des substitutions aléatoires, dispersées à travers le génome (Alexandrov et al., 2013; Nik-Zainal et al., 2012). Ces mutations sont essentiellement des transitions C→T et se trouvent sur le même brin d'ADN. Elles sont fréquemment associées avec des réarrangements génomiques, incluant le phénomène de *chromothripsis* (Stephens et al., 2011), qui est aussi décrit dans les cellules cancéreuses, où des dommages importants causés à l'ADN entraînent l'éclatement et le réassemblage de chromosomes entiers. Ces hypermutations groupées ont été proposées être le résultat d'une désamination processive de l'ADN simple brin exposé par la résection des DSBs durant la réparation (Sakofsky et al., 2014). Contrairement aux tumeurs non-B, les régions *kataegis* observées dans les DLBCLs sont initiées par AID au sein des « hotspots » WRCY (Alexandrov et al., 2013). AID désaminerait donc l'ADN simple brin exposé par la transcription en phase G1. Le traitement des lésions de dU par les mécanismes de BER et MMR conduirait alors à des cassures double brin occasionnelles. Celles-ci subiraient la résection en phase S ou G2/M provoquant l'exposition de fragments d'ADN simple brin, substrat d'AID, qui entraînerait de longues périodes de mutations qui sont indépendantes de la transcription (Casellas et al., 2016; Kruhlak et al., 2007; Shanbhag et al., 2010).

Afin d'éviter les aberrations génomiques induites par AID qui pourraient conduire à la transformation cellulaire et au cancer, les cellules B ont développé un ensemble

de mécanismes permettant de contrôler l'activité d'AID et de protéger l'intégrité du génome.

4.2. La régulation d'AID

L'expression d'AID joue un rôle important dans l'équilibre d'une activité physiologique versus pathologique. La présence de mécanismes qui régulent son abondance, sa localisation cellulaire et son activité permet de maintenir l'intégrité du génome (Figure 11).

4.2.1. La régulation transcriptionnelle

L'expression d'AID limitée dans les cellules B activées des GCs est acquise par la combinaison de facteurs de transcription activateurs et répresseurs qui agissent au niveau de quatre grandes régions *cis*-régulatrices se trouvant de part et d'autre du locus du gène *Aicda* codant pour AID. On retrouve donc la région 1 localisée en amont du site de démarrage de la transcription (TSS), la région 2 entre l'exon 1 et 2, la région 3 située à 17 kb en aval du TSS et enfin, la région 4 à 8 kb en amont du TSS (Crouch et al., 2007; Huong le et al., 2013; Tran et al., 2010). Cette régulation transcriptionnelle est dépendante de la reconnaissance de l'antigène ainsi que des interactions cellulaires et des cytokines. En effet, ces facteurs d'activation ou de répression de l'expression d'AID sont activés par différents stimuli issus des différentes cytokines reçues telles que l'interleukine 4 (IL-4), le TGF- β ou le LPS, la liaison du récepteur CD40 avec le récepteur CD154 des lymphocytes T ou bien l'activation de récepteurs Toll-like (TLR, de l'anglais Toll-like receptors) (Dedeoglu et al., 2004; Pone et al., 2012). Les facteurs de transcription généraux et spécifiques aux cellules B tels que NF- κ B, HoxC4, Stat6 (signal transducer and activator of transcription 6), CEBPs (CCAAT/enhancer-binding proteins), Sp1, Sp3, qui se lient à la région 4, les protéines E-box et Pax5, qui se lient à la région 2, et BATF (Basic leucine zipper, qui se lie à la région 3, transcriptional factor ATF-like) activent l'expression d'AID (Betz et al., 2010; Dedeoglu et al., 2004; Gonda et al., 2003; Huong le et al., 2013; Ise et al., 2011; Park et al., 2003; Park et al., 2009; Sayegh et al., 2003; Tran et al., 2010; Yadav et al., 2006). En revanche, Id2 (Inhibitor of differentiation), Id3, BLIMP1 (B lymphocyte-induced maturation protein 1), E2f, c-

myb et Sirt1 répriment l'expression d'AID (Gan et al., 2020; Gonda et al., 2003; Huong le et al., 2013; Sayegh et al., 2003; Tran et al., 2010). De plus, les hormones sexuelles jouent également un rôle dans la régulation transcriptionnelle par l'intermédiaire de leurs récepteurs nucléaires. L'expression d'AID est induite par le récepteur à l'œstrogène tandis qu'elle est inhibée par le récepteur à la progestérone (Mai et al., 2010; Pauklin and Petersen-Mahrt, 2009; Pauklin et al., 2009).

La répression du locus Aicda dans les cellules B naïves est aussi médiée par une hyperméthylation du promoteur (Fujimura et al., 2008). Une fois les cellules B activées, le locus est déméthylé, notamment par les protéines TET (ten-eleven-translocation) recrutées par BATF, permettant l'augmentation de l'accessibilité à la chromatine et l'induction de l'expression d'AID (Lio et al., 2019).

Tandis que l'expression d'AID est fortement induite dans les cellules B activées, elle est inhibée dans les cellules B naïves, les plasmocytes ou les cellules B mémoires ainsi que les cellules non-B. L'ensemble de ces facteurs activateurs et inhibiteurs permet donc de limiter l'expression d'AID à quelques populations cellulaires.

La stabilité des ARN messagers (ARNm) d'AID est aussi régulée par 6 microARN (miR) : miR-181b, miR-155, miR-361, miR93 et miR-29b (Basso et al., 2012; Borchert et al., 2011; Dorsett et al., 2008; Recaldin et al., 2018; Teng et al., 2008) qui se fixent sur la partie 3'UTR (UnTranslated Region) de l'ARNm d'AID en empêchant ainsi sa traduction. Leur absence est donc liée à une augmentation de la quantité protéique d'AID. De plus, la perte de miR-155 aboutit à un niveau élevé de translocations c-myc/IgH, montrant que miR-155 possède un rôle protecteur qui réduit potentiellement les translocations oncogéniques générées par AID (Dorsett et al., 2008; Teng et al., 2008). Cette translocation est une caractéristique des cellules du lymphome de Burkitt qui par ailleurs n'expriment pas miR-155 (Kluiver et al., 2006). Cependant, miR-155 et miR-361 sont inhibés par Bcl-6 afin d'induire l'expression d'AID et d'autres facteurs impliqués dans la maintenance des centres germinatifs (Basso et al., 2012). Enfin, les miR-29 et miR-181b sont exprimés dans les cellules B naïves et dans les cellules non-B afin de réduire l'expression basale d'AID au silence (de Yebenes et al., 2008; Recaldin et al., 2018).

4.2.2 La régulation de la localisation cellulaire

Bien qu'AID assume sa fonction de désaminase dans le noyau, sa localisation majoritaire se trouve dans le cytoplasme et apporte donc un accès limité à son substrat situé dans le noyau (Ito et al., 2004; Pasqualucci et al., 2004; Rada et al., 2002a). Cet équilibre dynamique, dans lequel seulement 10% d'AID est dans le noyau, est contrôlé par une balance complexe entre des mécanismes d'import et d'export nucléaire, de rétention dans le cytoplasme et une stabilité protéique différente dans chacun des compartiments. Cet équilibre permet de protéger le génome des cellules B de mutations mal ciblées.

AID expose une demi-vie beaucoup plus courte dans le noyau que dans le cytoplasme. Dans le cytoplasme, les protéines chaperonnes HSP90 (Heat Shock Protein-90) et HSP40/DNAJA1 (DnaH homolog subfamily A member 1) stabilisent AID en la protégeant de la dégradation dépendante de l'ubiquitination (Methot et al., 2015; Orthwein et al., 2010; Orthwein et al., 2012). L'absence d'un de ses deux facteurs mène à une baisse de la quantité d'AID et de son activité et par conséquent à un défaut de CI et d'HMS (Montamat-Sicotte et al., 2015; Orthwein et al., 2010; Orthwein et al., 2012). AID peut être retenue dans le cytoplasme par l'intermédiaire de son motif NES en C-ter (Patenaude et al., 2009). De plus, l'association d'AID et d'un complexe contenant le facteur d'élongation de la traduction eEF1A (eukaryotic Elongation Factor 1 alpha) est aussi responsable de la rétention cytoplasmique d'AID (Hasler et al., 2011; Methot et al., 2015). L'absence ou la rupture de cette interaction mène à une accumulation d'AID dans le noyau, et donc à une augmentation de la CI, mais aussi à plus de translocations (Methot et al., 2015), ce qui suggère que la majorité d'AID est mature et probablement active mais séquestrée dans le cytoplasme, à l'écart de son substrat.

AID fait la navette entre le cytoplasme et le noyau par l'intermédiaire d'un import actif mené par l'importine- α qui reconnaît un signal de localisation nucléaire bipartite localisé en N-ter d'AID (Patenaude et al., 2009). De plus, le facteur d'épissage CTNNBL1 (Beta-catenin-like protein1) contribue aussi à l'import d'AID dans le noyau (Ganesh et al., 2011). En revanche, l'absence de CTNNBL1 n'a pas d'impact sur l'efficacité de la CI dans les CH12 (Han et al., 2010) et un impact moyen

dans les cellules B murines primaires (Chandra et al., 2013), suggérant que son rôle est redondant avec d'autres facteurs. Enfin, la surexpression de GANP (Germinal center Associated Nuclear Protein) favorise l'accumulation d'AID dans le noyau (Maeda et al., 2010), suggérant un rôle dans l'import d'AID ou reflétant un rôle de GANP dans l'aide au recrutement d'AID aux loci des Ig durant l'HMS (Kawahara et al., 2004; Singh et al., 2013). Etant donné la petite taille d'AID (24 kDa) qui lui permettrait théoriquement de diffuser passivement dans le noyau, le mécanisme de rétention et l'ensemble de ses facteurs sont essentiels pour qu'AID puisse remplir son rôle physiologique de désaminase (Macara, 2001).

Dans le noyau, la dégradation ou l'exportation d'AID permet de limiter son potentiel oncogénique. La dégradation d'AID est réalisée par le protéasome grâce à deux mécanismes, exercée d'une part, par l'ubiquitine ligase E3 CUL7 et FBXW11 (Luo et al., 2019) et de manière dépendante de la poly-ubiquitination d'AID et de l'ATP (Aoufouchi et al., 2008); et d'autre part, par REG- γ de façon indépendante de l'ubiquitination et de l'ATP (Uchimura et al., 2011). Néanmoins, le facteur de transcription YY1 (Yin Yang 1) contrôle la stabilité d'AID dans le noyau (Zaprazna and Atchison, 2012) en régulant son accumulation nucléaire. Cette régulation est probablement due à l'interaction physique YY1-AID qui bloque l'ubiquitination d'AID (Zaprazna et al., 2018). Après la mitose, lorsque la membrane nucléaire se reforme, AID peut être piégée dans le noyau (Hu et al., 2013; Ito et al., 2004; Methot et al., 2018; Patenaude et al., 2009; Wang et al., 2017), corrélant ainsi avec les activités de désamination d'AID et de la réparation de l'ADN par le BER ou le MMR qui se produisent principalement en phase G1 du cycle cellulaire (Faili et al., 2002; Sharbeen et al., 2012).

Enfin, dans le but de réguler les niveaux d'AID nucléaires, CRM1 exporte activement AID du noyau, suite à la reconnaissance d'un signal d'export situé en C-ter d'AID (Brar et al., 2004; Geisberger et al., 2009; Ito et al., 2004; McBride et al., 2004).

4.2.3. La régulation de l'activité d'AID

L'activité d'AID est régulée par des modifications post-traductionnelles et plus particulièrement, par des phosphorylations. AID comporte 7 sites de

phosphorylation connus à ce jour : sérine 3 (S3), thréonine 27 (T27), sérine 38 (S38), sérine 41 (S41), sérine 43 (S43), thréonine 140 (T140) et tyrosine 184 (Y184) (Basu et al., 2005; Gazumyan et al., 2011; McBride et al., 2006; McBride et al., 2008; Pasqualucci et al., 2006). La phosphorylation à la S38 par PKA (cAMP-dependent protein kinase A) régule l'activité d'AID (Basu et al., 2005) et a un rôle important dans la CI et l'HMS qui sont presque abolies en absence de cette modification (Cheng et al., 2009; McBride et al., 2008; Vuong et al., 2013). La faible fraction d'AID qui est phosphorylée à la S38 (5-15%) est liée à la chromatine et provoque la plupart des mutations constatées durant l'HMS (McBride et al., 2006). Cette phosphorylation S38 permet l'association d'AID avec RPA (Replication Protein A) et APE1 (Apurinic/apyrimidic endonuclease 1). La mutation de la sérine 38 d'AID en alanine mène vers une abolition presque complète des mécanismes de CI et d'HMS tandis que son activité catalytique reste inchangée. L'abolition de la CI et de l'HMS s'explique par la perte de capacité à générer des DSBs, mécanisme dans lequel APE1 joue un rôle important (Cheng et al., 2009; McBride et al., 2008; Vuong et al., 2013). La reproduction de cette mutation de la sérine 38 chez des souris *Msh2*^{-/-} et *Ung*^{-/-} montre également une abolition de la CI et un défaut sévère de l'HMS, ce qui suggère que la phosphorylation d'AID à la S38 est requise pour que le MMR intervienne lors de l'absence d'Ung. (Choi et al., 2020).

Il a été montré que PKA s'associe aux régions S pendant la CI, ce qui suggère que PKA est recrutée dans ces régions pour induire la phosphorylation d'AID à la S38 afin d'activer la cascade de la réaction de CI (Vuong et al., 2009). Cependant, une autre étude (Vuong et al., 2013) a révélé qu'un mutant catalytique inactif d'AID, lié aux régions S, n'est pas phosphorylé efficacement. De plus, il a été observé une forte baisse du niveau de phosphorylation d'AID à la S38 dans des cellules *Ung*^{-/-} et *Msh2*^{-/-} qui sont dépourvues de DSBs. Ce défaut de phosphorylation d'AID a aussi été observé dans des cellules B murines primaires déficientes en ATM (Vuong et al., 2013). L'induction artificielle de DSBs, à l'aide de rayons ionisants, dans ces cellules a permis de restaurer la phosphorylation d'AID, montrant la nécessité des DSBs pour l'activation de la phosphorylation d'AID. L'ensemble de ces résultats a permis de décrire un premier modèle dans lequel AID induit des mutations dans les régions S, entraînant des DSBs suite à l'action du BER et du MMR qui vont activer ATM,

permettant à PKA de phosphoryler une plus grande quantité d'AID. Cette phosphorylation mènera donc à l'association d'AID avec RPA et APE1. Cette boucle de rétroaction positive renforce l'activité d'AID et entraîne une forte densité de DSBs au sein des régions S (Vuong et al., 2013).

Très récemment, il a été exposé que l'augmentation de la phosphorylation d'AID à la S38 provoque une augmentation des mutations hors cibles, des DSBs et des translocations c-myc/IgH alors que sa déficience atténue d'une part, l'association d'AID à la chromatine et d'autre part, la CI mais pas les mutations à c-myc. Ceci indique que la phosphorylation favorise un ciblage différentiel de l'activité mutationnelle d'AID pour induire des lésions oncogéniques (Mu et al., 2017).

Enfin, contrairement aux autres phosphorylations d'AID, la phosphorylation de la S3 par PP2A (proteine phosphatase 2) régule négativement l'activité d'AID ainsi que son potentiel oncogénique (Gazumyan et al., 2011). La phosphorylation de la T140 est catalysée par PKC *in vitro* (McBride et al., 2008). La mutation de cette modification en alanine n'influence pas l'activité d'AID et altère spécifiquement l'HMS, suggérant que cette modification soit importante pour l'association de facteurs spécifiques à l'HMS (McBride et al., 2008). La phosphorylation de la Y184 est encore très peu décrite mais a été observée *in vivo* (Basu et al., 2005; McBride et al., 2006) et n'a pas de rôle dans la localisation ou l'activité d'AID (Patenaude and Di Noia, 2010).

5. Le ciblage d'AID

Bien que les régions V et S soient des cibles physiologiques d'AID afin de promouvoir respectivement l'HMS et la CI, AID peut générer des mutations et des DSBs dans les gènes non-Ig. Des mécanismes permettant de cibler AID aux loci Ig pour promouvoir les processus de diversification des anticorps tout en limitant l'accès d'AID au reste du génome ont été mis évidence. Ces mécanismes additionnels agissent en aval du recrutement d'AID et permettent le maintien de l'intégrité du génome (**Figure 11**).

5.1. Les motifs

L'enzyme AID cible et désamine préférentiellement les cytosines présentes dans les motifs WRCY, RGYW sur le brin complémentaire où W=A/T ; R=A/G ; Y=C/T, (Betz

et al., 1993; Rogozin and Kolchanov, 1992). Les nucléotides entourant la cytosine cible d'AID, soulignée plus haut, ont un rôle important dans l'efficacité de la désamination (Pham et al., 2003; Yu et al., 2004). Il a été mis en évidence que la séquence privilégiée était AGCT et que les protéines adaptatrices 14-3-3 contribuent au recrutement d'AID à ces séquences durant la CI (Xu et al., 2010). Les motifs WRCY, aussi appelés motifs « hotspots », ont été identifiés dans les CDRs des régions V et S (Hackney et al., 2009). Les régions S présentent une haute densité du motif AGCT mais il apparaît que ce soit le nombre de motifs au sein d'une région S qui soit le facteur important (Hwang et al., 2015; Zhang et al., 2014).

Cependant, ces motifs ne sont pas absolument requis pour la CI. En effet, le remplacement de la région $S\gamma 1$ par une région riche en G mais dépourvue de motifs montre une réaction de CI vers IgG1 avec néanmoins, une efficacité plus faible que dans des cellules B wild-type (Shinkura et al., 2003). De plus, la suppression de la quasi-totalité de la région $S\mu$, et donc des motifs WRCY, ne mène pas à l'abolition complète de la CI (Khamlichi et al., 2004). Également, lors de l'HMS, ces motifs ne sont pas essentiels pour le ciblage, tandis que le contexte génomique de la région V au sein du locus IgH a un rôle important (Yeap et al., 2015).

5.2. Les éléments en *cis*

Des éléments *cis*-régulateurs dans les gènes Ig peuvent favoriser le recrutement d'AID. Les enhancers Ig contiennent des éléments conservés appelés DIVACs (diversification activators) qui sont essentiels pour le ciblage d'AID lors de l'HMS et la conversion génique (Kim and Tian, 2010; Kothapalli et al., 2008). De plus, il a été montré que les DIVACs sont suffisants pour induire l'HMS (Blagodatski et al., 2009). Très récemment, il a été exposé que certains TADs sont sensibles à l'HMS. Cette sensibilité est notamment due à la présence d'éléments DIVACs « -like enhancers ». En effet, l'insertion d'un élément DIVAC fort dans un TAD résistant à l'HMS a converti le TAD en un TAD sensible à l'HMS, révélant ainsi le potentiel d'un élément DIVAC à provoquer un mauvais ciblage de l'HMS (Senigl et al., 2019).

Les DIVACs contiennent de nombreux sites de liaison de facteurs de transcription qui sont nécessaires pour promouvoir l'activité d'AID tels que les membres de la famille d'Ikaros, les E-box, NF- κ B, MEF2 et PU.1-IRF4 (Buerstedde et al., 2014;

Dinesh et al., 2020; Kohler et al., 2012; Senigl et al., 2019). Cependant, le mécanisme d'activité des DIVACs n'est pas clair. Néanmoins, il a été montré que les DIVACS sont transcrits de façon bidirectionnelle, produisant de l'enhancer ARN (eARN) à partir des deux brins d'ADN (Senigl et al., 2019). D'autres études ont présenté que la transcription bidirectionnelle et la dégradation de l'eARN par l'exosome peuvent contribuer au ciblage de l'activité d'AID (Pefanis et al., 2014; Pham et al., 2019), effet potentiel des DIVACs.

La région régulatrice 3'RR du locus IgH est également impliquée dans le recrutement en *cis* d'AID aux régions V (Rouaud et al., 2013) et aux régions S (Saintamand et al., 2017) mais également en *trans* (Le Noir et al., 2017). En effet, le super-enhancer 3'RR permet le recrutement d'AID de manière bi-allélique grâce à une proximité inter-allélique (Le Noir et al., 2017).

5.3. La transcription

AID interagit avec l'ARN polymérase II (ARNP II) et ses protéines associées telles que Spt5 et les complexes PAF et FACT qui sont essentiels pour le ciblage d'AID aux loci Ig. L'ensemble des mécanismes liés à la transcription tels que la pause de l'ARNP II, les R-loops, ainsi que les marques épigénétiques est largement décrit dans la partie IV-1. De plus, l'ARNP II en pause peut recruter le complexe de l'exosome pour dégrader le transcrit d'ARN naissant et faciliter la désamination des deux brins d'ADN (Basu et al., 2011; Wang et al., 2014b).

5.4. Le complexe de l'exosome à ARN

L'énigme de l'accès d'AID au brin transcrit et non transcrit a été résolue par l'implication du complexe de l'exosome à ARN (Basu et al., 2011). Ce complexe est recruté aux régions S d'une manière dépendante d'AID et rend le brin transcrit accessible aux désaminations induites par AID en dégradant l'ARN en cours de transcription. La formation de l'ADN simple brin (ADNss), substrat d'AID, et sa stabilisation par la possible association de RPA à l'ADNss qui facilite l'accumulation d'AID phosphorylée à la sérine 38 permet à l'exosome de stimuler l'activité d'AID (Basu et al., 2011). La dégradation de l'ARN naissant peut être favorisée par la dégradation de l'ARNP II aux régions S par l'ubiquitine ligase Nedd4 qui permettrait

donc l'accès à la bulle de transcription à l'exosome (Anindya et al., 2007; Sun et al., 2013a). De plus, l'exosome dégrade aussi l'ARN non codant anti-sens issu de la transcription divergente (Pefanis et al., 2014). Cette dégradation permet de protéger l'intégrité du génome d'une augmentation de structures hybrides ARN : ADN, d'une terminaison prémature de la transcription et d'instabilité génomique. Enfin, l'exosome à ARN contrôle aussi l'expression des longs ARN non codants provenant d'éléments enhancer divergents qui sont engagés dans les interactions de longue distance et dans la régulation directe de l'activité super-enhancer du 3'RR (Pefanis et al., 2015). L'exosome permet donc de réguler les interactions chromosomiques de longues distances, connues pour favoriser la CI (Pefanis et al., 2015). Le complexe de l'exosome est donc impliqué dans le ciblage d'AID sur l'ensemble du génome (gènes Ig et non-Ig).

5.3.5. L'épissage des ARN

L'absence du site donneur de l'épissage d'I γ 1 dans des souris, montrant que la CI vers IgG1 était abolie en dépit d'une transcription active des régions S, a mis en évidence la machinerie de l'épissage (Hein et al., 1998; Lorenz et al., 1995). Il a été suggéré suite à ces expériences que les transcrits épissés ou la machinerie de l'épissage soient requis pour la CI. Ensuite, des expériences de co-IP ont révélé qu'AID est physiquement et fonctionnellement liée à des protéines en lien avec la machinerie de l'épissage comme CTNNBL1 (Conticello et al., 2008), des membres de la famille de hnRNPs (heterogeneous nuclear ribonucleoproteins) (Hu et al., 2015; Mondal et al., 2016), PTBP2 (polypyrimidine tract-binding protein2) (Matthews et al., 2014; Nowak et al., 2011) et SRSF1-3 (Singh et al., 2020).

HnRNPK semble être nécessaire à l'activité mutagène d'AID aux régions S et aux régions V, tandis que hnRNPL est dispensable pour l'activité d'AID (Mondal et al., 2016) mais est toutefois requise pour la CI dans les cellules CH12 (Hu et al., 2015). Le knockdown de PTBP2 affecte la CI et réduit AID aux régions S alors qu'il augmente la quantité d'AID associée à la région V adjacente (Matthews et al., 2014). Enfin, l'isoforme du facteur d'épissage SRSF1, SRSF1-3 est nécessaire pour l'HMS dans les cellules B de poulet DT40 afin de faciliter l'accès d'AID à l'ADN simple brin dans les régions V en régulant l'épissage des transcrits (Kanehiro et al., 2012; Singh

et al., 2020). De plus, en interagissant avec le domaine C-ter d'AID, SRSF1-3 est impliqué dans l'accumulation d'AID dans le noyau (Kawaguchi et al., 2017). La différence de recrutement de ces protéines au sein des régions S et V démontre que les facteurs du spliceosome jouent un rôle essentiel dans la localisation d'AID et donc dans l'orientation des mécanismes de diversification des anticorps. Suite à l'épissage, l'ARN non codant qui est sous forme de lasso, est linéarisé par l'enzyme DBR1 (debranching enzyme 1). Le transcrit forme alors une structure secondaire de type G-quadruplex qui s'associe avec AID et la guide aux régions S par complémentarité de séquence (Zheng et al., 2015). Le rôle des G-quadruplex lors de la CI ainsi que les dernières études sur ces structures sont décrits plus en détail dans la partie IV.1.

Malgré l'apparition de nouveaux mécanismes et de nouvelles connaissances, le mécanisme par lequel AID mute préférentiellement les gènes Ig n'est pas encore compris. Le ciblage génomique d'AID semble être multi-couches, avec de la redondance mais suffisamment robuste pour garantir que la plupart du génome soit épargné de l'activité d'AID. La contribution de chacun de ces mécanismes dans le ciblage d'AID pour la CI, l'HMS et la mutagénèse hors-cible reste encore à être mesurée. Il serait également intéressant de savoir si certains de ces mécanismes sont liés ou s'ils sont indépendants les uns des autres.

Figure 11: Les mécanismes régulateurs d'AID. Légende sur la page suivante.

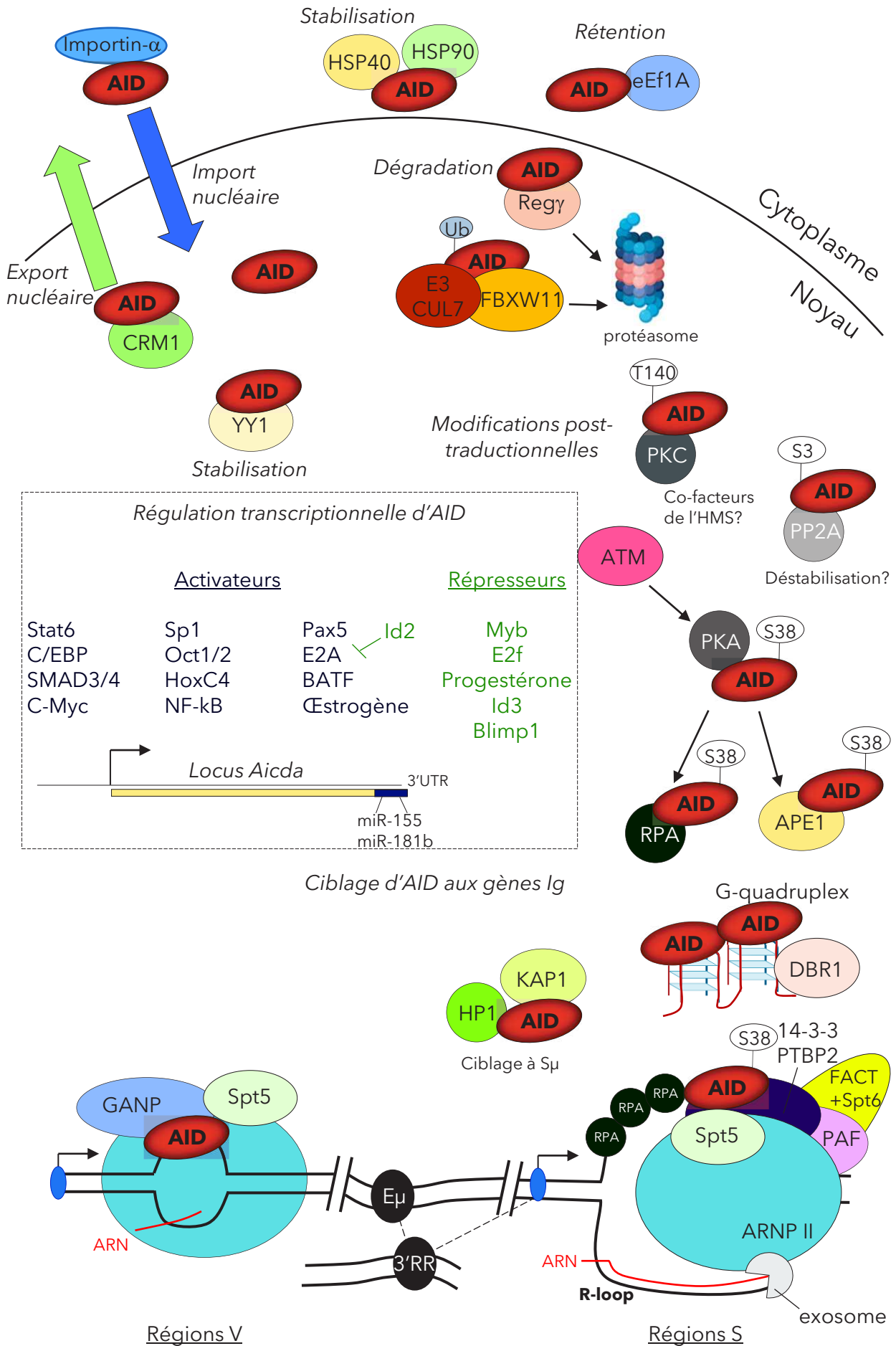


Figure 11: Les mécanismes régulateurs d'AID. AID, du fait de son action mutagène, est finement contrôlée dans les lymphocytes B. L'expression d'AID est tout d'abord régulée au niveau transcriptionnel par des facteurs de transcription activateurs (Stat6, C/EBP, SMAD3/4, c-Myc, Sp1, Oct1/2, HoxC4, NF-kB, 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 l'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 au contrôle de la localisation sous-cellulaire d'AID par son export nucléaire médié par CRM1 et par son import nucléaire avec l'importin- α , de sa stabilisation cytoplasmique via son interaction avec Hsp40 et Hsp90 et de sa rétention dans le cytoplasme en interagissant avec eEF1A. Il contribue aussi à sa dégradation dans le noyau par ubiquitination (Ub) induite par l'E3 CUL7 et FBXW11 et par interaction avec REG γ , et à sa stabilisation nucléaire via YY1. L'activité d'AID est aussi contrôlée par des modifications post-traductionnelles, comme la phosphorylation aux S3, T140 et S38, et des modifications épigénétiques, non détaillées ici. L'activation d'ATM induit la phosphorylation à la S38 par PKA, ce qui permet l'association d'AID avec RPA, facteur qui permet de stabiliser l'ADN simple brin durant la transcription et APE1, membre de la voie BER.

Le ciblage d'AID aux régions S et IgV requiert une transcription active afin d'exposer l'ADN simple brin à AID par la formation d'une bulle de transcription et l'extrusion du brin non transcrit. Ce microenvironnement chromatinien inclut des interactions longues distances entre promoteurs (ovale bleu) et enhancers (E μ 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 la 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 l'exposition 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, 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, apparaissant dans l'ARN transcrit et dans l'ADN du brin non-transcrit, qui semblent aussi requises pour le bon recrutement d'AID aux loci Ig.

III- La génération de mutations induites par AID durant l'HMS.

Durant l'hypermutation somatique, AID désamine les cytosines en uraciles, introduisant des mésappariements dU : dG dans les régions variables des Ig (IgH et IgL). L'enzyme AID cible et désamine préférentiellement les cytosines présentes dans les motifs $WR\text{C}Y$ ($W=A/T$; $R=A/G$; $Y=C/T$), aussi appelés motifs « hotspots » (Betz et al., 1993; Rogozin and Kolchanov, 1992) qui ont été identifiés dans les régions déterminant la complémentarité du site de liaison de l'antigène (CDRs) (Hackney et al., 2009). De plus, il a été observé que les mutations des gènes IgV sont plus courantes à environ 100-200 pb en aval du TSS et peuvent s'étendre jusqu'à 1.5-2 kb en aval du promoteur (Lebecque and Gearhart, 1990; Rada and Milstein, 2001), en épargnant les régions constantes.

Durant l'hypermutation somatique, les uraciles introduits par la désamination d'AID peuvent être reconnus par différentes voies de réparation pouvant mener soit à un profil non-mutagène, soit à un profil mutagène (Rada et al., 2004; Rada et al., 2002b) (**Figure 10**). Tout d'abord, l'enzyme UNG, appartenant à la voie de réparation BER, peut exciser l'uracile et donc créer un site abasique, aussi appelé site AP (Cristina Rada, 1998). Celui-ci est reconnu par les endonucléases apuriniques/apyrimidiques (APE) qui introduisent des cassures simples brins à l'ADN (SSBs, de l'anglais single stranded DNA breaks). Ces lésions peuvent être réparées de manière fidèle médiées par la polymérase β .

Par ailleurs, les mésappariements dU : dG sont traités par les voies détournées du BER et du MMR à des fins mutagènes. Le BER est principalement impliqué dans la phase 1 de l'HMS et entraîne des mutations aux bases C et G (Rada et al., 2002b) tandis que dans la seconde phase, ce sont des mutations aux bases A et T qui sont introduites par la voie du MMR (Chahwan et al., 2012). Environ 50% des mutations apparaissent au niveau des paires A/T.

Dans la phase 1A, la réplication des uraciles via une polymérase d'une haute-fidélité peut introduire des transitions $C \rightarrow T$ ou $G \rightarrow A$, ou bien 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). La polymérase Rev1, une désoxycytidine transférase, introduit une C en face du site abasique et est responsable de la plupart des mutations de transversions aux bases C et G (Jansen et al., 2006).

Dans la phase 2, les mésappariements dU : dG sont traités par la voie de réparation des mésappariements non canonique (ncMMR) qui génère la majorité des mutations A/T. En effet, l'hétérodimère MSH2/MSH6 reconnaît le mésappariement puis un trou au niveau d'un des deux bras de l'ADN est produit afin de créer un point d'entrée pour l'exonucléase, Exo1 (Bardwell et al., 2004; Frieder et al., 2009b). Depuis plusieurs années, il a été montré que contrairement à la voie canonique du MMR, PMS2 ET MLH1 sont dispensables durant 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). En premier lieu, il a été proposé qu'UNG et APE2 aident la voie du MMR pour la mutagenèse A-T (Frieder et al., 2009b; Stavnezer and Schrader, 2014) en créant le trou requis pour la résection de l'ADN par Exo1. Une étude (Girelli Zubani et al., 2017) a récemment montré que finalement l'espace permettant l'entrée de l'exonucléase Exo1 serait initié par le complexe PMS2/MLH1 mais pourrait être compensé par l'action d'UNG. Les facteurs permettant la génération d'un espace pour l'entrée d'Exo1 pendant l'HMS ne sont pas encore très clairs. Exo1 va donc exciser et éliminer l'ADN simple brin comprenant la lésion ce qui va former une brèche (Bardwell et al., 2004; Frieder et al., 2009b). Cette brèche va être ensuite resynthétisée par l'ADN polymérase translésionnelle η qui insère de préférence des nucléotides T spécifiquement aux motifs WA/TW (Delbos et al., 2007; Rogozin et al., 2001; Zeng et al., 2001). Cette polymérase est recrutée grâce à la mono-ubiquitination spécifique de PCNA (Proliferating cell nuclear antigen) à la lysine 164 (Langerak et al., 2007; Peña-Díaz et al., 2012; Roa et al., 2008; Zlatanou et al., 2011), facteur appelé PCNA-Ub.

Plus récemment, l'existence d'une 4^{ème} et d'une 5^{ème} voie a été proposée pour le modèle de diversification des mutations lors de l'HMS. La 4^{ème} voie, appelée la voie hybride UNG+MSH2, requiert au moins deux uraciles et nécessite les deux voies

permettant d'une part, la génération d'un espace au niveau d'un des deux brins de l'ADN par MSH2/MSH6 et d'autre part, la génération d'un site AP par UNG (Frieder et al., 2009b; Krijger et al., 2009). Dans cette voie, le mésappariement U:G est reconnu par le complexe MSH2/MSH6. Une partie des nucléotides entourant le mésappariement est alors éliminé par Exo1, laissant un vide dans l'ADN simple brin. Cependant, si sur le brin opposé un site est créé par UNG, les polymérases translésionnelles peuvent insérer une base à l'opposé du site AP, ce qui entraîne des transversions au niveau des bases G et C (Krijger et al., 2013). Enfin, la 5^{ème} voie, appelée la voie longue du BER (long patch BER pathway) a tout d'abord été décrite dans un modèle murin déficient en MSH2 (Rada et al., 2004). Elle est dépendante d'UNG, de PCNA-Ub et de la polymérase η et représente 8 à 20% des mutations dans les paires de base A/T (Delbos et al., 2007; Rada et al., 2004). (Ijspeert et al., 2019; Pilzecker and Jacobs, 2019).

Il est important de noter que parmi les uraciles glycosylases, UNG a une fonction privilégiée durant l'HMS car SMUG1 ne peut pas entièrement compenser la perte d'UNG (Di Noia et al., 2006; Dingler et al., 2014), excepté si SMUG1 est surexprimé (Di Noia et al., 2007b).

IV- Le rôle d'AID dans la commutation isotypique

La commutation isotypique est une réaction de recombinaison de plusieurs étapes qui est dépendante de la transcription des régions de « switch » et de l'expression d'AID qui mène à la formation de DSBs qui sont ensuite réparées par le NHEJ (**Figure 12**). Etant le principal mécanisme de diversification des anticorps que j'ai étudié durant ma thèse, toutes les étapes nécessaires à la commutation isotypique seront décrites.

1. La transcription des régions de switch

La commutation isotypique est initiée par la transcription germinale des promoteurs des régions S donneuse et acceptrice en réponse des cytokines et des signaux de co-stimulation reçus. La transcription à partir de ces promoteurs produit des transcrits germinaux (GLTs) qui sont épissés pour éliminer la région intronique de « switch » et joindre l'exon-I à l'exon constant C_H , générant ainsi des transcrits stériles (ne codant aucune protéine détectable) (Chaudhuri and Alt, 2004) qui sont polyadénylés. L'analyse de mutations qui altèrent ou suppriment le promoteur de l'exon-I a montré que l'efficacité de la CI est perturbée (Bottaro et al., 1994; Harriman et al., 1996; Jung et al., 1993; Zhang et al., 1993), indiquant que les transcrits stériles sont essentiels pour une réaction de CI efficace.

La transcription permet de rendre accessible les régions S et d'exposer le brin non transcrit sous forme d'ADN simple brin qui va être ciblé par AID via la formation de bulle de transcription ou de boucle. Elle joue donc un rôle primordial dans le ciblage d'AID.

Pour ce modèle d'accessibilité, la transcription des GLTs a trois grands rôles dans la commutation isotypique :

Le premier est de fournir le substrat d'AID pour la désamination. En effet, des structures telles que des R-loops ont été mises en évidence durant la transcription de la région S (Tian and Alt, 2000; Yu et al., 2003; Yu et al., 2004; Yu et al., 2005), dans lesquelles l'ARN naissant s'hybride à l'ADN matrice pour former un hybride ARN:ADN stable qui expose l'ADN non transcrit sous forme d'ADN simple brin, substrat de l'enzyme AID. La formation de ces R-loops est favorisée par les

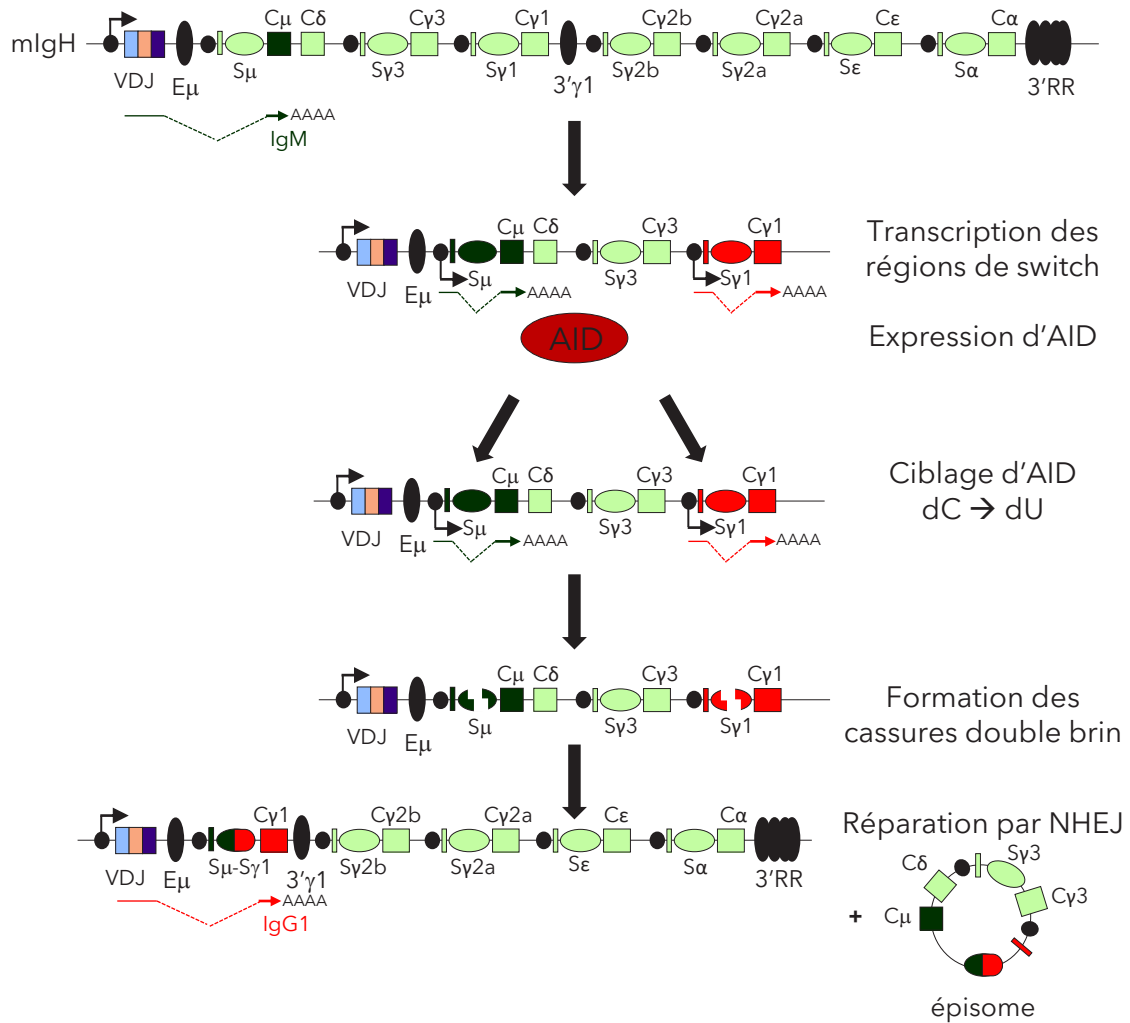


Figure 12: La réaction multi-étapes de la commutation isotypique. La commutation isotypique est un événement constitué de plusieurs étapes qui débute par la transcription germinale des régions S donneuse (ligne verte en pointillée) et acceptrice (ligne rouge en pointillé). Sous l'expression d'AID et son ciblage aux régions S, la désamination des cytosines en uraciles dans l'ADN introduit des mésappariements qui sont traités en cassures double brin (DSBs) par les voies BER et MMR. Enfin, les DSBs sont réparées via la voie de réparation NHEJ, ce qui aboutit à l'expression d'un nouvel isotype (de IgM à IgG1, ici)

séquences répétitives des régions S qui sont d'une part, enrichies par des motifs « hotspots » GAGCT (Obata et al., 1981; Yu et al., 2004; Yu et al., 2005; Zarrin et al., 2004) pour le ciblage d'AID, avec le court palindrome AGCT représentant la forme canonique du motif RGYW, et d'autre part particulièrement riches en G sur le brin non transcrit. Il a été proposé que les R-loops jouent un rôle important dans la commutation isotypique en fournissant un substrat stable pour AID, menant à des taux de mutations plus élevés, qui à leur tour, augmentent l'efficacité de la CI (Yu et al., 2003). Il a été mis en évidence d'autres structures secondaires de type G-quadruplex ou G4 qui se forment avec les répétitions en G sur le brin non transcrit (Duquette, 2004). Une étude récente a proposé un modèle dans lequel les ARN non codant produits par la transcription des régions de switch, linéarisés par l'enzyme DBR1 (Debranching enzyme) suite à l'épissage, peuvent former des G-quadruplex ou des structures d'ARN de type G4 qui s'associent avec AID, la guidant ainsi préférentiellement à l'ADN de la région S complémentaire de manière spécifique à la séquence (Zheng et al., 2015). Cette interaction est cruciale pour le recrutement d'AID au niveau des régions S ainsi que pour la CI (Zheng et al., 2015). Deux études très récentes montrent également d'une part qu'AID a une préférence intrinsèque pour les substrats structurés et que la reconnaissance de ces structures G4, qui facilitent l'oligomérisation d'AID aux régions S, est essentielle lors de la CI (Pucella and Chaudhuri, 2017; Qiao et al., 2017). Il a été montré *in vitro* qu'AID a une liaison de substrat bifurqué qui peut aider à la liaison d'une seule molécule d'AID à 2 extrémités débordantes simple brin à proximité. Ainsi, l'association d'AID aux structures G4 favorise l'oligomérisation d'AID qui peut être responsable de mutations plus importantes au niveau de la région S menant aux DSBs (Qiao et al., 2017). D'autre part, l'ARN hélicase DDX1 se lie aux structures G4 présentes dans l'ARN non codant produit par la transcription des régions S et favorise la formation de R-loop sur l'ADN de la région S, ciblant ainsi AID sur le locus IgH (Ribeiro De Almeida et al., 2018). L'ensemble de ces travaux met en évidence un important mécanisme de ciblage d'AID aux régions S médié par l'ARN (Qiao et al., 2017; Ribeiro De Almeida et al., 2018; Zheng et al., 2015). Il est intéressant de noter qu'AID semble reconnaître à la fois l'ADN et l'ARN avec des affinités similaires, ce qui pourrait expliquer comment la liaison d'AID aux structures d'ARN G4 a un impact sur la

commutation isotypique (Pucella and Chaudhuri, 2017; Zheng et al., 2015). Bien évidemment, il reste encore beaucoup de questions en suspens concernant le rôle et l'importance exacts de toutes ces structures. Un nouveau modèle de régulation de la CI a été décrit dans lequel, les protéines NME1 et NME2 agissent comme un couple coordonné inhibiteur-stimulateur pour réguler la CI grâce à leur liaison aux régions S qui est favorisée par la formation de structure en G4 par le facteur stimulateur NME2 (Zheng et al., 2019).

Deuxièmement, la transcription des GLTs peut aussi favoriser la CI, en induisant le remodelage de la chromatine via des modifications post-traductionnelles des histones, afin de favoriser l'accessibilité et le recrutement d'AID aux régions S. En effet, les régions S donneuse et acceptrices portent de nombreuses marques d'activation et de méthylation sur les histones H3 et H4 associées avec le statut d'ouverture de la chromatine, telles que H3K9ac/K14ac, H3K27ac, H4K8ac et H3K4me3 (Chowdhury et al., 2008; Kuang et al., 2009; Vaidyanathan et al., 2014; Wang et al., 2009). La région donneuse S_μ présente un état transcriptionnel actif même dans les cellules B naïves, ce qui suggère qu'elle est toujours prête à participer à la CI. Les modifications d'histones comme H3K4me3, H3K9ac sont présentes à S_μ dans les cellules B naïves et augmentent lors de la stimulation. Tandis que les régions S acceptrices sont inaccessibles à la machinerie de la CI au statut basal. L'hyperacétylation de H3 des K9 et K14 corrèle avec la transcription germinale aux régions de « switch » (Wang et al., 2006a) tandis que l'acétylation de H4 a été montrée être indépendante de l'expression d'AID. L'association d'AID aux régions S peut être facilitée par certains cofacteurs d'AID par l'intermédiaire des marques épigénétiques. Notamment, HP1 (heterochromatin protein 1) et KAP1 (KRAB domain-associated protein 1) forment un complexe avec AID qui permet le recrutement et la rétention d'AID aux régions S_μ par la reconnaissance et l'interaction de la modification H3K9me3 (Jeevan-Raj et al., 2011). De plus, PTIP, un composant des complexes de type leucémique à lignées mixtes, qui sont d'importants régulateurs de la méthylation de H3K4, participe dans la réaction de CI en régulant l'accessibilité de la chromatine couplée à la transcription. Les cellules B déficientes en PTIP présentent une altération de la triméthylation de H3K4 ce qui provoque un défaut sévère de CI en raison d'une diminution de la transcription

germinale des exons C_H en aval, et une réparation de l'ADN compromise (Daniel et al., 2010; Schwab et al., 2011). Il a été aussi montré que PTIP peut être impliqué dans les interactions de longue distance entre le 3'RR et ses promoteurs I cibles (Schwab et al., 2011). Également, les protéines 14-3-3, qui permettent la stabilisation d'AID aux régions S, sont recrutées aux régions S par l'intermédiaire des motifs AGCT (Xu et al., 2010) et de la phosphorylation combinatoire de H3K9ac et H3S10 (H3K9acS10ph) déposée par le complexe GCN5/PCAF (Li et al., 2013). De nombreux autres facteurs, participant au recrutement d'AID aux régions S ou à la réparation de l'ADN, sont régulés par l'intermédiaire des marques épigénétiques (Begum et al., 2012; Hsiao and Mizzen, 2013; Pei et al., 2013; Stanlie et al., 2010).

Enfin, l'importance de la transcription des GLTs dans le ciblage d'AID a été montrée par l'étude de la distribution de l'ARN polymérase II (ARNP II) sur le locus IgH. Une expérience de CoIP a tout d'abord montré qu'AID co-immunoprécipite avec l'ARN polymérase II (Nambu et al., 2003), indiquant une association physique avec la machinerie de la transcription. Ensuite, à l'aide de CHIP, il a été montré que l'ARNP II phosphorylée sur la sérine 5 du domaine Cter (ARNP II-p-ser5) s'accumule tout au long des régions S_{μ} et $S_{\gamma 3}$ impliquant que l'ARNP II est en pause à de multiples sites dans l'ADN des régions S (Rajagopal et al., 2009; Wang et al., 2009). Cette accumulation corrèle exactement avec la distribution des mutations induites par AID, indiquant l'importance de l'ARNP II dans le ciblage et l'activité d'AID (Canugovi et al., 2009; Wang et al., 2009). La présence d'uraciles dans l'ADN générés par AID et de structures secondaires comme les R-loops et les G-quadruplex peut être impliquée dans le statut de pause de l'ARNP II à proximité du promoteur (entre 20 et 50 pb en aval du TSS) (Sun et al., 2013b). L'identification du facteur d'élongation et de pause de la transcription Spt5 lors d'un criblage à l'aide de shARN (Pavri et al., 2010) donne une preuve considérable du rôle de la pause transcriptionnelle dans le ciblage d'AID. En effet, Spt5 est une sous-unité du complexe DSIF (DRB Sensitivity Inducing Factor) qui a un rôle dans la stabilisation de l'ARNP II en pause. De plus, des expériences de CHIP-seq révèlent que Spt5, AID et l'ARNP II en pause sont colocalisés sur l'ensemble du génome, et particulièrement aux régions V et S (Maul et al., 2014; Pavri et al., 2010). Les cellules B dépourvues de Spt5 présentent un impact sévère de la CI, conséquence d'une diminution du recrutement d'AID aux

régions S des Ig (Pavri et al., 2010). Le recrutement d'AID aux gènes non-Ig est également fortement réduit en absence de Spt5 (Pavri et al., 2010). Spt5 qui s'associe à l'ARNP II en pause, permet donc le recrutement d'AID aux gènes transcrits. L'association d'AID avec Spt5 ainsi que l'ARNP II en pause a été suggérée être requise afin qu'AID puisse avoir le temps de désaminer les cytosines dans l'ADN (King et al., 2015; Larijani and Martin, 2012; Pham et al., 2016).

Néanmoins, les mécanismes qui soulignent le ciblage de la désamination productive versus non productive d'AID ne sont pas encore très clairs. Très récemment, l'étude de 3 mutants d'AID où 3 résidus arginines (R171, R174 et R178) sont mutés a révélé qu'AID s'associe avec le complexe d'élongation de la transcription (Methot et al., 2018). Ces mutants, catalytiquement actifs, s'associent avec Spt5 ainsi que l'ARNP II en pause et occupent le promoteur des gènes cibles d'AID. Cependant, ils ont perdu leur capacité d'accès au corps des gènes cibles du fait de la perte de leur association avec le facteur d'élongation Spt6. Ces 3 résidus arginines groupés définissent donc un domaine fonctionnel d'AID qui lui permet de gagner l'accès au corps des gènes cibles en circulant avec l'élongation de la transcription (Methot et al., 2018).

De plus, avant l'analyse de ces trois mutants d'AID, plusieurs études ont montré qu'AID s'associe avec des complexes qui permettent la modulation de la transcription tels que les facteurs PAF (RNAP II-associated factor) (Willmann et al., 2012), Spt6 (Okazaki et al., 2011), le complexe FACT (facilitates chromatin transcription) composé de Spt16 et SSRP1 (Stanlie et al., 2010) et le complexe de l'exosome (Basu et al., 2011). L'ensemble de ces facteurs sont requis pour une CI efficace. Le complexe FACT dépose le variant d'histone H3.3 aux régions V et S (Aida et al., 2013), ce qui promeut potentiellement la diversification des gènes IgV en augmentant la formation d'ADN simple brin accessible à AID (Romanello et al., 2016). FACT induit aussi l'accumulation de la marque d'histone H3K4me3 aux régions S (Stanlie et al., 2010), citée précédemment comme requise pour la CI. Il facilite aussi la transcription avec Spt6 en remodelant la chromatine (Saunders et al., 2006). Spt6 permet aussi le dépôt de la marque H3K4me3 aux régions S par son association avec la méthyltransférase Set1 (Begum et al., 2012). Le recrutement de FACT et Spt6 à la machinerie de la transcription requiert le complexe PAF qui agit

comme une plateforme de recrutement pour coordonner l'élongation de la transcription. Le complexe PAF est aussi impliqué dans le recrutement d'AID dans les loci Ig (Willmann et al., 2012). Une compréhension plus précise de la transition entre la pause de la polymérase et l'élongation est nécessaire pour mesurer son impact sur l'activité d'AID (Feng et al., 2020).

2. La séquence et la spécificité structurale du locus IgH

De nombreuses études mutationnelles ont été réalisées afin de souligner l'importance de la transcription des GLTs et le rôle des différentes régions du locus IgH dans les mécanismes de diversification des Ac. L'intégrité des promoteurs de l'exon-I est requis pour une commutation isotypique efficace. La suppression de la majorité de la région donneuse $S\mu$ entraîne un défaut sévère de CI vers tous les isotypes en aval (Khamlichi et al., 2004; Luby et al., 2001) tandis que la suppression de la région acceptrice $S\gamma 1$ abolit presque la CI sans affecter les C_H en aval (Shinkura et al., 2003), ce qui indique que la spécificité de l'isotype dépend de la région elle-même.

Le locus IgH contient trois éléments enhancers : l'enhancer intronique $E\mu$, le 3'RR et le $\gamma 1E$. $E\mu$ est localisé entre l'exon J_{H4} et le 5' de $S\mu$. Le super-enhancer $E\mu$ promeut la recombinaison V(D)J (Marquet et al., 2014; Perlot et al., 2005), mais n'est pas crucial pour la CI et l'HMS dans les cellules B matures (Li et al., 2010; Marquet et al., 2014); il est donc principalement considéré comme un super-enhancer dans les cellules pro-B/pre-B mais n'est qu'un simple enhancer dans les cellules B matures. Le super-enhancer 3'RR est localisé en 3' de l'exon $C\alpha$ et s'étend sur environ 30 Kb. Il est composé de plusieurs sites hypersensibles (hs) à la DNase I : hs3a, hs1,2, hs3b, hs4, hs5, hs6 et hs7. Dans des expériences d'inactivation des 4 enhancers de la région 3'RR, il a été montré que hs3a et hs1,2 sont dispensables pour la CI tandis que hs3b et hs4 sont indispensables pour la transcription des GLTs et la CI vers tous les isotypes, excepté pour IgG1 (Manis et al., 1998b; Pinaud et al., 2001). L'absence des sites hs5,6 et 7 n'a pas d'impact sur l'efficacité de CI, suggérant ainsi que ce sont des isolateurs (Volpi et al., 2012). L'absence de la région régulatrice 3'RR complète affecte sévèrement la transcription des régions S, excepté $S\gamma 1$, et la commutation

isotypique (Vincent-Fabert et al., 2010). Ce super-enhancer contrôle la transcription germinale de tous les exons C_H (Vincent-Fabert et al., 2010), excepté $C\gamma_1$ (Manis et al., 1998b; Pinaud et al., 2001) et $C\alpha$ dans les cellules B1 (Issaoui et al., 2018), en établissant des interactions de longue distance avec les promoteurs I. Les souris déficientes de la région 3'RR démontrent que seules la transcription et la CI vers $C\gamma_1$ sont partiellement indépendantes de ce contrôle, suggérant que des éléments autres que la région 3'RR sont soutenus par la transcription de γ_1 (Pinaud et al., 2001; Vincent-Fabert et al., 2010). Enfin, il existe un troisième enhancer γ_1E , situé en aval du gène $C\gamma_1$ qui joue un rôle dans la CI vers plusieurs isotypes (Amoretti-Villa et al., 2019). Ayant participé à cette étude, j'évoquerai ce dernier enhancer plus en détails dans la partie Résultats.

Durant la CI, il a été proposé un modèle dans lequel la recombinaison entre les régions de switch donneuse et acceptrice est favorisée par des changements de conformation tridimensionnelle qui impliquent la génération de boucles d'ADN entre le promoteur des régions de switch acceptrices et les enhancers $E\mu$ et 3'RR (Kenter et al., 2012; Thomas-Claudepierre et al., 2016; Wuerffel et al., 2007). Il a été montré que dans les cellules B naïves, $E\mu$ et 3'RR interagissent grâce à la formation de boucles d'ADN (Pinaud et al., 2001) et qu'après l'activation des cellules B, la région de switch acceptrice est également recrutée dans cette boucle (Kenter et al., 2012). Il a été démontré que ces boucles dépendent de $hs3b$ et $hs4$ car l'absence de ces enhancers abolit la formation de boucles (Wuerffel et al., 2007). La formation de celles-ci dépend également des complexes Cohesin (Thomas-Claudepierre et al., 2013) et Mediator (Thomas-Claudepierre et al., 2016) ainsi que des facteurs INO80 (Kracker et al., 2015) et 53BP1 (Feldman et al., 2017). L'absence de chacun de ces facteurs entraîne un défaut de CI.

Ensemble, ces données montrent la nécessité de la transcription, de la séquence et de la structure du locus IgH pour une CI efficace.

3. La formation de DSBs par le BER et le MMR

La commutation isotypique requiert les voies de réparation du BER et du MMR pour générer des cassures doubles brins dans les régions S recombinantes (**Figure 14**).

Le BER est le premier mécanisme qui est responsable de l'élimination et la réparation des petites lésions de l'ADN, qui, par ailleurs, provoquerait des mutations ou des cassures double brin de l'ADN durant la réplication. Ce mécanisme est initié par la reconnaissance et l'élimination de bases endommagées par les ADN glycosylases aboutissant dans la formation de sites AP. Ces sites AP sont hautement mutagènes et nécessitent un traitement ultérieur par les endonucléases AP ou par l'activité AP lyase des glycosylases bifonctionnelles, qui entaillent le squelette phosphodiester du site AP. Les cassures d'ADN simple brin qui en résultent peuvent être transformées en DSB ou être réparées par la voie BER de synthèse longue avec déplacement (long-patch BER) ou de synthèse courte sans déplacement (short-patch BER). La voie long-patch du BER a principalement lieu dans les cellules en prolifération et est caractérisée par la génération et le comblement d'un vide de 2 à 16 nucléotides tandis que la voie short-patch du BER se produit dans des cellules en prolifération ou non dans laquelle un vide d'un seul nucléotide est généré et comblé (Stratigopoulou et al., 2020).

Alors que les uraciles sont généralement réparés efficacement par un BER sans erreur, ce processus est étonnamment sujet aux erreurs au niveau des loci Ig des cellules B. Cette réparation mutagène peut être couplée à une inactivation du short-patch BER dans les cellules B des GCs se divisant rapidement (Bahjat et al., 2020). L'absence de réparation ou la dégradation de ce processus de haute-fidélité dans les loci non-Ig a été liée à l'accumulation de mutations présentant l'empreinte d'AID (Xue et al., 2006), incluant les oncogènes associés aux lymphomes des cellules B, tels que Bcl-6, Pim1 et c-Myc (Liu et al., 2008). Il est encore difficile aujourd'hui de donner le mécanisme précis menant à la réparation mutagène des résidus uraciles induits par AID dans les loci Ig.

Chez les mammifères, il existe quatre uracile-ADN glycosylases qui sont MBD4, SMUG1, TDG et UNG (Stratigopoulou et al., 2020). L'ADN glycosylase UNG est l'uracile glycosylase majeure chez les mammifères qui est responsable de l'activité d'excision des uraciles dans l'ADN durant les mécanismes de CI et d'HMS (Akbari, 2004; Petersen-Mahrt et al., 2002; Rada et al., 2002b). Les trois autres uracile-ADN glycosylases MBD4, SMUG1 et TDG ne peuvent pas compenser la perte d'UNG dans

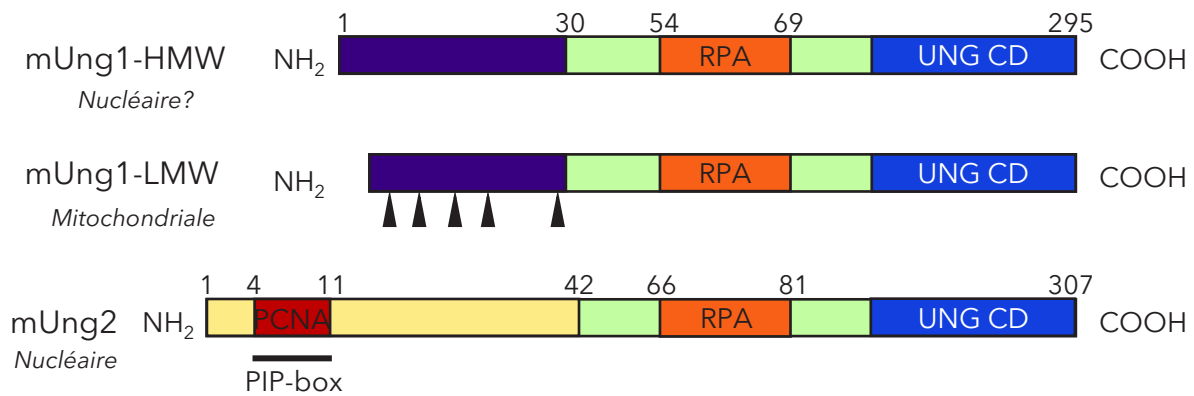


Figure 13: Les domaines fonctionnels des isoformes de l' Uracil-DNA-Glycosylase (UNG). La transcription du locus Ung aboutit à l'expression de 2 isoformes appelées Ung1 et Ung2. Ces deux isoformes varient par leur séquence N-terminale (rectangle violet pour Ung1 et jaune pour Ung2). Ung1 existe sous la forme de deux variants : un de haut poids moléculaire appelé UNG1-HMW (high molecular weight, 31kDa) et un de plus faible poids moléculaire, Ung1-LMW (low molecular weight, 26kDa). Ces deux variants présentent un site de liaison à RPA (résidus 54-69) ainsi que le domaine catalytique Ung (UNG CD) en C-ter qui est aussi présent dans l'isoforme Ung2. La séquence N-ter du variant Ung1-LMW est ciblée par un processus protéolytique conduit par la MPP (mitochondrial processing peptidase) qui permet de confirmer la présence de ce variant dans la mitochondrie. Une récente étude a montré que le variant Ung1-HMW se trouve dans le noyau comme l'isoforme Ung2 (Sarno et al. 2019). De plus, l'isoforme Ung2 contient un site de liaison à PCNA, appelé PIP-box, dans sa séquence N-ter et à RPA (résidus 66-81). Adapté de (Sarno et al. 2019).

les mécanismes de diversification des anticorps (Bardwell et al., 2003; Di Noia et al., 2006; Di Noia et al., 2007b; Imai et al., 2003b).

Le gène UNG code pour deux isoformes UNG1 et UNG2 qui diffèrent par leur séquence N-terminale, mais possèdent le même domaine catalytique et le même motif de liaison pour RPA. UNG1 se trouve dans la mitochondrie tandis qu'UNG2 est localisé dans le noyau (Otterlei et al., 1998). UNG1 comporte deux variants de tailles différentes avec la forme à plus haut poids moléculaire, appelée UNG1-HMW de 31 kDa et la forme à plus faible poids moléculaire, appelée UNG1-LMW de 26 kDa qui a subi un traitement par la peptidase de transformation mitochondriale (MPP) (Bharati et al., 1998; Sarno et al., 2019). L'isoforme nucléaire UNG2 a été décrit comme l'isoforme jouant un rôle clé dans la CI et l'HMS (Akbari, 2004; Doseth et al., 2012; Kavli et al., 2002; Krijger et al., 2009; Pettersen et al., 2007; Sharbeen et al., 2012). Néanmoins, très récemment, il a été montré, à l'aide de western blots d'extraits nucléaires à partir de la lignée cellulaire CH12 pWT et de cellules B primaires, que la forme UNG1-HMW est présente dans le noyau et qu'elle constitue 25% d'UNG total nucléaire dans les CH12 pWT (Sarno et al., 2019). De plus, ce variant d'UNG1 permet de soutenir la CI en absence d'UNG2 dans la lignée cellulaire CH12 (Sarno et al., 2019). Toutefois, son rôle précis reste à être déterminé. En plus de son interaction avec RPA, UNG2 peut interagir avec PCNA par son motif PIP-box N-terminal (UNG2⁴⁻¹¹) (Rodriguez et al., 2017). Il faut d'ailleurs noter qu'UNG2 comporte 2 sites de liaison pour RPA (Otterlei et al., 1999) UNG2⁶⁶⁻⁸¹ (Sarno et al., 2019) et un deuxième site qui n'est pas encore totalement déterminé (Otterlei et al., 1999; Weiser, 2020). L'isoforme UNG2 humain contient plusieurs sites de phosphorylation qui peuvent potentiellement réguler l'activité glycosylase d'UNG2. En effet, la phosphorylation de Thr6 ou Tyr8 sur UNG2 peut entraver l'interaction PCNA-UNG2 sans affecter l'activité catalytique d'UNG2 ou son interaction avec RPA (Weiser et al., 2017). En revanche, la phosphorylation de Thr6 et Thr126 sur UNG2 a permis de voir une augmentation de l'activité glycosylase d'UNG2 *in vitro*. De plus, il a été montré que PPM1D peut inhiber la voie BER en déphosphorylant UNG2 sur Thr6 et Thr126 pour faciliter son inactivation après l'achèvement de la réparation de l'ADN (Lu et al., 2004). Ceci est contredit par des études montrant qu'UNG non phosphorylée est catalytiquement hautement active

(Hagen et al., 2008; Kavli et al., 2002; Torseth et al., 2012). L'ensemble de ces modifications post-traductionnelles n'est pas conservé chez la souris mais révèle une régulation très bien orchestrée de l'isoforme UNG2 humain. De plus, un variant hétérozygote d'UNG2 à un seul nucléotide menant à la substitution Arg88Cys (R88C) dans le motif d'interaction de RPA a été observé chez l'homme. Cette substitution abolie l'interaction RPA-UNG2 ainsi que le recrutement d'UNG à l'ADN simple brin (Torseth et al., 2012). UNG2 est la glycosylase prédominante dans l'élimination de l'U et est active sur l'ADN simple brin (ss) et double brin (ds) (Stratigopoulou et al., 2020). Son expression culmine en phase S, mais UNG2 est également présente et requise durant la phase G1 du cycle cellulaire (Schrader et al., 2007b; Sharbeen et al., 2012) afin d'éliminer les uraciles induits par AID durant les mécanismes de CI et d'HMS (Imai et al., 2003b; Rada et al., 2002b). L'excision de l'uracile par UNG2 en phase G1 induit principalement une réparation correcte ainsi qu'un traitement mutagène (Sharbeen et al., 2012).

La déficience des facteurs du BER, comme UNG et APE1 dans les cellules B humaines (Imai et al., 2003b; Kavli et al., 2005) et murines (Rada et al., 2002b; Schrader et al., 2005) entraîne un défaut dans la formation des DSBs dans les régions S (Imai et al., 2003b) et donc une altération sévère de la CI (Imai et al., 2003b; Kavli et al., 2005; Rada et al., 2002b; Ranjit et al., 2011; Schrader et al., 2005; Yousif et al., 2014).

Également, l'absence des facteurs du MMR, tels que MSH2, MSH6, Exo1, MLH1 et PMS2 provoque une diminution de l'efficacité de commutation isotypique, avec des degrés variables (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). Enfin, un défaut dans les deux voies combinées (UNG^{-/-} MSH2^{-/-}) abolit complètement la CI (Xue et al., 2006).

Dans la voie détournée du BER, les uraciles sont donc reconnus et éliminés par UNG2 (Pettersen et al., 2007; Sarno et al., 2019; Sharbeen et al., 2012), menant à la création d'un site AP. L'endonucléase APE1 (Pettersen et al., 2007) va reconnaître le site abasique et introduire une SSB dans l'ADN. Des SSBs adjacentes sur des brins d'ADN complémentaire génèrent des DSBs. La voie BER est la voie majoritaire dans

l'élimination des uraciles durant la CI *in vitro* et *in vivo* (Imai et al., 2003b; Kavli et al., 2005; Rada et al., 2002b; Schrader et al., 2005). Il a été montré que l'altération du recrutement d'UNG aux régions S dans les cellules B Rev1^{-/-} réduit la CI (Zan et al., 2012).

Les mésappariements dU:dG peuvent aussi être reconnus par l'hétérodimère MSH2/MSH6 qui ensuite mobilise les protéines adaptatrices du complexe MLH1/PMS2 afin de recruter les facteurs PCNA-Ub et EXO1. EXO1 va ensuite dégrader l'ADN ébréché pour produire des cassures simple brin (Choi et al., 2020). PCNA-Ub est important pour une CI efficace mais son rôle exact n'est pas encore connu (Chahwan et al., 2012).

De plus, les facteurs du MMR peuvent être impliqués dans la génération de DSBs suite à des SSBs générées par UNG2 et APE1 (Schrader et al., 2007b). En effet, Exo1 peut exciser l'ADN de la SSB 5' la plus proche générée par UNG2 et APE1 jusqu'à ce qu'elle atteigne une discordance sur l'autre brin, créant ainsi une DSB avec une extrémité sortante 5' (Schrader et al., 2005). Cette extrémité sortante peut être complétée par les polymérases translésionnelles (η et θ), en laissant des DSBs à extrémités franches (Schrader et al., 2005) ou presque. Ces DSBs vont activer les facteurs de la réponse des dommages à l'ADN (DDR) afin de favoriser la recombinaison à longue distance et la réparation via la voie classique des extrémités non homologues (C-NHEJ) (Casellas et al., 1998; Manis et al., 1998a) ou bien la voie alternative (A-EJ) (Eccleston et al., 2011).

4. La DDR

L'observation de foyers de réparation formés par γ H2AX et Nbs1, qui s'accumulent dans le locus IgH des cellules B subissant la commutation isotypique avec une formation et localisation dépendante d'AID, relie pour la première fois la réponse des dommages de l'ADN au mécanisme de CI (Petersen et al., 2001) (**Figure 14**). Cette étude montre que le rôle d'AID se situe en amont de la formation des DSBs durant la CI et que la DDR est impliquée dans le maintien de l'intégrité génomique des lymphocytes B pour faire face aux dommages de l'ADN induits par AID (Petersen et al., 2001). De nombreuses études portant sur le rôle des différents facteurs de la DDR ont été réalisées d'une part, sur des patients déficients pour

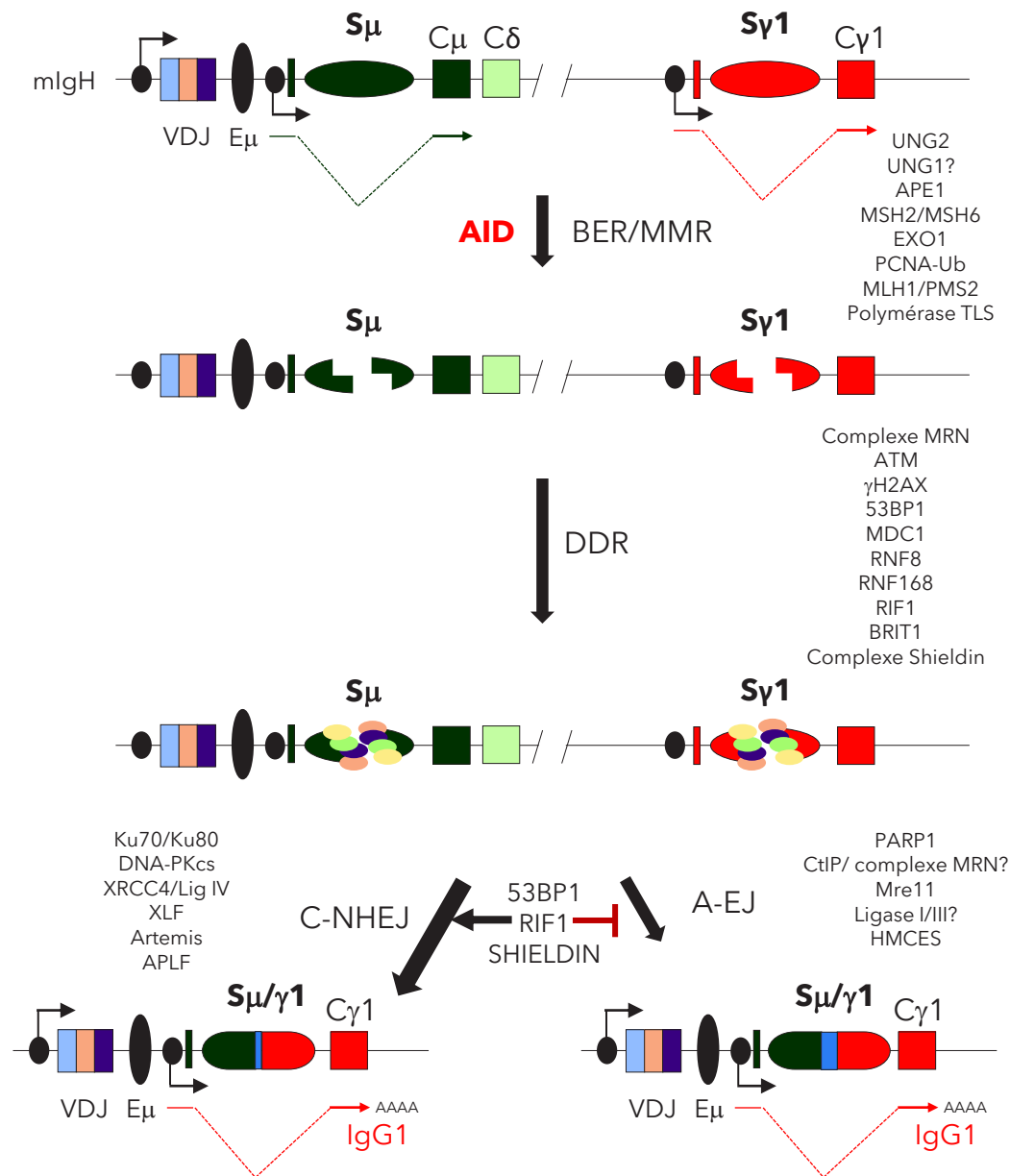


Figure 14: Formation et réparation des cassures double brin (DSBs) durant la CI. Durant la CI, la transcription germinale des régions S donneuse ($S\mu$) et acceptrice ($S\gamma1$) mène à l'exposition d'ADN simple brin, substrat d'AID. AID va donc induire la désamination des cytosines dans l'ADN en entraînant des mésappariements dU:dG. Ces mésappariements sont ensuite traités par des voies détournées du BER et MMR qui vont permettre la formation de DSBs, nécessaires à la recombinaison entre la région donneuse $S\mu$ et une région acceptrice Sx , ici $S\gamma1$). Les DSBs générées vont être tout d'abord reconnues par la voie de réponse des dommages à l'ADN (DDR), puis réparées par les voies des jonctions non-homologues classique (C-NHEJ) ou bien alternative (A-EJ).

certaines acteurs et d'autre part, par des analyses d'inactivation ou de mutations dans le modèle souris ou dans la lignée CH12. En effet, l'inactivation des acteurs tels que ATM (Pan-Hammarstrom et al., 2003; Pan et al., 2002; Reina-San-Martin et al., 2004), le complexe MRN composé de Mre11, Rad50 et Nbs1 (Dinkelmann et al., 2009b; Kracker et al., 2005; Lahdesmaki et al., 2004; Reina-San-Martin et al., 2005), H2AX (Celeste et al., 2002; Franco et al., 2006; Reina-San-Martin et al., 2003), MDC1 (Lou et al., 2006), RNF8 et RNF168 (Li et al., 2010; Ramachandran et al., 2010; Santos et al., 2010) provoque des défauts de CI avec une réduction de 30 à 80% et augmente les anomalies chromosomiques et les translocations oncogéniques (Bassing et al., 2003; Celeste et al., 2002). La déficience de 53BP1 montre une abolition presque totale de la CI (90% de réduction de CI), avec une présence significative d'aberrations chromosomiques, suggérant un second rôle particulièrement important dans la CI. 53BP1 favorise la synapse S-S en coopération avec les facteurs ATM et H2AX, et l'assemblage de longue portée des régions S (Dudley et al., 2002; Manis et al., 2004; Reina-San-Martin et al., 2007; Ward et al., 2004). 53BP1 protège aussi les extrémités libres d'ADN avec l'aide d'H2AX pour diriger la réparation de l'ADN vers la voie NHEJ (Bothmer et al., 2011; Bothmer et al., 2010), en inhibant la voie de la HR (Bunting et al., 2010). Cette protection des extrémités libres de l'ADN est appuyée par le recrutement de Rif1 (Chapman et al., 2013; Di Virgilio et al., 2013; Escribano-Diaz et al., 2013; Feng et al., 2013; Zimmermann et al., 2013) suite à la phosphorylation de 53BP1 par ATM. De plus, le complexe 53BP1-RIF1 inhibe la réparation des DSBs par la voie A-EJ (Bothmer et al., 2011; Bothmer et al., 2010; Nicolas et al., 2018). Très récemment, il a été montré que le complexe Shieldin, composé de Shld1, Shld2, Shld3 et Rev7, empêche la résection des extrémités de l'ADN en concert avec 53BP1 et Rif1 (Ghezraoui et al., 2018; Gupta et al., 2018; Liang et al., 2020; Noordermeer et al., 2018). Le mécanisme du complexe Shieldin permettant d'empêcher la résection n'est pas encore clair mais requiert le recrutement de Shld2 à l'ADN simple brin exposé aux DSBs.

D'une manière générale, ATM est recrutée par le complexe MRN au niveau des cassures, puis va subir une auto-phosphorylation et va ensuite phosphoryler d'autres protéines effectrices comme H2AX (Burma et al., 2001), Nbs1 (Falck et al., 2005), 53BP1 (Anderson et al., 2001; Rappold et al., 2001), MDC1 (Lou et al., 2006)

et KAP1 (White et al., 2006; Ziv et al., 2006). La phosphorylation de H2AX sur la sérine 139 forme γ H2AX (Rogakou et al., 1998) et mène au recrutement de MDC1 (Stucki et al., 2005), 53BP1 (Ward et al., 2003) et Nbs1 (Kobayashi et al., 2002) qui s'accumulent et s'étendent au site de dommage en formant des foci. La phosphorylation de MCD1 entraîne le recrutement de l'ubiquitine ligase RNF8, responsable de l'ubiquitination de l'histone de type H2A (Huen et al., 2007; Kolas et al., 2007; Mailand et al., 2007) qui a son tour recrute RNF168, impliquée dans la stabilisation de 53BP1 aux DSBs (Doil et al., 2009).

Très récemment, le facteur BRIT1 a été identifié comme étant un nouveau composant de la DDR (Yen et al., 2017). Il a été montré que BRIT1 est requis pour la CI et qu'il est recruté dans le locus IgH de manière dépendante d'AID et de H2AX avec lequel il interagit grâce à son domaine BRCT (Yen et al., 2017). Lié aux DSBs du locus IgH, il est proposé que BRIT1 pourrait servir de plateforme pour recruter les facteurs qui résolvent les DSBs induites par AID pour mener à bien la CI ; l'identité des facteurs est encore inconnue (Yen et al., 2017).

5. La réparation des DSBs

Dans les cellules de mammifères, les cassures double brin à l'ADN sont réparées par deux voies de réparation à l'ADN : la recombinaison homologue (HR) et la religation d'extrémités non homologues (Alt et al., 2013). La HR est restreinte à la phase S/G2 du cycle cellulaire et nécessite de grandes étendues d'homologie tandis que le NHEJ est actif tout au long du cycle cellulaire et ne demande pas ou peu d'homologie. Puisque les DSBs associées à la CI sont observées principalement pendant la phase G1, phase durant laquelle la HR n'opère pas, et n'ont pas un niveau d'homologie suffisante pour fournir le substrat de la HR (Chaudhuri et al., 2007), le NHEJ est donc considéré comme la voie majeure de réparation des DSBs durant la CI (Alt et al., 2013; Boboila et al., 2012). De plus, la plupart des jonctions obtenues suite à la recombinaison entre deux régions S expose des caractéristiques de réparation par C-NHEJ car elles ont tendance à présenter des jonctions directes ou des jonctions qui impliquent des micro-homologies (MH) très courtes (<4 nucléotides) (Boboila et al., 2010; Yan et al., 2007) ou bien des insertions courtes (Dunnick et al., 1993; Yan et al., 2007).

5.1 La réparation des DSBs par le C-NHEJ

Le C-NHEJ est initié par le recrutement de l'hétérodimère Ku70-Ku80 aux extrémités des DSBs qui va à son tour recruter et activer la DNA-PKcs au site de cassure double brin pour former un complexe établissant une synapse de longue-portée (Graham et al., 2016) (**Figure 14**). La formation de l'holo-complexe DNA-PKcs-Ku-ADN va activer la fonction kinase de DNA-PKcs, essentielle pour la réparation par C-NHEJ (Dobbs et al., 2010; Jette and Lees-Miller, 2015; Jiang et al., 2015). Les facteurs Ku70, Ku80 et DNA-PKcs sont requis pour la CI (Bjorkman et al., 2015; Casellas et al., 1998; Cook et al., 2003; Franco et al., 2008; Manis et al., 1998a). Le recrutement du complexe de l'ADN ligase IV et XRCC4 par l'holo-complexe va activer la ligation des extrémités de l'ADN (Grawunder et al., 1997). La nucléase Artemis (Franco et al., 2008; Rivera-Munoz et al., 2009) et le facteur XLF (appelé aussi Cernunnos) (Du et al., 2012; Li et al., 2008) sont aussi impliqués dans la réparation des DSBs induites par AID. En effet, si les extrémités de l'ADN ne sont pas compatibles pour la ligation, DNA-PKcs mobilise la nucléase Artemis afin de modifier les extrémités d'ADN (Franco et al., 2008). Le facteur XLF ainsi qu'APLF favorisent la ligation en stimulant l'activité du complexe de l'ADN ligase IV et de XRCC4 (Ahnesorg et al., 2006; Callebaut et al., 2006; Frit et al., 2019) dans une synapse de courte portée (Graham et al., 2016).

Alors que les facteurs Ku70, Ku80 et DNA-PKcs peuvent être impliqués en dehors de la voie NHEJ (Sekiguchi et al., 2001), l'ADN ligase IV et XRCC4 sont considérés comme les facteurs essentiels du NHEJ car ils n'ont aucun rôle connu en dehors de cette voie. L'observation d'une présence résiduelle de commutation de classe en l'absence des facteurs essentiels du C-NHEJ et l'analyse des jonctions des cellules appartenant à cette présence résiduelle a mené à la découverte d'une voie alternative, appelée aujourd'hui A-EJ. En effet, les travaux de trois différents groupes ont montré tout d'abord que les facteurs ADN ligase IV et XRCC4 sont requis pour une CI efficace, impliquant la voie du NHEJ dans la résolution des DSBs durant la CI car leur absence altère significativement l'efficacité de la CI. Cependant, l'analyse des jonctions de « switch » dans les cellules qui ont réussi à effectuer la CI a montré une augmentation significative de l'utilisation de micro-homologies plus

longues et une perte complète des jonctions franches obtenues par ligation directe (Soulas-Sprauel et al., 2007; Yan et al., 2007). L'augmentation de la longueur des micro-homologies (>7pb) a été rapporté chez des patients porteurs de mutations dans le gène codant de l'ADN ligase IV (Pan-Hammarström et al., 2005). Suite à ces travaux, il a aussi été montré qu'en l'absence des facteurs de reconnaissance des DSBs Ku, la voie A-EJ permet la résolution des DSBs durant la CI (Boboila et al., 2012; Boboila et al., 2010). Ensemble, ces résultats montrent la participation d'une voie alternative de la jonction des extrémités dans la réparation des DSBs durant la CI qui est distincte du NHEJ, laquelle est basée sur l'utilisation de longues micro-homologies. Cette voie A-EJ, n'est aujourd'hui pas encore bien définie (**Figure 14**).

5.2 La réparation des DSBs par le A-EJ

La voie A-EJ montre une préférence pour la jonction des extrémités d'ADN qui présentent une micro-homologie de 4 à 20 nucléotides (Pannunzio et al., 2018). Il a donc été proposé que la réparation par A-EJ commence par une étape de résection limitée au DSB afin d'exposer de courts segments d'ADN simple brin homologues à l'autre extrémité. Par la suite, ces derniers s'hybrident avant d'entrer dans l'étape de ligation. Les mécanismes précis de la réparation par A-EJ ne sont pas encore connus et il est encore difficile à dire si l'A-EJ représente une ou plusieurs voies. Parp1 est mobilisé dans la reconnaissance des extrémités de l'ADN (Audebert et al., 2004; Robert et al., 2009; Wang et al., 2006b) et favoriserait par son interaction avec Mre11, le recrutement de Mre11 (Dinkelmann et al., 2009a; Rass et al., 2009; Xie et al., 2009) et de CtIP (Bennardo et al., 2008; Lee-Theilen et al., 2011; Yun and Hiom, 2009) qui entraînent la résection des extrémités. L'état phosphorylé de CtIP semblerait stimuler la fonction endonucléase de MRN et donc initier le traitement des extrémités de l'ADN (Chang et al., 2017). L'étape de ligation peut impliquer par ailleurs, de multiples complexes contenant des ADN ligases. En effet, il a été montré que l'ADN ligase III (Lu et al., 2016; Paul et al., 2013; Simsek et al., 2011) et I (Lu et al., 2016; Paul et al., 2013) interviennent lors de la ligation durant la réparation par A-EJ, mais le mécanisme précis n'est pas clair (Masani et al., 2016). Il semblerait que les ADN ligase I et III peuvent se compenser l'une envers l'autre (Lu et al., 2016; Masani et al., 2016). Il a récemment été exposé que Rad52 facilite la voie A-EJ en

rentrant en compétition avec le complexe Ku, par sa liaison aux DSBs (Zan et al., 2017). Encore plus récemment, une étude a révélé que le facteur HMCES (5-Hydroxymethylcytosine binding, ES-cell-specific) connu pour son rôle de capteur et bouclier des sites AP dans l'ADN simple brin (Mohani et al., 2019), est requis durant la CI et joue un rôle important dans la voie A-EJ (Shukla et al., 2020). En effet, les cellules B primaires et la lignée cellulaire CH12 déficientes en HMCES ont montré une réduction significative de CI et l'utilisation de micro-homologies dans les régions de switch comparé aux WT (Shukla et al., 2020). Il a été proposé que HMCES se lie aux régions de switch et protège les extrémités sortantes d'ADN simple brin via son domaine peptidase associé à la réponse SOS (SOS-response-associated-peptidase domain)(Halabelian et al., 2019).

Il est important de noter que cette voie alternative A-EJ est beaucoup moins fidèle que le C-NHEJ, c'est en effet une voie très exposée aux erreurs et provoque des délétions et des insertions. La voie A-EJ est aussi reconnue pour catalyser de nombreux réarrangements génomiques, conduisant couramment à des translocations chromosomiques oncogéniques (Roth, 2002; Simsek and Jasin, 2010; Wang et al., 2008; Yan et al., 2007). Plusieurs études ont montré que la voie A-EJ n'est pas seulement présente dans des cellules déficientes en C-NHEJ, elle joue aussi un rôle important dans la réparation des DSBs dans des cellules avec une voie C-NHEJ intacte. En effet, il a été révélé que des cellules compétentes pour le C-NHEJ montre des défauts de CI et de jonctions de switch en absence de facteurs impliqués dans la réparation par A-EJ (Cortizas et al., 2013; Lee-Theilen et al., 2011; Robert et al., 2009). Il semble donc important de continuer à décortiquer cette voie alternative afin d'élucider les mécanismes précis impliqués.

Les mécanismes déterminant le choix entre l'utilisation de la voie A-EJ ou C-NHEJ restent toujours inconnus. Il a été proposé un modèle dans lequel la densité des désaminations induites par AID et la nature de la lésion déterminent le choix de réparation par C-NHEJ et A-EJ (Cortizas et al., 2013). Cependant, des résultats contradictoires ont été obtenus, montrant que 53BP1 est impliqué dans le choix entre la voie classique et alternative-EJ durant la CI par sa capacité à interférer avec l'étape de résection (Bothmer et al., 2010). De plus, il est possible que la voie A-EJ

soit en concurrence avec la C-NHEJ pour réparer les mêmes DSBs. En effet, plusieurs études ont mis en avant que PARP1 (Wang et al., 2006b) ou bien Rad52 (Zan et al., 2017) rentrent en compétition avec le complexe Ku pour initier la réparation des DSBs.

V- Les immunodéficiences de la commutation isotypique

Les immunodéficiences de la CI (CSR-ID), aussi appelées les syndromes Hyper-IgM (HIGM) sont des immunodéficiences primaires rares (1/100 000 naissances) dues à des défauts génétiques qui mènent à une déficience de commutation isotypique (CSR-D). Elles sont caractérisées par la présence d'un taux normal ou élevé d'IgM associé à une absence ou une présence à un niveau très bas des autres isotypes (IgG, IgA et IgE). Elles peuvent être aussi combinées à un défaut d'hypermutation somatique (PéRon et al., 2007). Ces syndromes HIGM ont pour conséquence une susceptibilité particulière aux infections bactériennes. Une commutation isotypique altérée peut être due à un défaut d'interaction entre les cellules T et les cellules B, d'activation de la voie intracellulaire dans les cellules B, de mécanismes intrinsèques des cellules B (Durandy et al., 2005) et enfin de machinerie de réparation de l'ADN (PéRon et al., 2007).

1. Les CSR-ID dues aux défauts de l'activation des cellules B

Les mécanismes de CI et de d'HMS sont initiés par l'interaction entre les cellules B et les cellules T qui se produit via l'interaction entre le CD40 ligand (CD40L), exprimé à la surface des lymphocytes T CD4+ activés, et CD40, exprimé à la surface des lymphocytes B ainsi que des cellules dendritiques. Cette interaction est indispensable pour la maturation des lymphocytes B au sein des organes lymphoïdes secondaires (Nonoyama et al., 1993; Ochs et al., 1994). Des mutations dans le gène codant pour le ligand de CD40 (Allen et al., 1993; Aruffo et al., 1993; Disanto et al., 1993; Fuleihan et al., 1993; Korthäuer et al., 1993) causent la plus fréquente des CSR-ID, appelée HIGM de type 1 (HIGM1), provoquant chez les patients de récurrentes infections bactériennes dues à l'absence de production des isotypes IgA et IgG. Cette déficience de CD40L a une transmission liée au chromosome X et les cellules B des patients sont incapables de proliférer et de former des centres germinatifs dans les organes lymphoïdes secondaires, exposant donc un défaut de CI et d'HMS *in vivo* (Van Zelm et al., 2014). Cependant, les cellules B sont capables de subir la CI vers les autres isotypes *in vitro* (Durandy et al., 1993). La détection d'IgA dans le sérum et d'HMS chez certains patients suggère qu'une

voie alternative de diversification peut se produire : la CI vers IgA pourrait être médiée par le ligand inducteur de prolifération (APRIL) dans la lamina propria de l'intestin (He et al., 2007), tandis que l'HMS indépendante des cellules T pourrait éventuellement refléter un mécanisme de défense inné (Scheeren et al., 2008). De plus, ce défaut affecte aussi les lymphocytes T en entraînant une incapacité d'interagir avec les monocytes et les cellules dendritiques.

Un défaut génétique de transmission autosomique récessive dans le gène CD40 a été décrit chez quelques patients (Ferrari et al., 2001), montrant l'absence d'expression de CD40 à la surface des cellules B et des monocytes (Ferrari et al., 2001). Cette nouvelle forme d'HGM, appelée HGM3, montre les mêmes profils cliniques et immunologiques que l'HGM1, excepté que les cellules B sont incapables de subir une CI *in vitro*.

La voie NF- κ B activée par l'interaction entre CD40L et CD40 est aussi impliquée dans l'activation des cellules B. Des défauts dans l'expression de NEMO, modulateur essentiel de la voie NF- κ B, aboutissent aussi à une CSR-ID, appelée EDA-ID (ectodermal dysplasia associated with immunodeficiency) (Döffinger et al., 2001; Hanson et al., 2008; Jain et al., 2004). L'absence de l'expression de NEMO cause des niveaux très bas d'IgA et d'IgG dans le sérum et des réponses immunitaires affectées, conduisant à la susceptibilité d'infections myco-bactériennes. Dû à l'hétérogénéité de ce syndrome, la CI et l'HMS *in vitro* peuvent être soit défectueuses, soit normales (Hanson et al., 2008; Jain et al., 2004).

2. Les déficiences de la CI dues aux défauts intrinsèques des cellules B

Bien que les défauts de la signalisation médiée par CD40 conduisent à un déficit immunitaire qui implique en plus des cellules B, d'autres cellules immunitaires ; d'autres formes de syndromes hyper-IgM sont causées par des défauts intrinsèques et sélectifs aux cellules B.

2.1. La déficience de la CI causée par des mutations dans le gène Aicda

La déficience de la CI à transmission autosomale récessive causée par des mutations dans le gène AICDA codant pour AID (Revy et al., 2000), appelée HIGM2, est caractérisée par une altération des deux mécanismes de diversification des anticorps, la CI et l'HMS (Durandy et al., 2006). Cependant, chez certains patients portant des mutations localisées dans la région C-terminale, le défaut de CI est associé à une réaction d'HMS normale, confirmant le rôle de la région C-terminale d'AID dans l'interaction de cofacteurs spécifiques de la CI (Durandy et al., 2006; Ta et al., 2003). Ces observations ont été confirmées par l'analyse d'un mutant artificiel d'AID, dans lequel les 10 derniers résidus sont absents, qui est capable de réaliser l'HMS et la conversion génique mais en aucun cas la CI (Barreto et al., 2003).

2.2 La déficience de la CI causée par un défaut de la réparation de l'ADN

Les lésions dans l'ADN induites par AID activent la machinerie de réparation de l'ADN conduisant à la reconnaissance et à l'élimination des résidus uraciles présents dans l'ADN par UNG2. La déficience d'UNG est une rare cause de déficience de la CI (Imai et al., 2003b), appelée HIGM5, avec seulement trois cas décrits qui présentent un défaut de commutation isotypique *in vivo* et *in vitro*. L'analyse du profil de l'HMS dans les souris déficientes pour UNG a montré un profil de mutations biaisé (Imai et al., 2005). Ces trois cas présentent des infections bactériennes récurrentes et une hyperplasie lymphoïde. Comme le facteur UNG fait partie de la voie BER qui est impliquée dans la réparation de mutations spontanées, il constitue un facteur majeur de défense anti-mutagène. D'ailleurs, les souris déficientes pour UNG ont développé des tumeurs en vieillissant (Nilsen et al., 2003). Il est donc possible que la déficience en UNG prédispose les patients à la formation de tumeurs à l'âge adulte.

De plus, des mutations dans le gène PMS2, facteur de la voie de réparation des mésappariements, mènent à la survenue précoce de divers cancers (Nakagawa et al., 2004; Niessen et al., 2006), et à de récurrentes infections bactériennes. Ceci est

dû à de faibles taux d'IgG2 et d'IgG4 dans le sérum et à des taux totalement réduits d'IgG ou d'IgA ou bien les deux chez certains patients. Ces patients présentent donc un défaut de CI *in vivo* et *in vitro* mais une réaction d'HMS normale (Péron et al., 2008). Cette CSR-ID est aussi appelée syndrome de Lynch.

Etant donné que les mécanismes de réparation de l'ADN du BER, MMR et du NHEJ sont impliqués dans la synapse de la région intra-switch, une déficience de CSR n'est pas surprenante dans les déficiences d'ATM (Etzioni et al., 2007; Noordzij et al., 2009), MRE11, NBS1, XLF, ADN ligase 4 ou bien Artemis (de Miranda et al., 2011).

Enfin, un dernier syndrome Hyper-IgM (HIGM4) a été décrit par Imai et al. (Imai et al., 2003a) chez des patients avec des syndromes CI/HIGM génétiquement indéfinis. Ces patients présentent les mêmes caractéristiques cliniques et immunologiques que la plupart des formes de syndromes HIGM citées précédemment.

Hypothèse de travail

Au cours des réponses immunes, le répertoire des cellules B est diversifié par les mécanismes d'hypermutation somatique et de commutation isotypique, menant à l'établissement d'une réponse humorale spécifique, adaptée et durable. La commutation isotypique modifie les propriétés effectrices des anticorps en induisant le changement de l'isotype exprimé. Ce mécanisme a lieu au sein du locus IgH et implique un événement de recombinaison longue distance au niveau des régions de switch. Mécaniquement, la CI est initiée par la désamination des cytosines dans l'ADN par l'enzyme AID. Les uraciles résultant de cette désamination sont principalement excisés par UNG, ce qui entraîne la génération de cassures double brin. Celles-ci activent la réponse des dommages de l'ADN puis, sont principalement réparées par la voie de jonction des extrémités non homologues (NHEJ). Des modèles récents ont permis d'expliquer plusieurs caractéristiques du mécanisme de la commutation isotypique, mais de nombreuses questions concernant, par exemple, le ciblage d'AID aux régions S et le choix de la réparation des lésions induites par AID, restent encore non élucidées. Notre hypothèse est qu'il existe des mécanismes encore inconnus qui permettent la régulation de la CI. Afin d'éclaircir l'ensemble des mécanismes moléculaires mis en jeu lors de la CI, l'objectif de ma thèse a été d'identifier les gènes impliqués lors de la CI par un criblage fonctionnel à large échelle à l'aide de la nouvelle technologie CRISPR/Cas9.

Cette nouvelle approche non biaisée permettant l'identification des co-facteurs de la CI sera décrite dans la première partie de mes résultats, suivie par la caractérisation fonctionnelle de *Fam72a*, un gène candidat issu du criblage fonctionnel, dans la CI et l'HMS. Les dernières parties de mes résultats sont constituées de deux projets auxquels j'ai contribué et portant sur la caractérisation fonctionnelle de *Zbtb24*, un gène impliqué dans le syndrome ICF, et de l'enhancer $\gamma 1$ dans la CI.

Résultats

CHAPITRE 1 :

**Vers l'identification de gènes impliqués
dans la commutation isotypique.**

Contexte

La commutation isotypique est un mécanisme complexe qui comporte plusieurs étapes : la transcription des GLTs, l'expression d'AID, le ciblage d'AID aux régions S, la désamination des cytosines, la formation des DSBs, leur reconnaissance et leur réparation pour donner lieu à une recombinaison efficace. Le grand nombre de facteurs et structures impliqués dans chacune de ces étapes montre qu'elles sont dépendantes d'un « effort d'équipe ». Malgré l'enrichissement de connaissances sur les différents mécanismes moléculaires impliqués durant la CI, la régulation de ceux-ci est encore floue et de nombreuses questions restent non élucidées. Notamment, les plus importantes du moment portent sur la compréhension des mécanismes menant au ciblage d'AID, à la réparation mutagène des résidus uraciles induits par AID dans le locus des Ig, au choix de la voie de réparation des cassures double brin ou encore si les mécanismes déjà étudiés dans le ciblage d'AID sont liés ou indépendants. Le projet de ma thèse a donc été d'identifier de manière globale tous les acteurs moléculaires impliqués durant la CI par une approche non biaisée.

Par le passé, plusieurs approches non biaisées ont été utilisées pour identifier les facteurs impliqués dans la commutation isotypique telles que l'identification de cofacteurs d'AID par co-immunoprécipitation et spectrométrie de masse (Jeevan-Raj et al., 2011; Thomas-Claudepierre et al., 2013) ou bien un criblage à large échelle à l'aide de shARNs (Pavri et al., 2010). Elles ont amélioré notre compréhension des mécanismes moléculaires contribuant à la CI et ont participé à établir le modèle actuel de recrutement couplé à la transcription d'AID. Cependant, ces approches présentent des inconvénients. En effet, la co-IP a été réalisée dans des cellules CH12 non stimulées surexprimant AID, cette condition ne correspond pas à ce que peut l'on retrouver de manière physiologique. Ensuite, l'utilisation de cribles à large échelle à l'aide de shARNs a montré la présence d'effets hors-cible et d'expression résiduelle ou faible de protéines qui peut masquer certains phénotypes, entraînant des taux élevés de faux-négatifs et de faux-positifs, et donc des résultats difficiles à interpréter (Peng et al., 2015; Shalem et al., 2015; Smith et al., 2017).

Nous avons donc choisi de réaliser un criblage fonctionnel CRISPR/Cas9 à l'échelle du génome murin afin d'identifier les cofacteurs de la CI.

I- Identification des gènes impliqués dans la commutation isotypique par criblage CRISPR/Cas9 à l'échelle du génome murin

La simplicité de la programmation et la grande flexibilité du système CRISPR/Cas9 (Cong et al., 2013) ont été exploitées pour permettre des criblages à l'échelle du génome dans des cellules de mammifères (Doench et al., 2016; Sanjana et al., 2014; Shalem et al., 2014), par la construction de bibliothèques de vecteurs permettant d'exprimer des collections d'ARN guides (ARNg).

Originellement décrit comme un système de défense chez la bactérie *Streptococcus Pyogenes* (Lander, 2016), le système CRISPR/Cas9 de type II (Clustered Regularly Interspaced Short Palindromic Repeats) a été optimisé pour en faire un outil d'édition du génome dans les cellules eucaryotes (Cong et al., 2013; Jinek et al., 2012; Mali et al., 2013). Celui-ci parvient à l'inactivation des protéines par des mutations induisant un décalage du cadre de lecture, introduites dans la région codante génomique cible. Ce système est composé d'un ARN guide (ARNg) spécifique de la région génomique cible, et de la nucléase Cas9. Cas9 reconnaît la séquence génomique cible sous la direction de l'ARNg qui s'hybride par complémentarité aux 20 nucléotides de la séquence cible, située en amont du motif NGG, aussi appelée la séquence PAM (Protospacer Adjacent Motif). Puis, la nucléase Cas9 induit des cassures double brin dans l'ADN, trois nucléotides avant le motif NGG. Ces cassures double brin sont ensuite réparées par la voie de jonction des extrémités non homologues (NHEJ). Cette réparation mutagène entraîne des délétions et des insertions, appelées « indels » pouvant mener à l'inactivation du gène cible provoquée par l'apparition d'un codon stop prématuré. D'autres « indels » dont la longueur est un multiple de trois ne donneront pas lieu à un décalage du cadre de lecture des gènes cibles. Dans ce scénario, lors d'un criblage fonctionnel CRISPR/Cas9, le produit du gène pourrait être fonctionnel (Cong et al., 2013; Shalem et al., 2014).

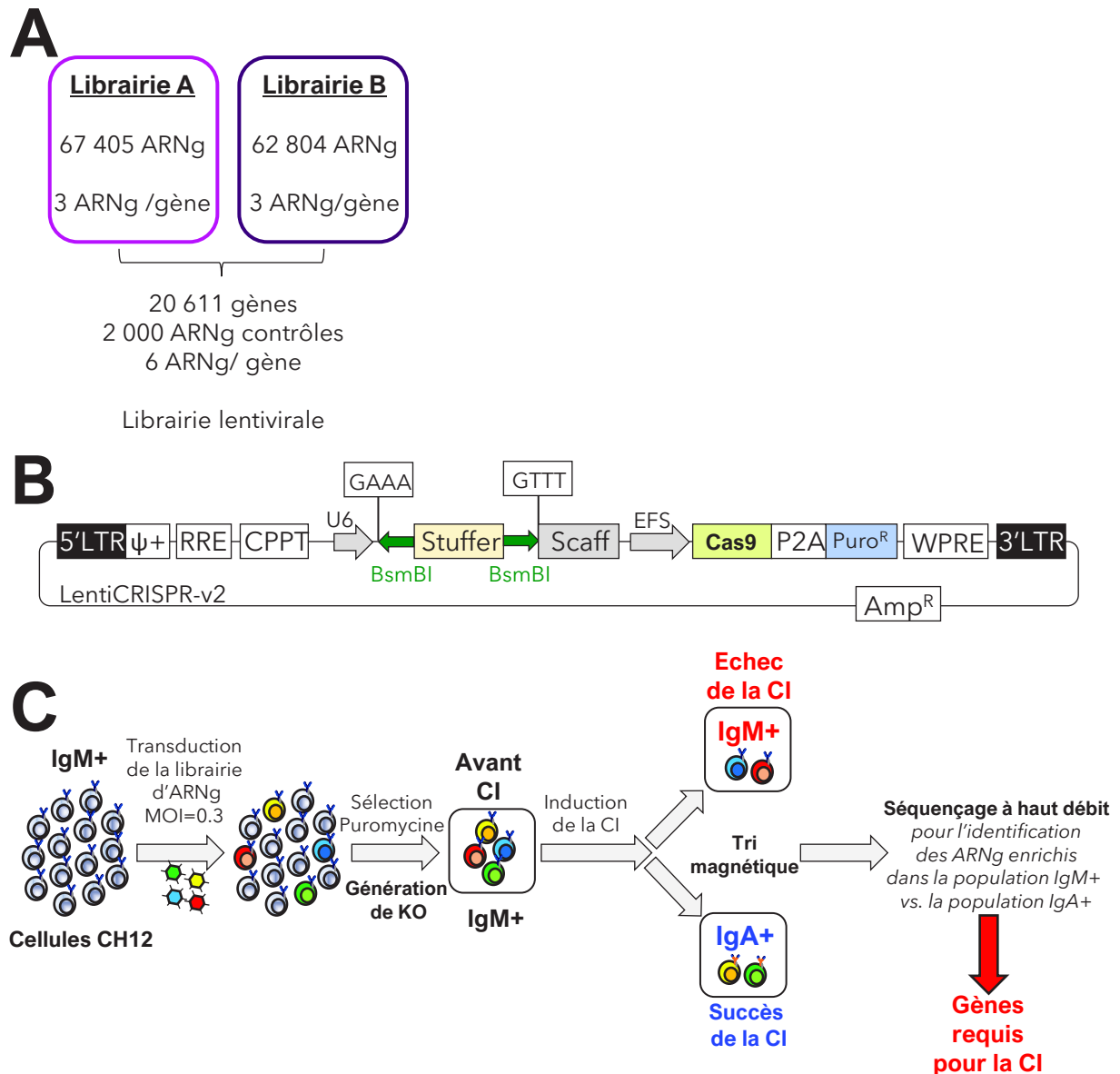


Figure 15: Stratégie du criblé d'inactivation par CRISPR/Cas9 sur la CI à large échelle.

A. La librairie lentivirale d'ARNg murine GeCKO-v2 (mGeCKO-v2) est constituée de deux sous bibliothèques d'ARNg, A et B qui contiennent ensemble 130 209 ARNg ciblant plus de 20 000 gènes avec 6 ARNg/gène, et 2 000 ARNg contrôles. **B.** Le vecteur de destination des bibliothèques d'ARNg, appelé LentiCRISPR-v2 est composé de deux cassettes d'expression, la nucléase Cas9 et la collection d'ARNg chimérique sous le contrôle du promoteur U6, de l'élément « scaffold » (Scaff) et d'une cassette de sélection permettant la résistance à la puromycine. **C.** Vue d'ensemble du criblage CRISPR/Cas9 à large échelle visant à identifier de nouveaux gènes impliqués dans la CI. Les cellules CH12 sont infectées avec la librairie lentivirale d'ARNg mGeCKO-v2 avec une MOI (multiplicity of infection) de 0.3 pour éviter de multiples infections/cellule. Les cellules infectées sont ensuite sélectionnées avec de la puromycine pendant plusieurs jours et enfin, elles sont soumises à l'essai fonctionnel de la CI avec l'ajout de TGF β , IL-4 et anti-CD40. Après 72h, les cellules incapables de réaliser la CI (restant donc IgM+) et celles qui ont subi des événements de CI productifs (exprimant IgA+) sont triées à l'aide de billes magnétiques couplées à des anticorps anti-IgM et anti-IgA. L'ADN génomique est extrait et les séquences des ARNg sont amplifiées par PCR pour ajouter les adaptateurs Illumina et les codes-barres. Enfin, le produit PCR est soumis au séquençage haut-débit pour identifier les ARNg qui sont enrichis dans la population IgM+ vs. la population IgA+.

1.1. Stratégie du criblage fonctionnel CRISPR/Cas9 sur la commutation isotypique

Afin d'identifier les cofacteurs de la CI, le criblage fonctionnel est effectué dans des cellules CH12 qui sont dérivées d'un lymphome murin à cellules B (Bishop and Haughton, 1986; Nakamura et al., 1996), en utilisant la bibliothèque lentivirale d'ARNg murine GeCKO-v2 (Genome-scale CRISPR-Cas9 knockout-Version2 (Sanjana et al., 2014)).

La lignée cellulaire CH12 est le modèle de référence utilisé pour étudier la commutation isotypique, car elle peut être efficacement induite à exprimer AID et à effectuer une recombinaison de changement de classe d'IgM à IgA après une stimulation de 72h avec un cocktail de TGF β , d'IL-4 et de l'anticorps @CD40 (Nakamura et al., 1996).

La bibliothèque mGeCKO-v2 conçue par le laboratoire de Feng Zhang couvre l'ensemble du génome murin, avec 130 000 ARNg ciblant environ 21 000 gènes dont des miARNs (4 ARNg/miARN) (**Figure 15A**). Chaque gène est ciblé par 6 ARNg répartis sur 3 ou 4 exons exprimés de manière constitutive. Cette librairie d'ARNg est divisée en deux sous-librairies, appelées A et B, chacune contenant 3 ARNg par gène et 1 000 ARNg contrôles non ciblés. La conception des ARNg est basée sur la sélection des exons constitutifs en excluant le premier et le dernier exon ainsi que tout exon contenant un site d'épissage alternatif. Chaque ARNg est sous forme (N)²⁰NGG, où le N représente un nucléotide. Afin de limiter les effets hors cible, la spécificité des séquences de (N)²⁰NGG des ARNg candidats a été testée sur l'ensemble du génome murin avec des paramètres acceptant jusqu'à 3 mésappariements de base. Enfin, pour inactiver les miARNs qui jouent un rôle clé dans la régulation transcriptionnelle, des ARNg orientant des mutations à la structure en épingles à cheveux du pré-miARN ont été additionnés (Sanjana et al., 2014).

Suite à la synthèse du « pool » des ARNg, la librairie d'ARNg peut être clonée dans le vecteur lentiviral appelé LentiCRISPR-v2 (**Figure 15B**). Celui-ci contient 2 cassettes d'expression, la nucléase Cas9 et la collection d'ARNg chimérique sous le contrôle du promoteur U6, un élément appelé « scaffold » qui permet la liaison de la

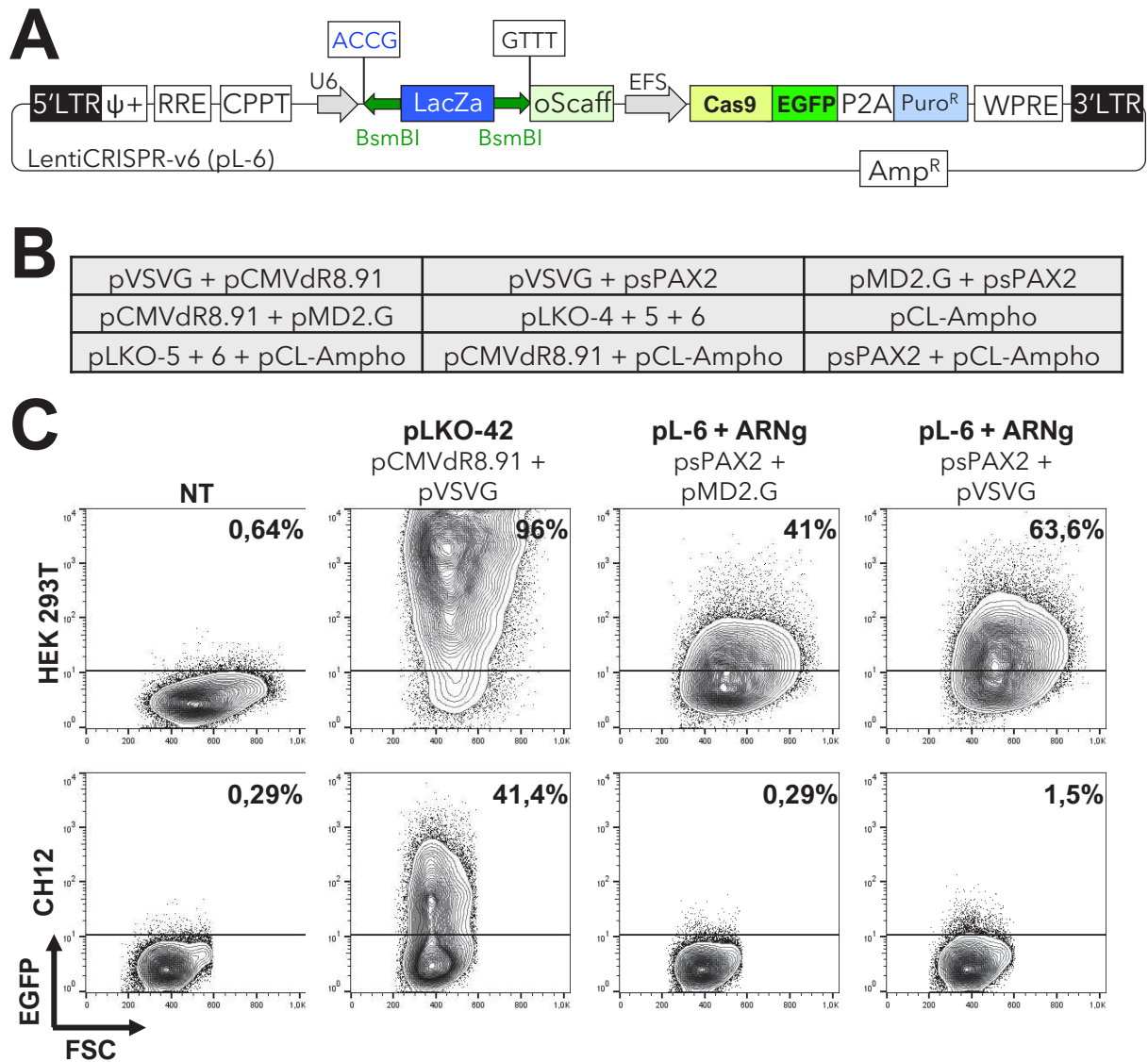


Figure 16: Le nouveau LentiCRISPR-v6 est fonctionnel mais n'infecte pas la lignée cellulaire CH12. A. Le nouveau LentiCRISPR-v6 (pL-6) a été amélioré par l'ajout d'une cassette LacZ à la place de l'élément « Stuffer » et d'un marqueur fluorescent EGFP, et le remplacement de l'élément « scaffold » par un scaffold optimisé (o-Scaff). **B.** Le test de fonctionnalité de ce nouveau pL-6 avec un ARNg cloné a été réalisé avec 9 combinaisons de vecteurs auxiliaires regroupées dans ce tableau. **C.** Des cellules Bosc23 ont été transfectées avec le nouveau vecteur pL-6, ainsi que le vecteur servant de contrôle pLKO-42 avec les différentes combinaisons de vecteurs auxiliaires regroupées dans le tableau en B afin de produire des particules virales. Des cellules HEK 293T et CH12 ont été infectées par les différents surnageants lentiviraux produits. 24h après infection, l'efficacité d'infection a été analysée par cytométrie en flux en regardant le pourcentage de cellules GFP+. Seulement les résultats de 2 combinaisons de vecteurs helper sont montrées.

nucléase Cas9 à l'ADN et la stabilisation de l'interaction Cas9-ARNg, et une cassette de sélection permettant la résistance à la puromycine. Afin de cloner la librairie d'ARNg dans le vecteur LentiCRISPR-v2, il est nécessaire de digérer de part et d'autre le « pool » des ARNg et le vecteur LentiCRISPR-v2 avec l'enzyme BsmBI (Sanjana et al., 2014).

La stratégie que nous avons mise en place afin de réaliser ce crible fonctionnel sur la CI est celle-ci (**Figure 15C**) :

Les cellules CH12 seront infectées avec la banque d'ARNg mGeCKO-v2 contenue dans le vecteur lentiviral LentiCRISPR-v2, avec une couverture permettant une représentation complète de la banque d'ARNg (300X), et à une faible multiplicité d'infection (MOI=0.3), assurant que les cellules soient infectées par une construction unique d'ARNg. Les cellules transduites seront ensuite sélectionnées avec la puromycine, afin d'éliminer les cellules qui n'expriment pas d'ARNg, et de permettre ainsi la génération de knockouts (KO). Puis, les cellules sélectionnées seront soumises à l'essai fonctionnel de la CI par stimulation pendant 72h, avec le cocktail de TGF β , d'IL-4 et de l'anticorps @CD40. Les cellules incapables d'effectuer la CI (restant IgM+) et les cellules qui ont subi des événements de CI efficaces (exprimant des IgA+) seront séparées et purifiées à l'aide de billes magnétiques couplées à des anticorps anti-IgM et anti-IgA. Enfin, les gènes dont la perte de fonction altère la CI seront identifiés par séquençage haut débit en comparant l'enrichissement des ARNg dans la population des IgM+ par rapport à la population d'IgA+.

1.2. D'une librairie lentivirale vers une librairie rétrovirale

1.2.1. La préparation du vecteur lentiCRISPR-V6

Afin d'améliorer le clonage de librairie d'ARNg, l'efficacité du criblage et la titration du virus produit, nous avons décidé d'ajouter des modifications au vecteur LentiCRISPR-v2. Tout d'abord, afin de discriminer efficacement les colonies ayant la bonne construction (librairie d'ARNg + LentiCRISPR-v2), l'élément de remplissage

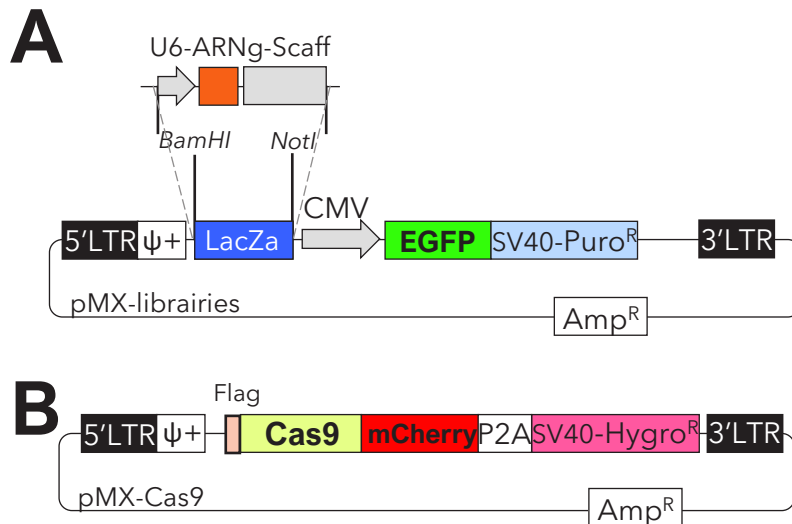


Figure 17: Nouvelle stratégie basée sur l'utilisation de vecteurs rétroviraux. Cette nouvelle stratégie a été conçue avec l'utilisation de deux vecteurs rétroviraux, d'une part, un vecteur permettant d'accueillir la collection d'ARNg (pMX-librairies) et d'autre part, un vecteur permettant la génération d'une lignée stable CH12 exprimant la nucléase Cas9 (pMX-Cas9). **A.** Le vecteur de destination pour les librairies d'ARNg est principalement composé d'un marqueur fluorescent EGFP sous le promoteur CMV et d'une cassette de sélection permettant la résistance à la puromycine. Les librairies d'ARNg sont sous clonées dans ce vecteur grâce aux sites de restriction des enzymes BamHI et NotI présentes de part et d'autre de la cassette LacZ dans le vecteur de destination et entourant la cassette U6-ARNg-Scaffold des librairies, après l'ajout des sites par PCR. **B.** Le vecteur permettant l'établissement d'une lignée cellulaire CH12 stable exprimant la nucléase Cas9 comprend une cassette Cas9-mCherry-P2A-SV40-gène de résistance à l'hygromycine. De plus, la nucléase Cas9 comporte une étiquette Flag en N-ter.

« staffer » entouré des sites de restriction de l'enzyme BsmB1 a été remplacé par une cassette LacZ. Ensuite, afin de perfectionner l'efficacité du criblage, l'élément « scaffold » a aussi été remplacé par un « scaffold » optimisé (« o-scaffold ») qui permet d'améliorer le repliement de l'ARNg et la stabilisation de l'interaction Cas9-ARNg (Cheng et al., 2015). Il est aussi indiqué comme pouvant réduire les effets hors-cible (Cheng et al., 2015). Enfin, de manière à évaluer plus facilement l'efficacité d'infection de la librairie d'ARNg par cytométrie en flux, l'insertion de la séquence codante pour la GFP a été réalisée en aval et fusionnée avec la séquence codante de la nucléase. Suite à l'ensemble de ces modifications, j'appellerai ce nouveau vecteur LentiCRISPR-v6 ou pL-6 (**Figure 16A**).

1.2.2. Le test de fonctionnalité du vecteur LentiCRISPR-v6

Dans le but de tester la fonctionnalité de ce nouveau vecteur, des cellules HEK 293T ont été transfectées avec le pL-6 exprimant un ARNg d'intérêt, et différentes combinaisons de plasmides auxiliaires (**Figure 16B**) pour produire des particules virales. Suite à une transfection efficace (40 à 65% de cellules GFP+ selon les combinaisons de vecteurs auxiliaires, données de transfection non montrées), les cellules CH12 et HEK 293T ont été infectées avec les différents surnageants lentiviraux. L'efficacité de la transduction a été évaluée en examinant le pourcentage de cellules GFP+ par cytométrie de flux. Les cellules HEK 293T ainsi que le vecteur pLKO-42 ont été utilisés comme contrôle pour l'infection. Les cellules HEK 293T ont correctement été infectées par ce nouveau vecteur pL-6 (20 à 64% de cellules GFP+ selon les combinaisons de vecteurs auxiliaires) contrairement à notre lignée de cellules B d'intérêt, les CH12 (0,3 à 1,5%). Seulement deux exemples de combinaisons de vecteurs auxiliaires sont montrés (**Figure 16C**). Nous avons conclu que notre nouveau vecteur pL-6 est fonctionnel mais qu'il est impossible d'infecter efficacement la lignée CH12 avec celui-ci.

1.2.3. Une nouvelle stratégie : l'utilisation de vecteurs rétroviraux

Pour réaliser notre criblage, nous nous sommes donc concentrés sur une nouvelle stratégie, et nous avons décidé d'utiliser le vecteur rétroviral pMXPIE, dont nous savons qu'il peut transduire les cellules B efficacement. Nous avons donc construit

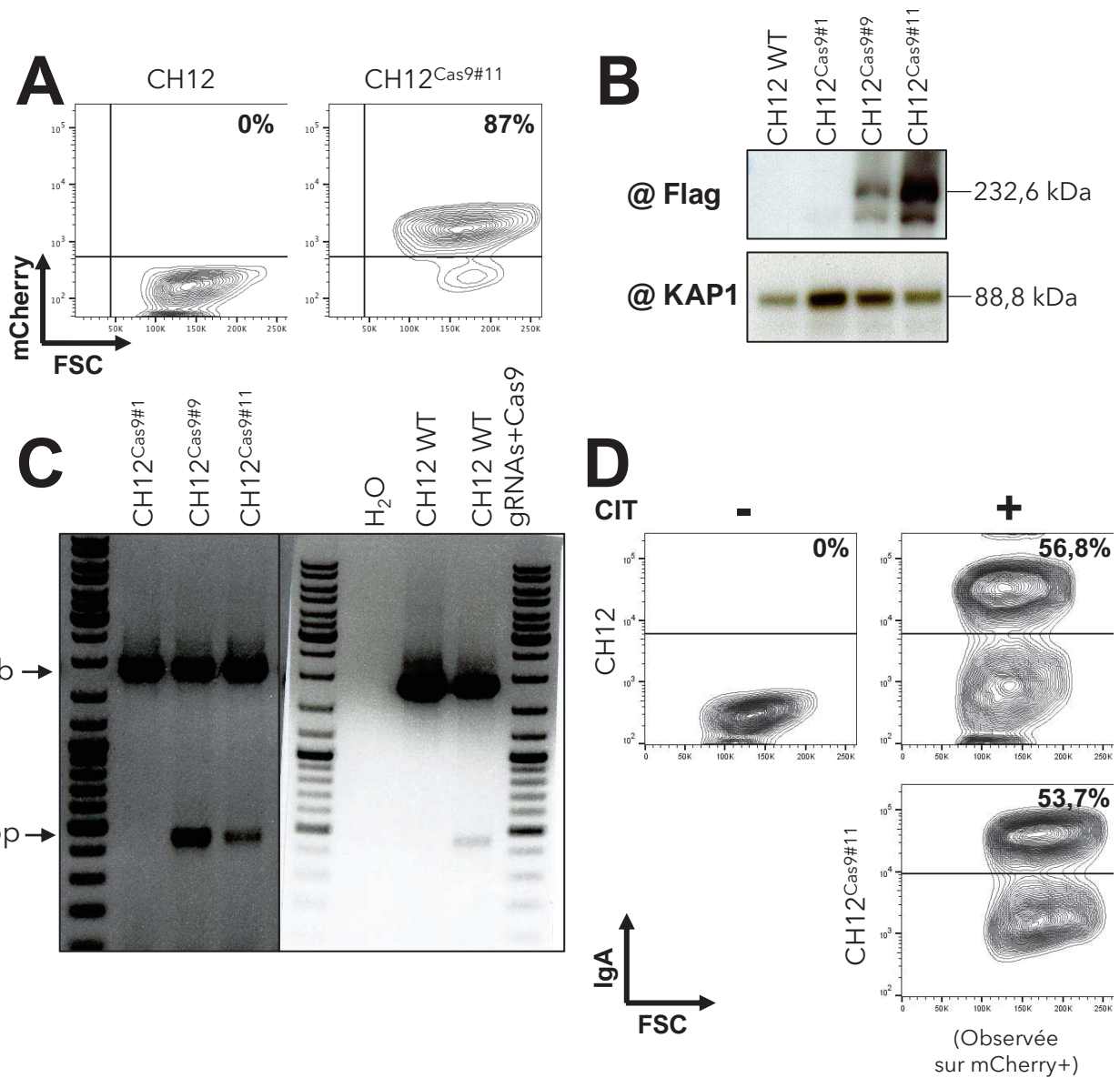


Figure 18: La lignée cellulaire CH12 exprimant la nucléase Cas9 (CH12^{Cas9}) est fonctionnelle. Suite à l'infection de cellules CH12 avec le vecteur rétroviral pMX-Cas9, ces cellules ont été sélectionnées 72h avec de l'hygromycine et les cellules mCherry positives ont été triées par cytométrie en flux afin de générer une lignée clonale CH12^{Cas9}. Les différents clones ont été caractérisés. **A.** L'expression stable de la mCherry qui est fusionnée à la Cas9 a été analysée par cytométrie en flux pour les cellules CH12^{Cas9} clone 11 (CH12^{Cas9}#11). **B.** L'expression de la nucléase Cas9 a été analysée par Western Blot à l'aide des anticorps anti-Flag qui est en N-ter de la Cas9 et anti-KAP1 servant de contrôle de chargement, à partir d'extraits provenant des différentes lignées clonales potentiellement CH12^{Cas9}. **C.** Les différentes lignées clonales exprimant la nucléase Cas9 ont été transfectées par 2 ARNg ciblant la région de l'enhancer E γ 1, puis sélectionnées avec de l'hygromycine pendant 3 jours. Différentes conditions contrôles ont été ajoutées: contrôles n'exprimant pas la nucléase Cas9: la lignée clonale CH12^{Cas9}#1 et les CH12 wild-type (WT); et des CH12 transfectées avec le vecteur d'origine ayant permis la génération des KO E γ 1 (ARNg+Cas9). L'ADNg de ces différentes conditions a été extrait et une PCR a été réalisée à l'aide d'amorces entourant la région E γ 1 ciblée par les ARNg donnant une taille de 1,8kb pour un allèle WT et 450pb pour un allèle KO. **D.** La lignée clonale CH12^{Cas9} choisie (CH12^{Cas9}#11) et les CH12 WT ont été stimulées pendant 72h avec du TGF β , IL-4 et anti-CD40, (CIT). L'expression d'IgA à la surface des cellules a été analysée par cytométrie en flux.

deux vecteurs rétroviraux, le premier afin de sous-cloner la banque d'ARNg (pMX-mGeCKO-v2) (**Figure 17A**) et un second vecteur rétroviral permettant l'expression de la nucléase Cas9 (pMX-Cas9) (**Figure 17B**), pour établir une lignée cellulaire CH12 stable exprimant Cas9-mCherry-P2A-Hygromycine.

1.2.3.1. La génération et la caractérisation d'une lignée cellulaire CH12 stable exprimant la nucléase Cas9

La lignée CH12 exprimant de façon stable la nucléase Cas9 (CH12^{Cas9}) a donc été générée par transduction rétrovirale. 72h après sélection avec de l'hygromycine, les cellules mCherry positives ont été triées par cytométrie en flux afin de générer une lignée clonale CH12^{Cas9}. Suite à l'obtention de plusieurs clones, ceux-ci ont été caractérisés. Tout d'abord, l'expression stable de la mCherry qui est fusionnée à la nucléase Cas9 a été analysée par cytométrie en flux. Seulement le clone CH12^{Cas9#11}, qui a été choisi finalement comme la lignée stable CH12^{Cas9} est montré ici avec 87% de cellules mCherry positives (**Figure 18A**). Puis, l'expression de la nucléase Cas9 a été analysée par western blot (WB) à l'aide des anticorps anti-Flag, l'étiquette Flag se trouvant en N-ter de la Cas9 et anti-KAP1, servant de contrôle de chargement. Ce WB a été réalisé à partir d'extraits de CH12 WT et de trois clones CH12^{Cas9} exprimant la mCherry par cytométrie en flux : les clones 1, 9 et 11. Le résultat du WB montre que seulement les clones CH12^{Cas9#9} et CH12^{Cas9#11} expriment la nucléase Cas9 mais à des niveaux différents (**Figure 18B**). J'ai ensuite évalué la capacité de la nucléase à induire des cassures double brin par la transfection de deux ARNg ciblant la région de l'enhancer Ey1, connus dans le laboratoire pour être efficaces. Les cellules transfectées sont ensuite sélectionnées avec de l'hygromycine pendant 3 jours et l'ADN génomique (ADNg) a été extrait. Des conditions négatives telles que la lignée clonale CH12^{Cas9#1} n'exprimant pas la Cas9 par WB et les cellules CH12 WT ont été ajoutées au même titre qu'une condition positive telle que les cellules CH12 WT transfectées avec le vecteur d'origine comportant la nucléase Cas9 et les deux ARNg ciblant l'Ey1. L'ADNg des différentes conditions a été amplifié par PCR avec des amorces entourant la région de l'Ey1 ciblée par les deux ARNg donnant une taille de 1,8kb pour un allèle WT et 450pb pour un allèle KO. La présence d'une bande à 450 pb sur gel d'agarose révèle une activité fonctionnelle de la nucléase Cas9 pour

Figure 19: Un crible à petite échelle permet d'identifier les gènes requis lors la CI. Légende page suivante.

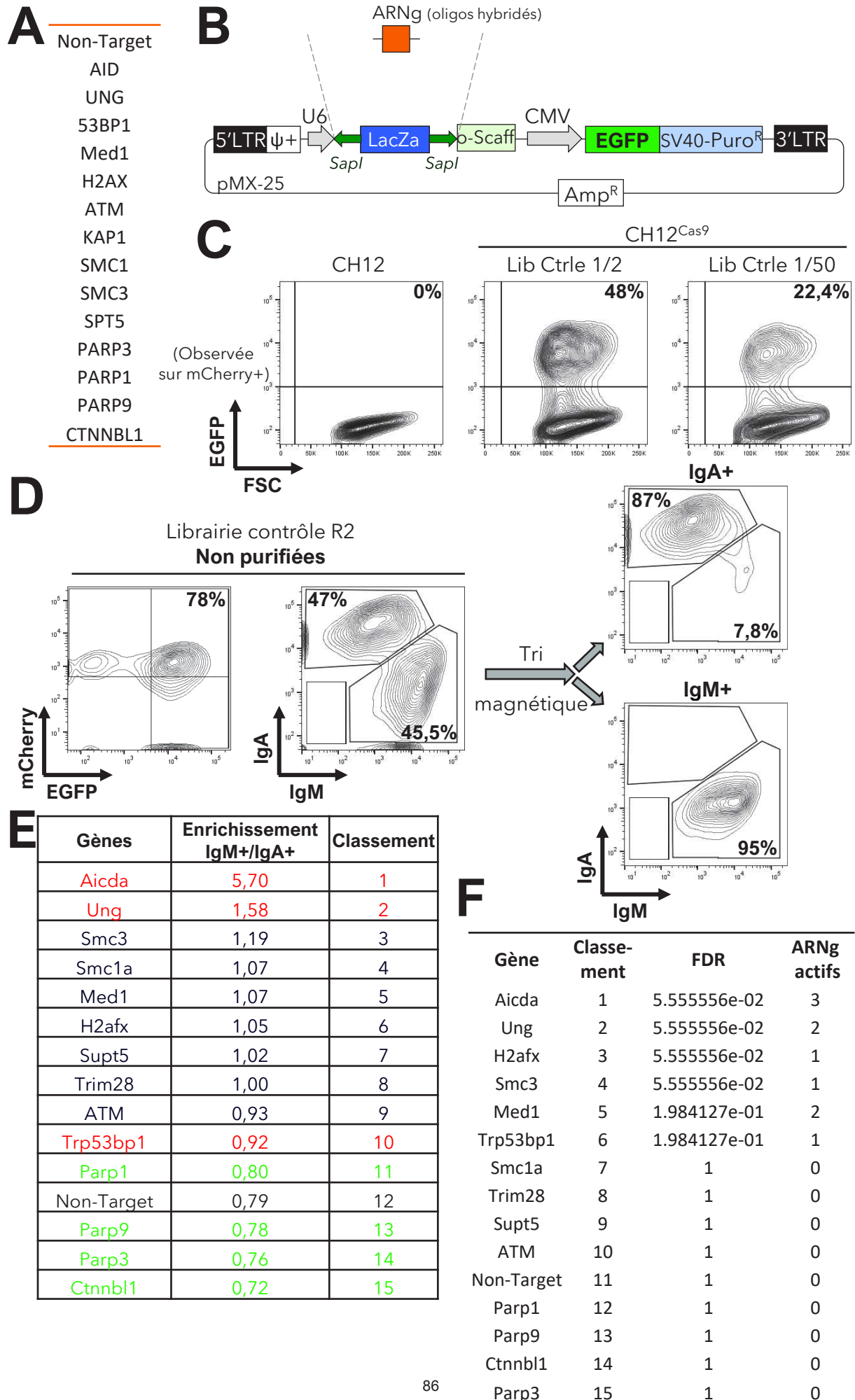


Figure 19: Un crible à petite échelle permet d'identifier les gènes requis lors la CI. La preuve de principe de notre stratégie a été réalisée avec un crible à petite échelle en utilisant des ARNg ciblant **A.** des gènes absolument requis pour la CI (AID, UNG et 53BP1), des gènes dont l'inactivation entraîne une diminution (Med1, H2AX, ATM et KAP1) ou une augmentation de l'efficacité de CI (PARP3), des gènes non essentiels pour la CI (PARP1, PARP9 et CTNBL1), des gènes léthaux (SMC1/3 et SPT5) ainsi que des ARNg contrôles (Non-Target). **B.** Les amorces hybridées des différents ARNg ont été sous clonées dans un vecteur rétroviral par digestion/ligation. Ce vecteur comprend une cassette U6-LacZ-o-Scaffold, un marqueur fluorescent EGFP sous le promoteur CMV et une cassette de sélection permettant la résistance à la puromycine. **C.** Des cellules BOSC23 ont été transfectées avec la librairie contrôle de 48 ARNg afin de produire des particules virales. Les cellules CH12^{Cas9} ont été infectées avec différentes dilutions du surnageant rétroviral afin d'effectuer la titration du virus. L'efficacité d'infection a été analysée par cytométrie en flux en regardant le pourcentage de cellules GFP+ dans la population mCherry+. **D.** L'analyse de l'expression d'IgA et d'IgM à la surface des cellules permettant d'évaluer l'efficacité de la CI et la purification par tri magnétique des cellules infectées et stimulées pendant 72h a été réalisée par cytométrie en flux. **E.** Analyse de l'enrichissement des ARNg dans la population IgM+ vs. la population IgA+ et du classement des différents gènes de la librairie contrôle. Les gènes en rouge représentent ceux qui sont absolument requis pour la CI, en bleu, leur inactivation entraîne une diminution de la CI et enfin en vert, ce sont les gènes dispensables lors de la CI. **F.** Analyse du crible CRISPR/Cas9 de la librairie contrôle avec HitSelect. FDR: False Discovery Rate.

les clones 9 et 11 (**Figure 18C**). Ce résultat montre que la nucléase Cas9 exprimée dans les lignées CH12^{Cas9#9} et CH12^{Cas9#11} est bien fonctionnelle. Enfin, j'ai évalué l'efficacité de la CI d'IgM vers IgA dans les deux lignées clonales CH12^{Cas9#9} et CH12^{Cas9#11} et les CH12 WT en les stimulant avec du TGFβ, de l'IL-4 et de l'αCD40. Seuls le clone CH12^{Cas9#11} et les cellules CH12 WT sont montrés ici et révèlent une efficacité de CI égale (**Figure 18D**). Ces résultats montrent que la lignée clonale CH12^{Cas9#11} exprime de façon stable la nucléase Cas9 qui a été montrée fonctionnelle, et qu'elle réalise la CI de manière efficace. Cette lignée clonale CH12^{Cas9#11} sera appelée par la suite CH12^{Cas9}.

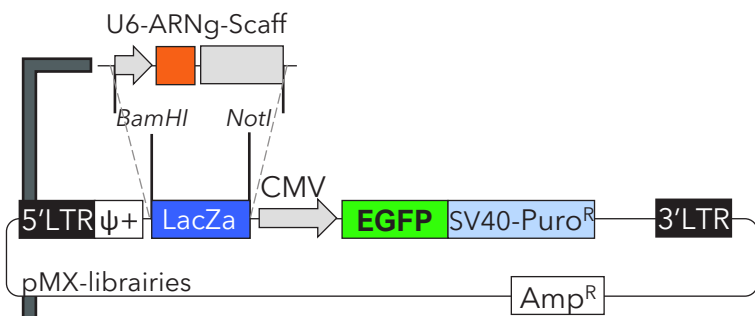
1.2.3.2. Preuve de principe : criblage fonctionnel sur la CI à petite échelle

Dans le but d'apporter la preuve de principe de notre criblage fonctionnel sur la CI, j'ai réalisé un criblage à petite échelle avec des ARNg ciblant des gènes absolument requis pour la CI (AID, UNG et 53BP1), des gènes dont l'inactivation entraîne une diminution (Med1, H2AX, ATM et KAP1) ou une augmentation de l'efficacité de CI (PARP3), des gènes non essentiels pour la CI (PARP1, PARP9 et CTNBL1), des gènes léthaux (SMC1/3 et SPT5) ainsi que des ARNg contrôles (**Figure 19A**). Les séquences des ARNg proviennent de manière générale des ARNg présents dans la librairie mGeCKO-v2 (Sanjana et al., 2014) ou la librairie Sabatini (Wang et al., 2014a), exceptés les guides ciblant AID qui sont des guides que nous avons conçus pour un précédent projet. Pour cela, j'ai cloné les ARNg ciblant ces gènes (3 ARNg/gène et 6 ARNg « Non-Target ») dans un nouveau vecteur rétroviral (pMX-25) (**Figure 19B**) et réalisé une librairie équimolaire comprenant les 48 ARNg. Des cellules BOSC23 ont été transfectées avec la librairie des 48 ARNg afin de produire des particules virales. Les cellules CH12^{Cas9} ont ensuite été infectées avec le surnageant contenant les particules virales dilué à plusieurs concentrations afin de titrer le virus. L'efficacité d'infection a ensuite été évaluée par cytométrie en flux. Afin de respecter une MOI de 0.3, selon la loi de Poisson, l'efficacité d'infection doit être équivalente à environ 26%. Dans ce cas, la dilution la plus proche et respectant le plus une MOI de 0.3 est celle au 1/50^{ème} (**Figure 19C**). Les cellules CH12^{Cas9} infectées présentant une MOI équivalente à 0.3 ont ensuite été mises sous sélection pendant

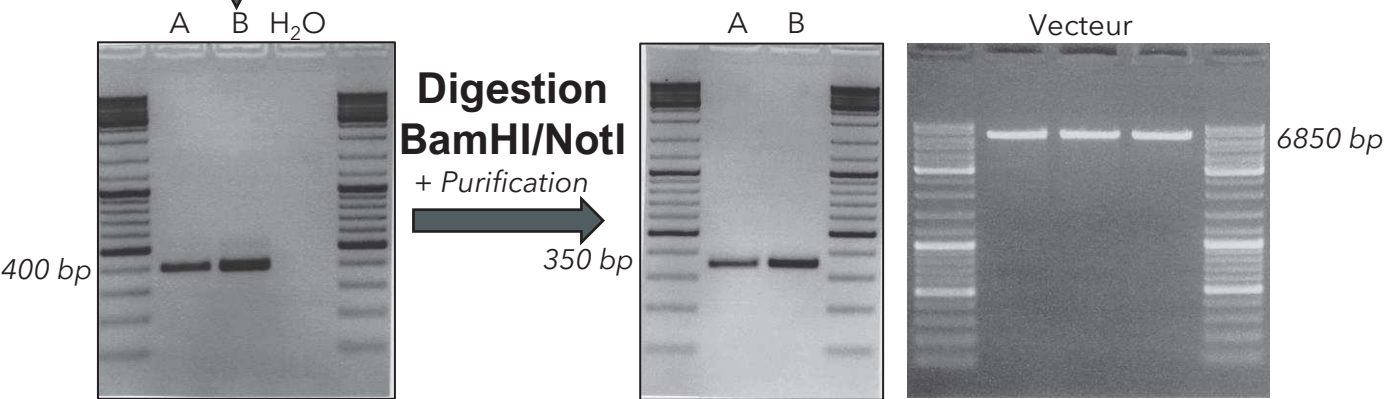
10 jours avec de la puromycine et de l'hygromycine, puis induites pour la CI. Après 72h de stimulation, les cellules ont été soumises à un tri magnétique afin d'isoler les populations IgM⁺ et IgA⁺. L'efficacité de la CI et du tri magnétique des cellules infectées ont été analysées par cytométrie en flux en évaluant l'expression d'IgA et d'IgM à la surface des cellules. Seulement un réplicat est montré. Les cellules CH12^{Cas9} transduites ont montré une efficacité robuste de CI (42 à 48% d'IgA⁺) et la purification des populations IgA⁺ et IgM⁺ a été efficace avec une pureté de 87% et de 87 à 95%, respectivement (**Figure 19D**). Puis, l'ADNg des populations d'IgA⁺ et d'IgM⁺ triées ainsi que les cellules infectées et sélectionnées mais non induites à la CI a été extrait et les séquences des ARNg ont été amplifiées par PCR. Cette amplification permet également d'ajouter les adaptateurs Illumina et les codes-barres, nécessaires au séquençage à haut débit. Les produits PCR ont été purifiés et analysés par séquençage haut débit. Ces expériences ont été réalisées en trois réplicats et la bibliothèque de plasmides a également été amplifiée par PCR afin d'être soumise au séquençage haut-débit.

L'extraction des données brutes a été réalisée avec MAGeCK (Li et al., 2014), puis celles-ci ont été normalisées selon le protocole du laboratoire de Feng Zhang (Shalem et al., 2014). Nous avons tout d'abord généré un premier classement des gènes en calculant l'enrichissement des ARNg des populations IgM⁺ vs. IgA⁺ (**Figure 19E**). Ensuite, nous avons utilisé le logiciel HitSelect (Diaz et al., 2015) qui prend en compte le nombre d'ARNg actifs et de réplicats biologiques pour effectuer un classement des gènes (**Figure 19F**). Pour précision, les guides actifs sont des guides qui ont un effet sur le phénotype observé.

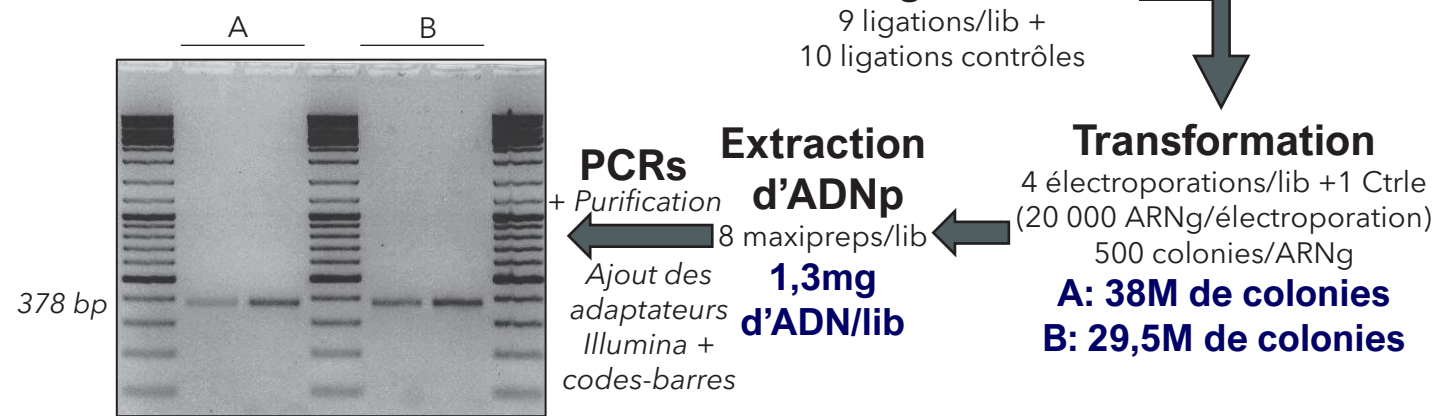
L'analyse du crible de la librairie contrôle a montré d'une part une répartition et une distribution égale des ARNg dans la banque de plasmides et dans les cellules transduites non stimulées ; d'autre part, un enrichissement important des ARNg correspondant aux gènes requis lors de la CI dans la population IgM⁺ vs IgA⁺ ; et enfin des gènes qui se classent comme prévu selon leur importance et leur implication ou non dans la CI (**Figure 19E**). Cependant, pour ces deux derniers points, nous avons une exception avec 53BP1. 53BP1 qui est nécessaire pour la CI n'est pas dans le haut de la liste avec AID et UNG. Le classement de 53BP1 est dû



4 PCR/lib
Ajout des sites
BamHI/NotI



Ligation
9 ligations/lib +
10 ligations contrôles



**Séquençage
haut-débit
+ Analyse**

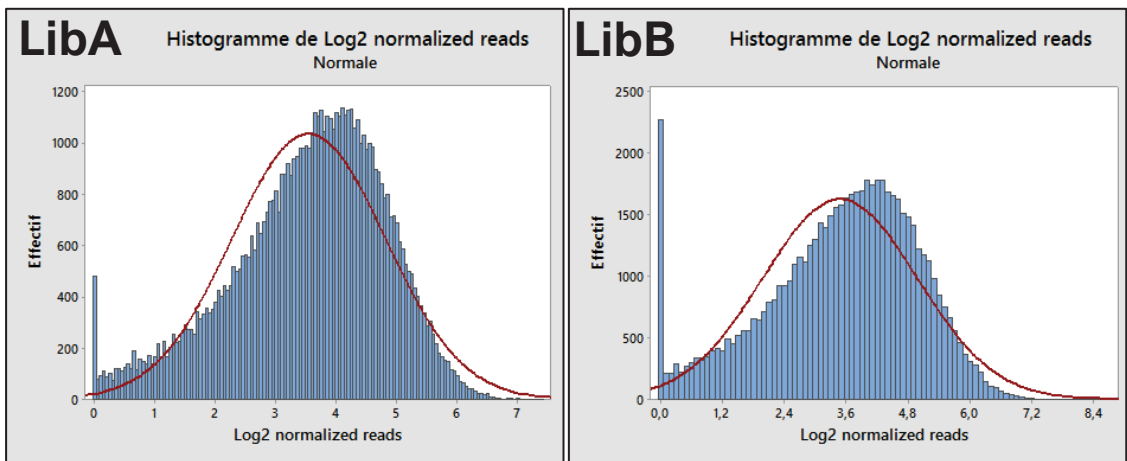


Figure 20: Le clonage de la librairie Rétro-mGeCKO-v2. Légende page suivante.

Figure 20: Le clonage de la librairie Rétro-mGeCKO-v2. La cassette U6-ARNg-Scaffold des librairies A et B mGeCKO-v2 a été amplifiée afin d'ajouter les sites de restriction BamHI/NotI de part et d'autre de la cassette. Puis, le vecteur de destination ainsi que les produits PCR de la cassette U6-ARNg-Scaff des deux sous-librairies ont été digérés, purifiés et enfin ligués. Les ligations ont été regroupées et précipitées afin d'être transformées dans des bactéries Stbl4. L'ADN plasmidique (ADNp) des deux librairies a été extrait et amplifié par PCR afin d'ajouter les adaptateurs Illumina et les codes-barres de part et d'autres des séquences des ARNg. Les produits PCR ont été purifiés et soumis au séquençage à haut débit. L'analyse de la qualité des deux sous-librairie A et B est représentée par un histogramme représentant le nombre d'ARNg en fonction du Log2 du nombre de reads normalisés.

au fait qu'il ne présente qu'un seul guide actif (**Figure 19F**). Ceci qui montre que les ARNg de la librairie mGeCKO-v2 ne sont pas forcément actifs comme c'est le cas pour plusieurs gènes. Pour conclure, ce criblage à petite échelle nous a permis d'identifier efficacement les gènes connus pour être impliqués dans les mécanismes de la CI, et valide notre nouvelle stratégie. Cependant, nous avons pu voir que des ARNg peuvent être inactifs. Cela pourrait entraîner des difficultés dans l'interprétation des résultats lorsque le criblage sera réalisé à grande échelle puisque les gènes avec des guides inactifs pourraient constituer de faux-négatifs. De plus, ce criblage a permis de mettre au point toutes les étapes de notre stratégie parmi lesquelles on retrouve notamment la détermination de la concentration en puromycine et en hygromycine pour la lignée CH12^{Cas9}, la détermination du nombre de cycle PCR pour l'ajout des adaptateurs Illumina et des codes-barres afin d'éviter les biais ou encore la mise au point du tri magnétique.

1.3. Le crible CRISPR/Cas9 pour identifier les gènes impliqués dans la CI avec la librairie mGeCKO-v2

1.3.1. La génération de la librairie Rétro-mGeCKO-v2

Afin de réaliser le crible fonctionnel, nous avons donc sous-cloné les collections d'ARNg venant de la librairie mGeCKO-v2 (U6-ARNg-scaffold) dans un vecteur rétroviral contenant la séquence codante pour le promoteur CMV, la GFP et le gène de résistance à la puromycine (**Figure 20**). Ce sous-clonage a été réalisé en respectant la représentation et la distribution de la collection d'ARNg, autrement dit une couverture de la bibliothèque d'ARNg de 300X comme décrit dans le protocole de l'équipe de Feng Zhang (Joung et al., 2016). La cassette U6-ARNg-scaffold a été amplifiée par PCR (4 PCRs de 100ng/chaque librairie) pour ajouter deux sites de restriction, BamHI et NotI en 5' et 3' de l'amplicon. Ensuite, ces produits PCR (librairie A et B) et le vecteur de destination ont été digérés, purifiés et reliqués ensemble (9 ligations/librairie et 10 ligations contrôles sans insert). 4 électroporations pour chaque banque ont été effectuées. Le calcul de l'efficacité de transformation a montré une bonne efficacité de transformation pour la librairie A (38 M de colonies contre 33,7M attendues) et une efficacité relativement correcte de la librairie B (29,5

M de colonies contre 31,4M attendues). Enfin, l'ADNp de chaque banque rétrovirale a été extrait et purifié (8 maxipreps/ librairie). Afin de pouvoir valider le crible fonctionnel, nous avons ajouté 6 ARNg actifs ciblant AID dans chacune des sous-librairies. Dans le but de valider la qualité des librairies rétrovirales, l'ADN plasmidique de celles-ci a été amplifié par PCR afin d'ajouter les adaptateurs Illumina et les barres-codes. Les produits PCR ont été purifiés et analysés par séquençage à haut débit (**Figure 20**). Afin de respecter la couverture de 300X des librairies, la prise en compte de l'obtention d'au moins 20 millions de « reads » par librairie a été considérée. Nous avons obtenu un total de 134 millions de « reads » pour la librairie A et 118 millions pour la librairie B avec un score de qualité Q30 de 91,56%. Le score Q30 est utilisé pour mesurer la précision de l'identité des nucléotides à partir d'un « run » de séquençage. Ces résultats montrent que la représentation et la distribution des ARNg ont bien été respectées durant le séquençage haut-débit et que les données sont exploitables.

L'extraction du nombre de « reads » de chaque ARNg a été réalisée en utilisant MAGeCK (Li et al., 2014). Afin d'évaluer la qualité des librairies, trois paramètres permettent de déterminer la qualité d'une librairie sous-clonée : la représentation globale, l'uniformité de la librairie et le pourcentage d'ARNg qui correspondent exactement (Joung et al., 2016). La représentation globale est le pourcentage d'ARNg sans « reads » qui doit être inférieur à 0,5%. La librairie A est à 0,696% tandis que la librairie B est à 3,46%. L'uniformité de la librairie correspond à la différence entre le 90^{ème} percentile et le 10^{ème} percentile qui doit être inférieure à 10X. La librairie A est à 13,33X et la librairie B est à 25,85X. Le pourcentage d'ARNg correspondant exactement doit être autour de 70%. La librairie A est à 69,6% et la librairie B est à 67,7%. De plus, la distribution des ARNg suit relativement la courbe de Gauss (**Figure 20**).

Ces résultats montrent que la librairie A est de très bonne qualité contrairement à la librairie B qui est de moins bonne qualité et affiche un nombre plus élevé de guides absents mais, elle reste tout de même relativement correcte. Enfin, l'ensemble des deux librairies montre une couverture de 98 % du génome murin.

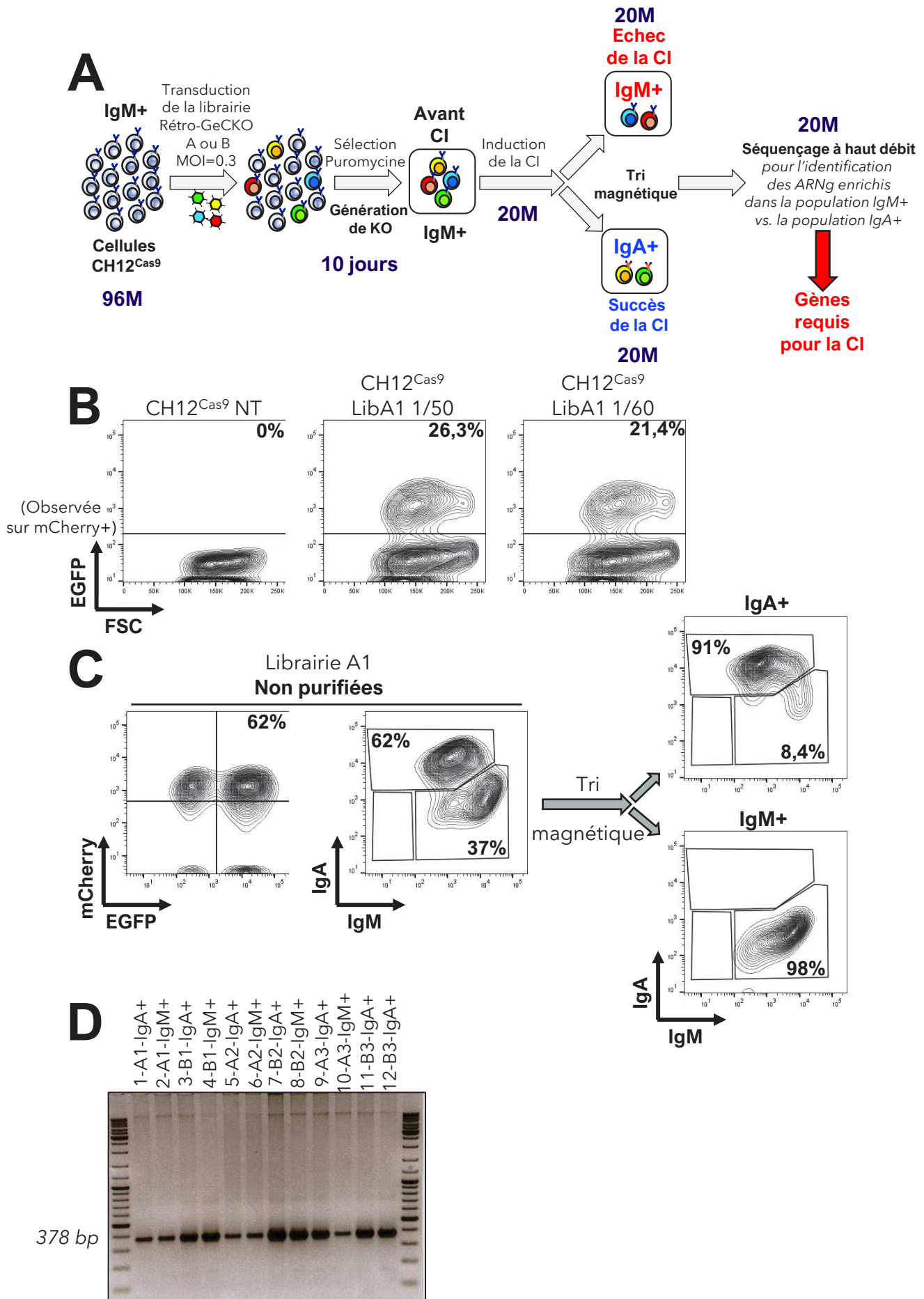


Figure 21: Criblage de la librairie Rétro-mGeCKO-v2. Légende page suivante.

Figure 21: Criblage de la librairie Rétro-mGeCKO-v2. **A.** Vue d'ensemble de la stratégie et des différentes étapes du crible avec les sous-librairies A et B. Le nombre de cellules indiqué en violet pour les différentes étapes correspond au respect de la couverture de 300X de la librairie. **B.** Les cellules CH12^{Cas9} ont été infectées avec différentes dilutions du surnageant rétroviral afin d'effectuer la titration du virus. L'efficacité d'infection a été analysée par cytométrie en flux en regardant le pourcentage de cellules GFP+ dans la population mCherry+. **C.** L'analyse de l'expression d'IgA et d'IgM à la surface des cellules permettant d'évaluer l'efficacité de la CI et la purification par tri magnétique des cellules infectées et stimulées pendant 72h a été réalisée par cytométrie en flux. **D.** L'ADNg des populations IgA+ et IgM+ a été extrait et amplifié par PCR afin d'ajouter les adaptateurs Illumina et les codes-barres de part et d'autre des séquences des ARNg. Les produits PCR ont été purifiés et soumis au séquençage à haut débit.

1.3.2. Le crible fonctionnel sur la CI avec la librairie Rétro-mGeCKO-v2

1.3.2.1. La réalisation du crible CRISPR/Cas9 à l'aide de la librairie Rétro-mGeCKO-v2

Le criblage fonctionnel sur la CI a été réalisé en 4 réplicats biologiques avec les librairies A et B Rétro-mGeCKO-v2 en utilisant une couverture de 300 fois pour chaque étape (**Figure 21A**). Des cellules BOSC23 ont été transfectées avec les librairies A et B afin de produire des particules virales. 4 surnageants rétroviraux différents pour chaque librairie ont été générés. Des cellules CH12^{Cas9} ont ensuite été infectées avec les surnageants rétroviraux dilués à plusieurs concentrations afin de les titrer. L'efficacité d'infection a ensuite été évaluée par cytométrie en flux en regardant les cellules GFP+ sur les cellules mCherry+. Seuls les résultats du réplicat A1 sont montrés pour les différentes étapes du criblage. Dans le cas du réplicat A1, la dilution au 1/50^{ème} a été choisie avec un pourcentage de 26,3 correspondant à une MOI de 0.3 (**Figure 21B**). J'ai donc ensuite transduit 96 millions de cellules CH12^{Cas9} avec le surnageant rétroviral de la librairie A ou B à une MOI de 0,3 pour chaque réplicat. Le nombre de cellules à infecter a été estimé en prenant en compte que nous voulions environ 20 millions de cellules infectées à une MOI de 0.3 et le nombre de cellules vivantes suite à une infection. J'ai pu évaluer la couverture de l'étape d'infection en analysant l'efficacité de celle-ci par cytométrie en flux et en comptant le nombre de cellules vivantes. Pour chaque réplicat, la couverture se situait approximativement entre 300 et 600X, avec une efficacité d'infection comprise entre 20% et 27%. Les cellules CH12^{Cas9} transduites ont été sélectionnées avec de la puromycine pendant 10 jours. Durant ces 10 jours, des passages de 40 millions de cellules transduites avec de la puromycine et de l'hygromycine ont été réalisés pour chaque librairie et réplicat. La commutation isotypique a été induite avec 20 millions de cellules pour chaque réplicat. 72h après, les cellules B incapables d'effectuer la CI (restant IgM+) ou qui ont subi des événements de CI efficaces (exprimant IgA+) ont été triées à l'aide de billes magnétiques couplées à des anticorps anti-IgM ou anti-IgA. Nous avons évalué l'efficacité de la CI et du tri magnétique des cellules en examinant l'expression d'IgA et d'IgM à la surface des

cellules par cytométrie en flux. Nous avons obtenu une efficacité de CI robuste de 53 à 63% ainsi qu'une purification efficace de la population IgA+ (86 à 93%) et de la population IgM+ (94 à 97%) (**Figure 21C**). L'ADN génomique de 20 millions de cellules des différentes populations cellulaires a été extrait et soumis à des PCR (20 PCR de 5µg/librairie/réplikat) pour amplifier les séquences des ARNg et ajouter des adaptateurs Illumina et des codes-barres. Ces PCRs ont été optimisées au préalable afin de déterminer le nombre de cycles nécessaires pour chaque échantillon/paire d'oligos pour éviter les biais. Les produits PCR ont été purifiés (**Figure 21D**), puis soumis au séquençage à haut débit pour identifier les ARNg enrichis dans la population IgM+ vs. la population IgA+. Afin de respecter la couverture de 300X des librairies, la prise en compte de l'obtention d'au moins 20 millions de « reads » par librairie a été considérée. Nous avons obtenu entre 20 et 85 millions de « reads » pour les différentes librairies et réplikat avec un score de qualité Q30 à environ 95,8%.

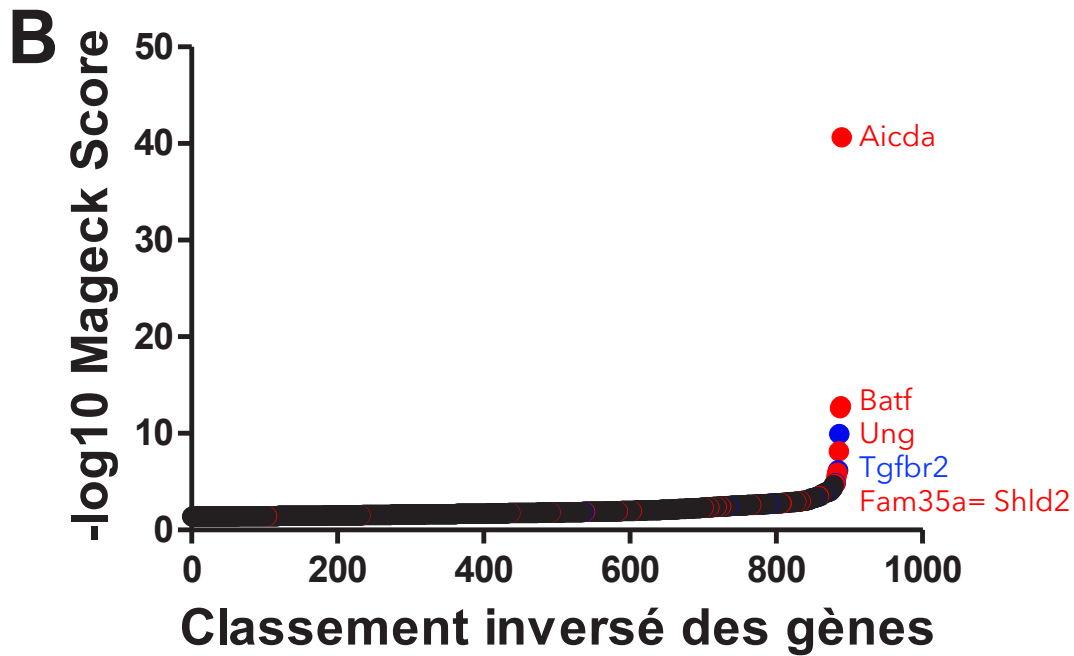
L'extraction des données brutes a été réalisée avec MAGeCK (Li et al., 2014) puis celles-ci ont été normalisées selon le protocole du laboratoire de Feng Zhang (Shalem et al., 2014). Pour analyser l'enrichissement des ARNg des populations IgM+ et IgA+, nous avons utilisé deux logiciels : HitSelect (Diaz et al., 2015) et MAGeCK (Li et al., 2014).

1.3.2.2. L'analyse du crible fonctionnel sur la CI réalisé avec la librairie Rétro-mGeCKO-v2.

Le classement des gènes obtenu avec l'analyse MAGeCK est relativement identique à celle obtenue avec HitSelect mais l'analyse MAGeCK s'est montrée plus stringente. C'est pourquoi, dans un premier temps les résultats obtenus avec l'analyse MAGeCK seront montrés (**Figure 22**). L'analyse du crible a permis l'identification de 640 gènes candidats potentiellement impliqués dans les mécanismes moléculaires de la CI (**Figure 22A-B**). Comme attendu, nous retrouvons des gènes enrichis dans la population IgM+ qui sont déjà connus pour être requis pour la CI tels qu'Aicda et Ung, des facteurs de transcription conduisant l'expression d'AID comme Batf et

A

Gène	Classement	ARNg actifs	p-Value
Aicda	1	14	2,08E-07
Batf	2	4	2,08E-07
Ung	3	5	2,08E-07
Tgfb2	4	4	2,08E-07
Mlh1	5	4	2,08E-07
Il2rg	6	4	3,12E-06
Pten	7	2	6,04E-06
Foxo1	8	3	1,81E-05
Junb	9	3	6,76E-05
Smc1a	10	3	8,26E-05
Irf4	11	3	8,72E-05



C

Gène	Classement	ARNg actifs	p-Value
Aicda	1	14	2,08E-07
Batf	2	4	2,08E-07
Ung	3	5	2,08E-07
Trp53bp1	23584	0	0,99164
Rpa1	10531	0	0,44766
Supt5	3312	0	0,14072

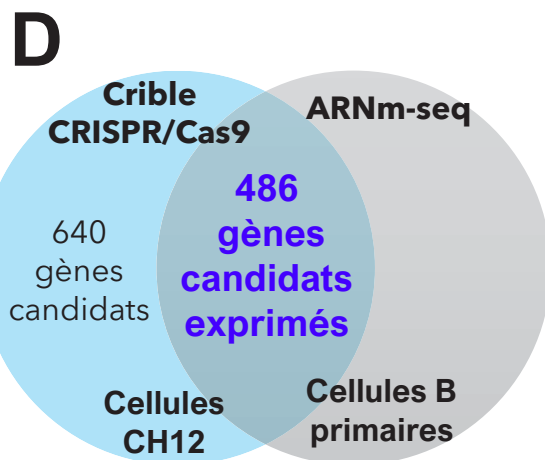


Figure 22: Le crible CRISPR/Cas9 avec la librairie Rétro-mGeCKO-v2 permet d'identifier des gènes connus pour être impliqués dans la CI et de nouveaux gènes candidats. Légende page suivante.

Figure 22: Le crible CRISPR/Cas9 avec la librairie Rétro-mGeCKO-v2 permet d'identifier des gènes connus pour être impliqués dans la CI et de nouveaux gènes candidats. **A.** Liste des 11 gènes les plus fortement enrichis issue de l'analyse MAGeCK. **B.** Graphique réalisé avec les données générées par l'analyse MAGeCK avec une p-Value inférieure à 0,05 et au moins 2 ARNg actifs, représentant le $-\log_{10}$ du Score Mageck en fonction du classement inversé des gènes. **C.** Liste de gènes comprenant les 3 gènes les plus fortement enrichis dans l'analyse MAGeCK (en marron) et 3 autres gènes connus pour être impliqués dans la CI mais qui ne sont pas enrichis dus à des ARNg inactifs (en vert). **D.** Diagramme de Venn représentant en bleu, les résultats du crible CRISPR/Cas9 avec la librairie Rétro-mGeCKO-v2 issus de l'analyse MAGeCK avec une p-Value inférieure à 0,05 et au moins 2 ARNg actifs et en gris, les résultats d'un ARNm-Seq réalisé dans des cellules B primaires. Les gènes en rouge représentent les gènes requis pour une CI efficace. Les gènes en bleu représentent les gènes impliqués dans la CI avec un rôle dans les voies de signalisation du TGF β , IL-4, CD40, BCR, NF- κ B. Les gènes en noirs représentent des gènes n'ayant pas encore de fonction connue dans la CI.

Pten, des facteurs de réparation de l'ADN tels que Ku80, Msh2, Msh6 et les gènes codants pour les sous-unités des récepteurs TGF β , IL-4 et CD40L, comme Tgfbr1/2, Smad2-4, CD40, Il2rg et Il2ra (**Figure 22A**). Également, nous avons aussi identifié un grand nombre de gènes candidats n'ayant pas encore de fonction connue dans les mécanismes de CI (**Figure 22B** ; gènes en noir). Ces résultats montrent que le crible fonctionnel sur la CI avec la librairie mGeCKO-v2 est fonctionnel et permet d'identifier de nouveaux cofacteurs de la CI dont Fam35a, plus connus sous le nom de Shieldin2 (Ghezraoui et al., 2018; Gupta et al., 2018; Noordermeer et al., 2018) et HMCES (Halabelian et al., 2019; Mohni et al., 2019; Shukla et al., 2020), inconnus lors de l'obtention des résultats de ce crible.

La présence d'ARNg inactifs issus de la librairie GeCKO-v2 observée lors du crible à petite échelle a été confirmée dans l'analyse de ce crible (**Figure 22C**). En effet, comme montré, des gènes nécessaires pour une CI efficace, tels que Trp53bp1, Rpa1 et Supt5 ne sont pas enrichis dans la population IgM⁺ et présentent 0 ARNg actif (**Figure 22C**). Ceci indique que l'inactivité des ARNg de la librairie mGeCKO-v2 entraîne la présence de faux-négatifs.

De plus, le recoupement de cette liste de gènes potentiellement impliqués dans les événements de CI dans les cellules CH12 avec une liste de gènes exprimés dans des cellules B primaires issue d'un ARNm-Seq effectué dans le laboratoire montre que seulement 486 gènes ont été trouvés exprimés dans les cellules B primaires (**Figure 22D**). Ceci indique que 26% des gènes enrichis dans la population IgM⁺ peuvent constituer des faux positifs. Même si des différences d'expression génique peuvent exister entre une lignée cellulaire ayant subi la transformation maligne, telle que la lignée CH12, dérivée d'un lymphome B murin, et des cellules « normales », la librairie d'ARNg mGeCKO-v2 semble générer un certain nombre de faux-positifs.

Dans l'ensemble, nous avons montré que notre criblage fonctionnel sur la CI à l'échelle du génome avec la librairie mGeCKO-v2 permet l'identification de gènes déjà connus pour être nécessaires pour une CI efficace et de nouveaux gènes candidats. Ceci indique que le crible est efficace. Toutefois, celui-ci pourrait être amélioré afin d'éviter les faux-négatifs et les faux-positifs.

1.4. Un crible CRISPR/Cas9 d'une plus grande précision pour identifier les cofacteurs de la CI avec la librairie Rétro-Brie

Afin d'améliorer le crible fonctionnel sur la commutation isotypique, nous avons décidé d'utiliser une nouvelle librairie qui couvre aussi l'ensemble du génome murin, la librairie lentivirale Brie (Doench et al., 2016). Celle-ci a été conçue de manière à maximiser l'activité des ARNg et minimiser les effets hors-cible. Doench et al., ont montré que des caractéristiques de la séquence de l'ARNg et de la séquence PAM permettent d'améliorer l'activité des ARNg (Doench et al., 2016; Doench et al., 2014). Le choix des ARNg a été réalisé en ciblant les transcrits codants d'un gène cible, annotés par la base de données appelée GENCODE qui identifie et classifie les caractéristiques des gènes des génomes humain et murin. Ils ont choisi un transcrit codant par gène en se basant sur 3 paramètres qui sont d'une part, les transcrits annotés avec des informations structurelles ou des résidus importants pour la fonction de la protéine ; d'autre part, l'annotation du locus d'où provient le transcrit (locus vérifié annoté manuellement ou automatiquement) ; et enfin, comment le transcrit s'aligne avec des banques d'ARNm ou d'ADNc. Ensuite, tous les ARNg ciblant chaque transcrit ont été classifiés selon leur score « on-target » et hors-cible. La combinaison des classements « on-target » et hors-cible a permis de fournir un classement final pour chaque ARNg ciblant un transcrit particulier. Enfin, pour choisir les ARNg pour chaque transcrit, l'ARNg le mieux classé qui cible dans les 5 à 65% de la région codante pour la protéine du gène cible a été choisi. Puis, les ARNg additionnels sont sélectionnés afin que tous les ARNg soient distancés d'au moins 5% de la position du site de coupure d'un autre ARNg.

Ainsi, la librairie lentivirale Brie est composée de 78 637 ARNg ciblant 19 674 gènes avec 4 ARNg par gène et 1 000 ARNg contrôles (Doench et al., 2016).

1.4.1. La génération de la librairie Rétro-Brie

Comme pour la conception de la librairie Rétro-mGeCKO-v2, nous avons donc sous-cloné la collection d'ARNg de la librairie Brie dans le vecteur rétroviral (**Figure 23**). Ce sous-clonage a aussi été réalisé en respectant une couverture de 300X et en suivant les recommandations du protocole de l'équipe de Feng Zhang (Joung et al.,

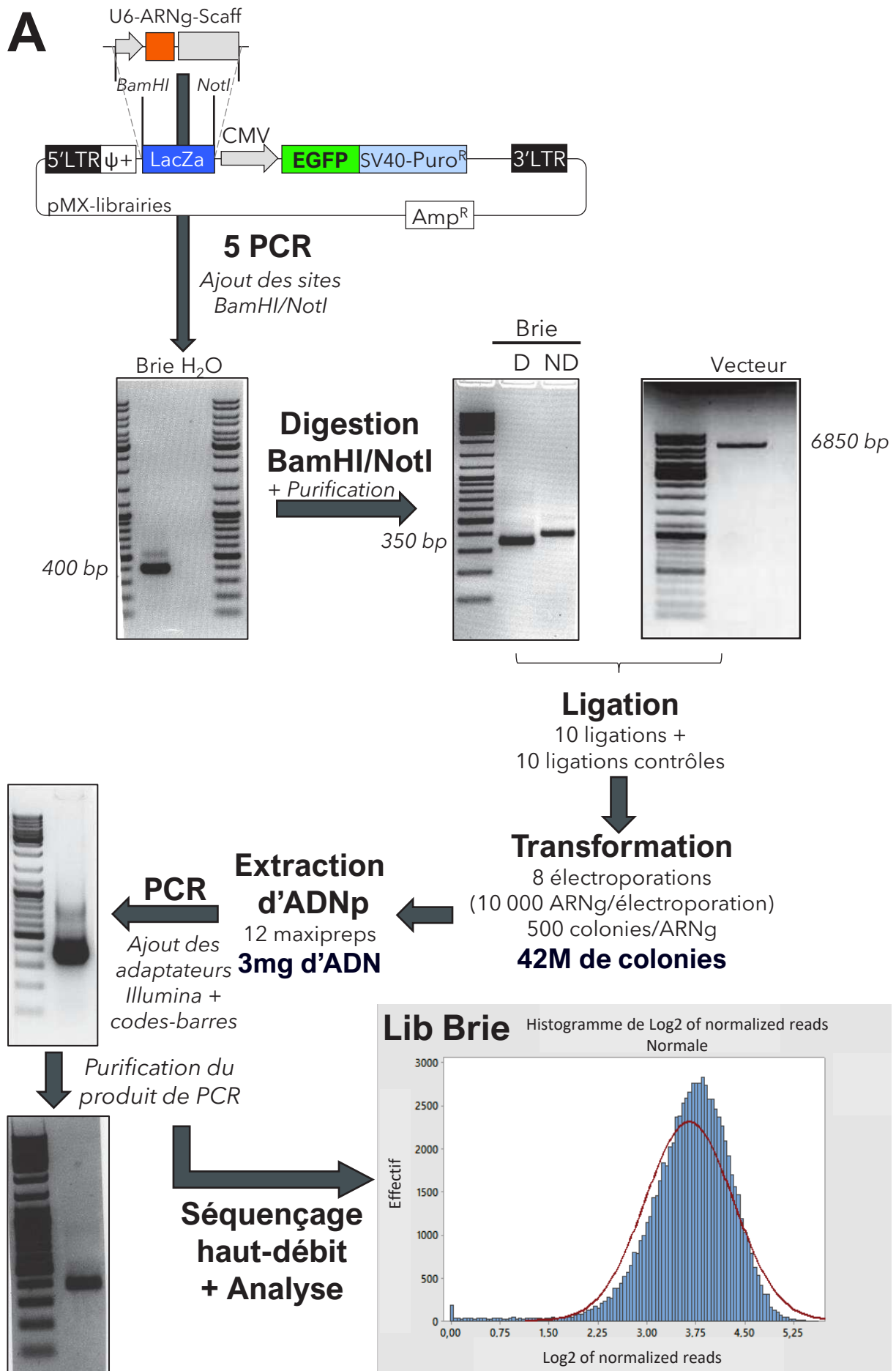


Figure 23: Le clonage de la librairie Rétro-Brie. Légende page suivante.

Figure 23: Le clonage de la librairie Rétro-Brie. La cassette U6-ARNg-Scaffold de la librairie Brie a été amplifiée afin d'ajouter les sites de restriction BamHI/NotI de part et d'autre de la cassette. Puis, le vecteur de destination ainsi que le produit PCR de la cassette U6-ARNg-Scaff de la librairie Brie ont été digérés, purifiés et enfin ligués. Les ligations ont été regroupées et précipitées afin d'être transformées dans des bactéries Stbl4. L'ADN plasmidique (ADNp) de la librairie Brie a été extrait et amplifié par PCR afin d'ajouter les adaptateurs Illumina et les codes-barres de part et d'autres des séquences des ARNg. Le produit PCR a été purifié et soumis au séquençage à haut débit. L'analyse de la qualité de la librairie Brie est représentée par un histogramme représentant le nombre d'ARNg en fonction du Log2 du nombre de reads normalisés.

2017). La cassette U6-ARNg-scaffold a donc été amplifiée par PCR (5 PCRs de 100ng) pour ajouter deux sites de restriction, BamHI et NotI en 5' et 3' de l'amplicon. Ensuite, le produit PCR et le vecteur de destination ont été digérés, purifiés et reliqués ensemble (10 ligations/ pour la librairie et 10 ligations contrôles sans insert). 8 électroporations ont été effectuées. Le calcul de l'efficacité de transformation a montré une bonne efficacité de transformation (42 M de colonies contre 39,3 attendues). Enfin, l'ADNp de la librairie rétrovirale (pMX-mBrie) a été extrait et purifié (12 maxipreps). Afin de pouvoir valider le crible fonctionnel, nous avons ajouté 2 ARNg actifs ciblant AID dans la librairie. Dans le but de valider la qualité de la librairie Rétro-Brie, l'ADN plasmidique de celle-ci a été amplifié par PCR afin d'ajouter les adaptateurs Illumina et les barres-codes. Le produit PCR a été purifié et analysé par séquençage à haut débit (**Figure 23**). Afin de respecter la couverture de 300X de la librairie, la prise en compte de l'obtention d'au moins 24 millions de « reads » par librairie a été considérée. Nous avons obtenu un total de 104 millions de « reads » avec un score de qualité Q30 de 93,47%. Ces résultats montrent que la représentation et la distribution des ARNg ont bien été respectées durant le séquençage haut-débit et que les données sont exploitables.

L'extraction du nombre de « reads » de chaque ARNg a été réalisée en utilisant MAGeCK (Li et al., 2014). Afin d'évaluer la qualité des librairies, les trois paramètres déterminant la qualité d'une librairie clonée ont été calculés. Le pourcentage d'ARNg sans « reads » est à 0,21% et donc inférieur à 0,5%. La différence entre le 90^{ème} percentile et le 10^{ème} percentile est de 3,14X et donc inférieure à 10X. Le pourcentage d'ARNg qui correspondent exactement est à 79,75% et donc supérieur à 70%. De plus, la distribution des ARNg suit parfaitement la courbe de Gauss (**Figure 23**). Ces résultats montrent que la librairie Rétro-Brie est très bien représentée et est de haute qualité.

1.4.2. Le crible fonctionnel sur la CI à l'aide de la librairie Rétro-Brie

1.4.2.1. La réalisation du crible à l'aide de la librairie Rétro-Brie

Comme pour le crible Rétro-mGecKO-v2, j'ai effectué le crible fonctionnel sur la CI en 4 réplicats biologiques avec la librairie Rétro-Brie en utilisant une couverture de

300 fois pour chaque étape (**Figure 24A**). Des cellules BOSC23 ont été transfectées avec la librairie Rétro-Brie afin de produire 4 surnageants rétroviraux différents. La titration de ces surnageants a été réalisée en infectant des cellules CH12^{Cas9}. L'efficacité d'infection a ensuite été évaluée par cytométrie en flux en regardant les cellules GFP+ sur les cellules mCherry+. Seuls les résultats du réplicat R3 sont montrés pour les différentes étapes du criblage. Dans le cas du réplicat R3, la dilution au 1/50^{ème} a été choisie avec un pourcentage de 26,4 correspondant à une MOI de 0.3 (**Figure 24B**). J'ai donc ensuite transduit 132 millions de cellules CH12^{Cas9} avec le surnageant rétroviral à une MOI de 0,3 pour chaque réplicat afin d'obtenir approximativement 24 millions de cellules infectées. Pour chaque réplicat, la couverture se situait à environ 1 000X, avec une efficacité d'infection comprise entre 24% et 30%. Les cellules CH12^{Cas9} transduites ont été sélectionnées avec de la puromycine et de l'hygromycine pendant 15 jours. Durant ces 15 jours, des passages de 40 millions de cellules transduites avec de la puromycine ont été réalisés pour chaque réplicat. La commutation isotypique a été induite avec 35 millions de cellules pour chaque réplicat. 72h après, les cellules B incapables d'effectuer la CI (restant IgM+) ou qui ont subi des événements de CI efficaces (exprimant IgA+) ont été triées à l'aide de billes magnétiques couplées à des anticorps anti-IgM ou anti-IgA. J'ai évalué l'efficacité de la CI et du tri magnétique des cellules en examinant l'expression d'IgA et d'IgM à la surface des cellules par cytométrie en flux. Nous avons obtenu une efficacité de CI robuste de 51 à 63% ainsi qu'une purification efficace de la population IgA+ (88 à 93%) et de la population IgM+ (97%) (**Figure 24C**). L'ADN génomique de 24 millions de cellules des différentes populations cellulaires a été extrait et soumis à des PCR (24 PCR de 5µg/librairie/réplicat) pour amplifier les séquences des ARNg et pour ajouter des adaptateurs Illumina et des codes-barres. Ces PCR ont été optimisées au préalable afin de déterminer le nombre de cycles nécessaires pour chaque échantillon/paire d'oligos pour éviter les biais. Les produits PCR ont été purifiés (**Figure 24D**) puis soumis au séquençage à haut débit pour identifier les ARNg enrichis dans la population IgM+ par rapport à la population IgA+. Afin de respecter la couverture de 300X de la librairie, la prise en compte de l'obtention d'au moins 24 millions de

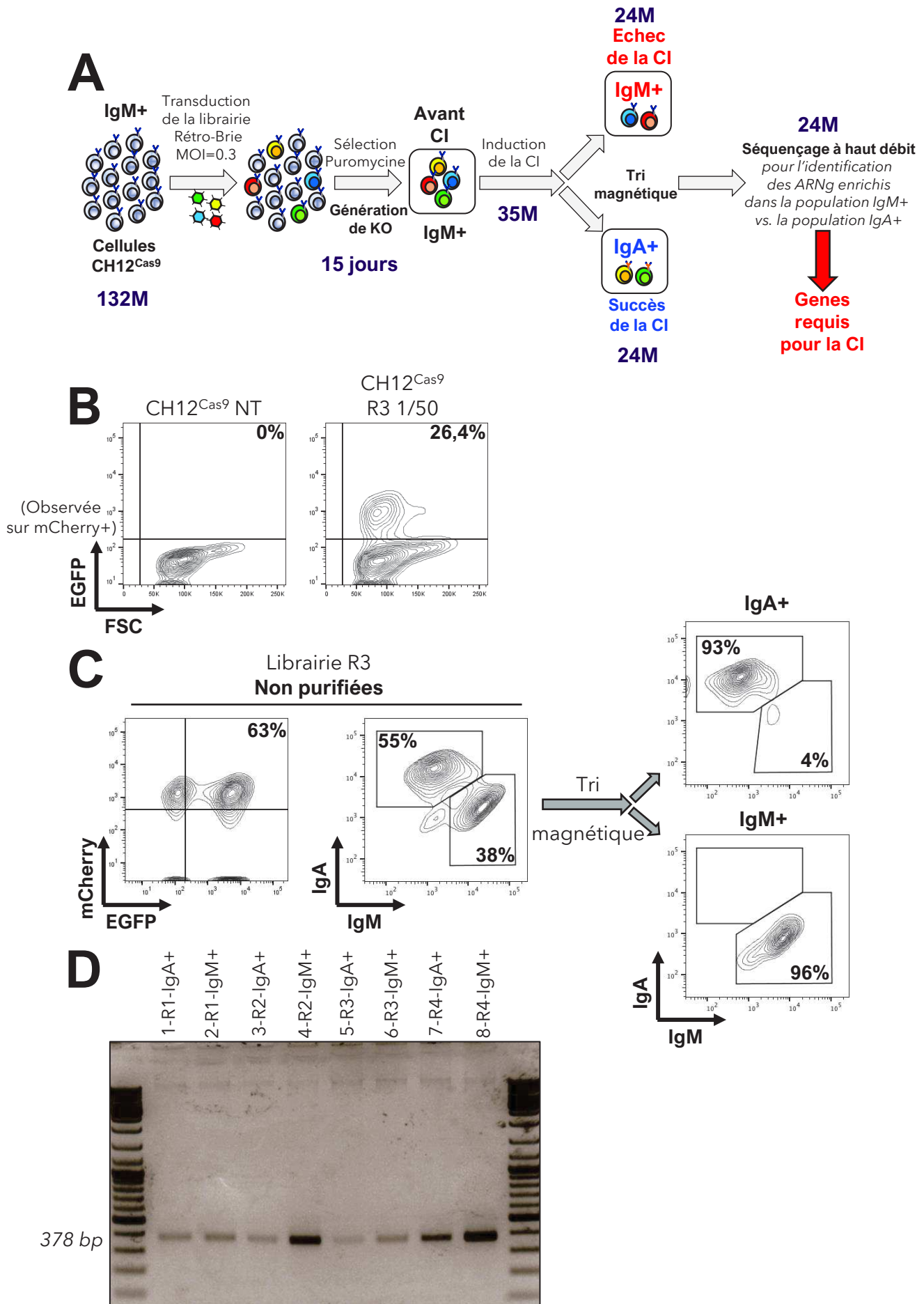


Figure 24: Criblage de la librairie Rétro-Brie. Légende page suivante.

Figure 24: Criblage de la librairie Rétro-Brie. **A.** Vue d'ensemble de la stratégie et des différentes étapes du crible avec la librairie Brie. Le nombre de cellules indiqué en violet pour les différentes étapes correspond au respect de la couverture de 300X de la librairie. **B.** Les cellules CH12^{Cas9} ont été infectées avec différentes dilutions du surnageant rétroviral afin d'effectuer la titration du virus. L'efficacité d'infection a été analysée par cytométrie en flux en regardant le pourcentage de cellules GFP+ dans la population mCherry+. **C.** L'analyse de l'expression d'IgA et d'IgM à la surface des cellules permettant d'évaluer l'efficacité de la CI et la purification par tri magnétique des cellules infectées et stimulées pendant 72h a été réalisée par cytométrie en flux. **D.** L'ADNg des populations IgA+ et IgM+ a été extrait et amplifié par PCR afin d'ajouter les adaptateurs Illumina et les codes-barres de part et d'autre des séquences des ARNg. Les produits PCR ont été purifiés et soumis au séquençage à haut débit.

« reads » par librairie a été considérée. Nous avons obtenu entre 30 et 46 millions de « reads » pour les différents réplicats avec un score de qualité Q30 à environ 95. L'extraction des données brutes a été réalisée avec le logiciel PoolQ,(Broad Institute) et elles ont été normalisées selon le protocole du laboratoire de Feng Zhang (Shalem et al., 2014). Pour analyser l'enrichissement des ARNg des populations IgM+ et IgA+, nous avons aussi utilisé les deux logiciels HitSelect et MAGeCK.

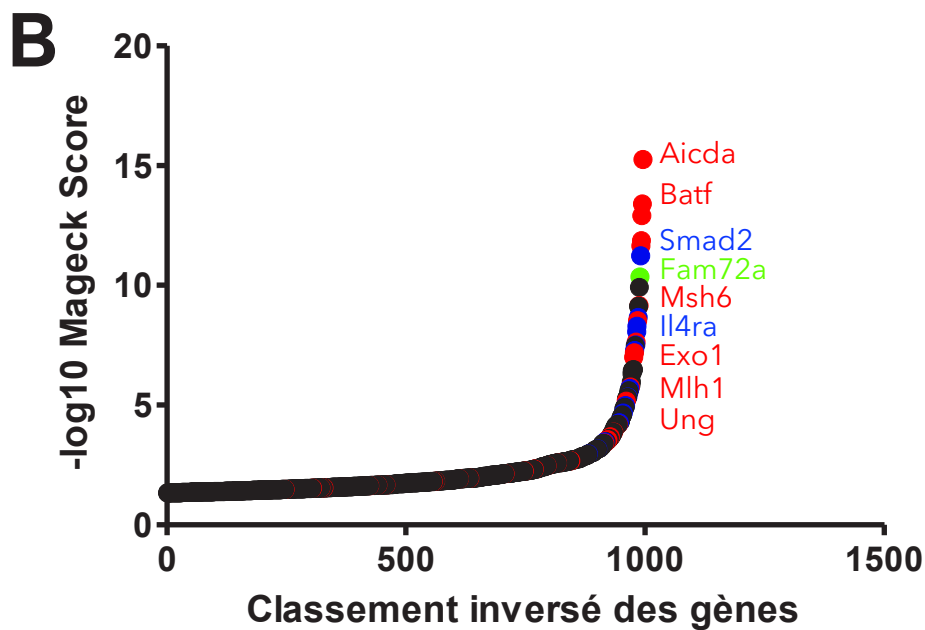
1.4.2.2. L'analyse du crible fonctionnel sur la CI réalisé avec la librairie Rétro-Brie.

Ce crible a permis l'identification de 654 gènes candidats potentiellement impliqués dans les mécanismes moléculaires de la CI (**Figure 25A-B**). Comme le précédent crible, nous retrouvons des gènes enrichis dans la population IgM+ qui sont déjà connus pour être requis pour la CI tels qu'Aicda et Ung, des facteurs de transcription conduisant l'expression d'AID comme Batf et Pten, des facteurs de réparation de l'ADN tels que Ku80, Msh2, Msh6 et les gènes codant pour les sous-unités des récepteurs TGF β , IL-4 et CD40L, comme Tgfbr1/2, Smad2-4, CD40, Il2rg et Il2ra (**Figure 25A**). Également, nous avons aussi identifié un grand nombre de gènes candidats ne possédant pas encore de fonction connue dans les mécanismes de CI (**Figure 25B** ; gènes en noir). Parmi ces gènes candidats se trouvent Fam72a, un gène codant pour une protéine mal caractérisée qui se lie à Ung2 (Guo et al., 2008) (**Figure 25A-B** ; gène en vert fluo).

Suite à l'identification de faux-négatifs et faux-positifs dans le précédent crible, j'ai examiné les gènes connus pour être impliqués dans la CI dans l'ensemble de la liste et seulement quelques-uns sont retrouvés non enrichis dans la population IgM+ avec aucun ou un seul ARNg actif, dont Mcm3, Med1, Fam35a et Hnrnpf (**Figure 25C**). Cependant, l'analyse des données de séquençage avec HitSelect montre que les gènes Mcm3, Fam35a et Hnrnpf sont enrichis dans la population IgM+. Ceci indique que suivant l'algorithme utilisé par les logiciels, les résultats peuvent être différents et que pour une analyse plus complète du crible, il est intéressant de compiler les listes obtenues avec MAGeCK et HitSelect.

A

Gène	Classe-ment	ARNg actifs	p-value
Aicda	1	6	2,3947E-07
Batf	2	4	2,3947E-07
Pten	3	3	2,3947E-07
Foxo1	4	4	2,3947E-07
Irf4	5	4	2,3947E-07
Smad2	6	4	2,3947E-07
Fam72a	7	4	2,3947E-07
Plekha2	8	4	2,3947E-07
Msh6	9	4	2,3947E-07
Atxn7l3	10	4	2,3947E-07
Runx3	11	4	2,3947E-07



C

Gène	Classe-ment	ARNg actifs	p-value
Mcm3	11358	0	0,56333
Med1	8598	0	0,4262
Fam35a	2250	1	0,11523
Hnrnpf	2548	1	0,13046

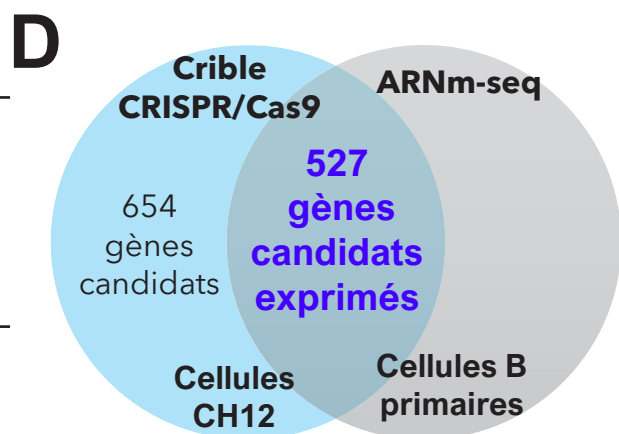


Figure 25: Le crible CRISPR/Cas9 avec la librairie Rétro-Brie permet d'identifier des gènes connus pour être impliqués dans la CI et de nouveaux gènes candidats. Légende page suivante.

Figure 25: Le crible CRISPR/Cas9 avec la librairie Rétro-Brie permet d'identifier des gènes connus pour être impliqués dans la CI et de nouveaux gènes candidats. **A.** Liste des 11 gènes les plus fortement enrichis issue de l'analyse MAGeCK. **B.** Graphique réalisé avec les données générées par l'analyse MAGeCK avec une p-Value inférieure à 0,05 et au moins 2 ARNg actifs, représentant le $-\log_{10}$ du Score Mageck en fonction du classement inversé des gènes. **C.** Liste de gènes connus pour être impliqués dans la CI mais qui ne sont pas enrichis dus à des ARNg inactifs. **D.** Diagramme de Venn représentant en bleu, les résultats du crible CRISPR/Cas9 avec la librairie Rétro-Brie issus de l'analyse MAGeCK avec une p-Value inférieure à 0,05 et au moins 2 ARNg actifs et en gris, les résultats d'un ARNm-Seq réalisé dans des cellules B primaires. Les gènes en rouge représentent les gènes requis pour une CI efficace. Les gènes en bleu représentent les gènes impliqués dans la CI avec un rôle dans les voies de signalisation du TGF β , IL-4, CD40, BCR, NF- κ B. Les gènes en noir représentent des gènes n'ayant pas encore de fonction connue dans la CI et enfin le gène Fam72a en vert fluo est le gène candidat choisi pour une caractérisation plus approfondie.

Le recoupement de la liste de gènes issue de l'analyse MAGeCK avec la liste des gènes exprimés dans les cellules B primaires issue de l'ARNm-Seq montre que 527 gènes ont été trouvés exprimés dans les cellules B primaires (**Figure 25D**). Ceci indique que 20% des gènes enrichis dans la population IgM+ peuvent constituer des faux-positifs. Ainsi, la librairie d'ARNg Brie semble aussi générer des faux-positifs.

Dans l'ensemble, nous avons montré que notre crible fonctionnel sur la CI à l'échelle du génome murin à l'aide de la librairie Brie permet l'identification de gènes déjà connus pour être nécessaires pour une CI efficace, et de nouveaux cofacteurs de la CI, indiquant que le crible est efficace. Même si le crible avec la librairie Brie est plus précis que celui avec la librairie mGeCKO-v2, notre liste de gènes candidats comporte également des faux-négatifs et des faux-positifs. Néanmoins, le crible nous a permis d'identifier *Fam72a*, un gène codant pour une protéine mal caractérisée qui se lie à UNG2 (Guo et al., 2008) que nous avons décidé de caractériser plus en profondeur (*partie II*). De plus, nous avons aussi identifié *Erc612* qui a été récemment publié (Liu et al., 2020).

1.5. Analyse globale des deux cribles.

Enfin, dans le but d'avoir une analyse globale des deux cribles réalisés pour l'identification de nouveaux cofacteurs de la CI, j'ai établi un recoupement des listes générées avec MAGeCK (MG) et HitSelect (HS) des deux cribles (**Figure 26**). Celui-ci montre que d'une part, toutes les listes ont seulement 30 gènes en commun qui sont principalement tous déjà connus pour être impliqués dans la CI ; d'autre part, les listes générées par les deux logiciels pour la librairie Brie ont 259 gènes en commun ; et enfin, les listes générées par les deux logiciels pour la librairie mGeCKO-v2 ont 200 gènes en commun (**Figure 26A**). L'analyse globale des deux cribles avec MAGeCK, identifiant 1232 gènes candidats, montre un enrichissement significatif des processus biologiques impliqués dans les mécanismes de CI tels que la réparation de l'ADN, la transcription ou la recombinaison (**Figure 26B**). Enfin, le regroupement de tous les gènes ayant déjà un rôle connu lors de la CI (**Figure 26C**)

montre la puissance des cribles CRISPR/Cas9 dans l'identification des cofacteurs de la CI.

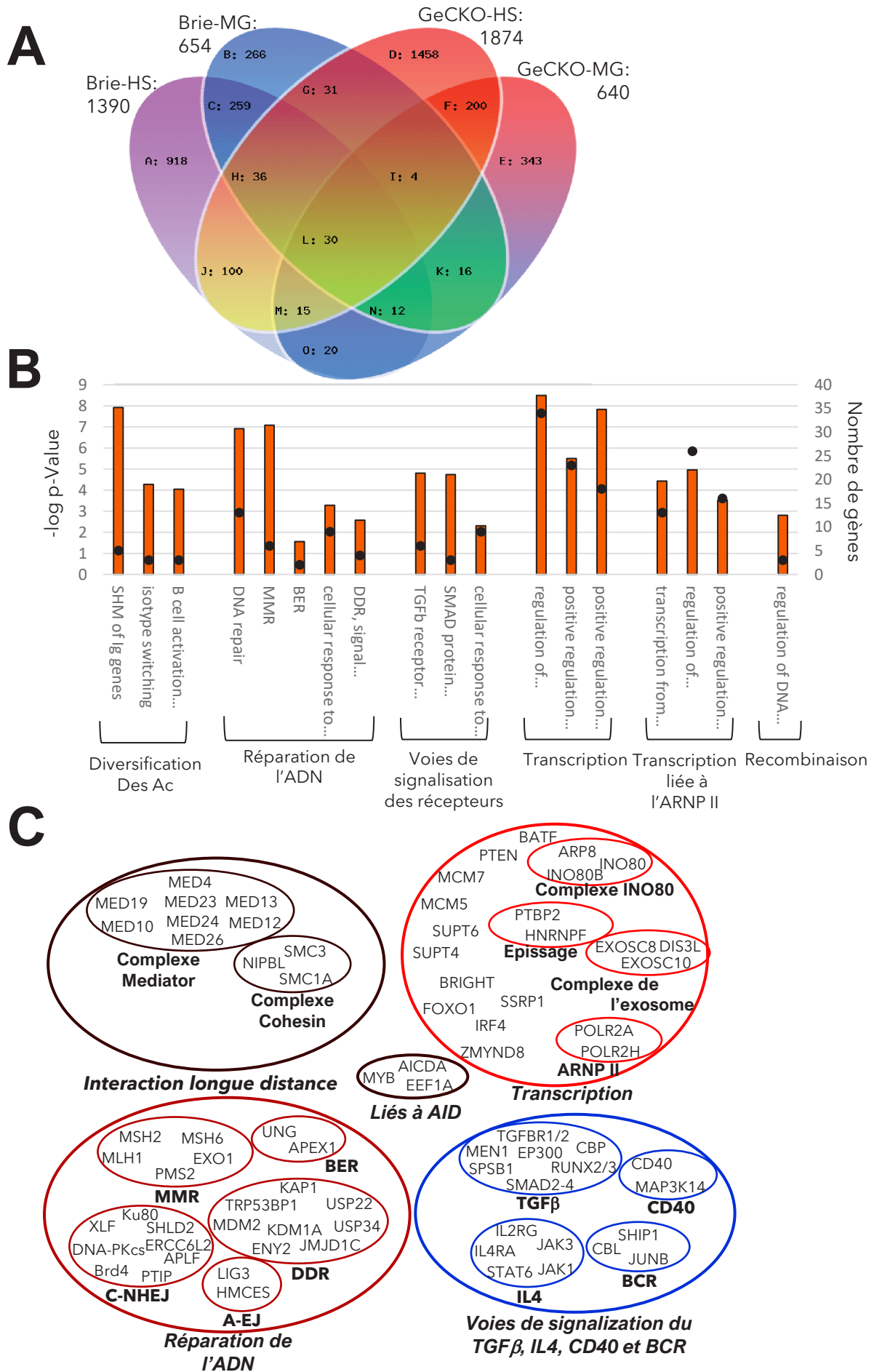


Figure 26: Analyse des deux cribles CRISPR/Cas9 réunis. Légende page suivante.

Figure 26: Analyse des deux cribles CRISPR/Cas9 réunis. A. Diagramme de Venn représentant la réunion des listes générées avec HitSelect et MAGeCK des cribles avec les bibliothèques Brie et mGeCKO-v2 (avec une FDR ou p-Value inférieure à 0,05 et au moins 2 ARNg actifs). MG: MAGeCK; HS: HitSelect. **B.** Processus biologiques significativement enrichis dans la liste d'analyse MAGeCK des cribles avec les bibliothèques mGeCKO-v2 et Brie avec une p-Value inférieure à 0,05 et au moins 3 ARNg actifs. Les barres oranges représentent le $-\log_{10}$ de la valeur p-Value calculée par le logiciel Enrichr (Kuleshov et al. 2016). Les points noirs représentent le nombre de gènes enrichis dans le processus biologique. **C.** Gènes groupés selon leur rôle lors de la CI présents dans les listes générées avec HitSelect et MAGeCK des cribles CRISPR mGeCKO-v2 et Brie avec une FDR ou p-Value inférieure à 0,05 et avec au moins 2 ARNg actifs.

II- Discussion

La CI est un mécanisme complexe qui est régulé à plusieurs niveaux impliquant des dommages « physiologiques » à l'ADN nécessaires pour la diversification des Ac, mais qui peuvent être également dans certains cas, la cause de l'apparition de lymphomes (Küppers, 2005; Nussenzweig and Nussenzweig, 2010). Les mécanismes qui régulent la CI ont fait l'objet d'études intensives durant ces dernières années, laissant des questions non élucidées. Nous avons donc développé et mis en place une stratégie dans le but d'identifier les facteurs impliqués dans les mécanismes moléculaires permettant une CI efficace. Cette stratégie repose sur la réalisation d'un crible CRISPR/Cas9 à l'échelle du génome murin. L'aboutissement de deux cribles fonctionnels nous a permis l'identification de 1232 gènes candidats engagés dans les mécanismes de CI. Parmi ceux-ci, nous retrouvons de nombreux gènes déjà connus pour être requis dans la CI (Aicda, Ung), des facteurs de la réparation de l'ADN (Ku80, Msh2, Msh6, 53bp1), des facteurs de transcription conduisant l'expression d'AID (Batf, Pten) ainsi que les membres des voies de signalisation TGF- β , Il-4 et CD40 (Tgfb1/2, Il4r, CD40). L'identification de ces facteurs déjà connus nous permet de conclure que les cribles visant à déterminer les cofacteurs de la CI ont été fonctionnels. De plus, la publication récente de nouveaux facteurs impliqués dans les mécanismes de CI tels que les facteurs Shieldin2 (Ghezraoui et al., 2018; Gupta et al., 2018; Liang et al., 2020; Noordermeer et al., 2018; Setiaputra and Durocher, 2019), Zmynd8 (Delgado-Benito et al., 2018), Errc6l2 (Liu et al., 2020), HMCES (Mohni et al., 2019), que nous retrouvons également dans notre liste appuie la puissance de notre crible CRISPR/Cas9. Enfin, nous avons aussi identifié de nombreux gènes candidats n'ayant pas de fonction connue dans la CI. Parmi ceux-ci, nous avons trouvé Fam72a, un gène qui code une protéine mal caractérisée qui a été montré s'associer avec Ung2 (Guo et al., 2008), facteur requis pour l'élimination des uraciles dans l'ADN induits par AID.

Néanmoins, nous avons rencontré quelques limitations lors de la réalisation des cribles qui peuvent avoir un impact dans l'interprétation des résultats. La première limitation rencontrée concerne le développement des criblages CRISPR/Cas9 uniquement basé sur l'utilisation de vecteurs lentiviraux. En effet, j'ai montré qu'il est impossible de transduire notre lignée cellulaire CH12 avec le LentiCRISPR-v6. Suite à des discussions avec d'autres équipes ayant rencontré ce même problème, ce vecteur ne semble pas avoir une transduction efficace dans les cellules du système immunitaire (cellules NK, cellules B et T). Nous avons pu dépasser cette limitation en sous-clonant les bibliothèques d'ARNg dans un vecteur rétroviral (pMXPIE) capable de transduire efficacement les cellules B CH12. Cependant, durant les étapes de sous-clonage, il est tout à fait possible de perdre un certain nombre d'ARNg. Nous avons observé une perte de 459 ARNg pour la bibliothèque A, 2177 pour la bibliothèque B et 197 pour la bibliothèque Brie. Nous avons très probablement perdu ces ARNg durant les étapes de sous-clonage. Néanmoins, il est aussi possible que certains n'étaient pas présents dans le pool d'ARNg reçu. Toutefois, un des trois paramètres permettant de valider la qualité d'une bibliothèque d'ARNg prend en compte cette perte d'ARNg. La seconde limitation concerne la présence d'ARNg inactifs dans les bibliothèques qui peut mener à des faux-négatifs. Alors que la cohérence des différents ARNg ciblant le même gène est le fondement de l'analyse de ces cribles CRISPR/Cas9, de nombreuses études précédentes ont montré une grande variabilité dans l'efficacité des ARNg à produire des allèles nuls (Chu et al., 2016; Doench et al., 2014; Wang et al., 2014a). Les propriétés « on-target » des ARNg ont alors fait l'objet d'études qui ont permis de déterminer des caractéristiques et des motifs dans la séquence du PAM et de l'ARNg qui affectent l'efficacité des ARNg (Doench et al., 2014; Wang et al., 2014a). Ces deux études ont permis notamment de générer de nouvelles bibliothèques d'ARNg avec des propriétés « on target » maximisées, telles que la bibliothèque Brie (Doench et al., 2016) et Sabatini (Wang et al., 2014a). Cependant, comme je l'ai montré avec les résultats du second crible réalisé avec la bibliothèque Brie, des ARNg inactifs sont encore présents. Si ces caractéristiques ou motifs ont contribué à améliorer les critères de conception des ARNg, ils ne permettent pas de dire avec certitude qu'un ARNg est inefficace. Toutefois, il a été également mentionné que la structure locale de la chromatine est un facteur majeur pouvant affecter la capacité

de la nucléase Cas9 à trouver le PAM et à commencer à se lier à l'ADN avec la région « seed » de l'ARNg (les 10-12 premières bases de l'ARNg adjacentes à la séquence PAM) (Kuscu et al., 2014; Wu et al., 2014). La troisième limitation rencontrée est la présence de faux-positifs que nous avons détecté en comparant les listes de gènes candidats engagés dans la CI obtenues dans la lignée cellulaire CH12 qui est dérivée d'un lymphome B murin et une liste de gènes exprimés dans les cellules B primaires issue d'un ARNm-Seq. Même si des différences d'expression génique peuvent être observées entre ces deux lignées, le pourcentage de gènes enrichis n'étant pas exprimés dans des cellules B primaires est bien trop élevé pour conclure que celui-ci est seulement dû à une différence d'expression génique. En effet, nous pensons que des effets hors-cible sont également responsables de la présence de ces faux-négatifs. Afin de contrer ces deux dernières limitations, des études très récentes ont décrit des outils flexibles pour la conception d'ARNg présentant une activité et une spécificité optimales en limitant les potentiels sites hors-cible (Chu et al., 2016; Dhanjal et al., 2019; Graf et al., 2019; Lee et al., 2018; Meier et al., 2017; Michlits et al., 2020; Veeneman et al., 2020). Enfin, la dernière limitation de ces cribles d'inactivation par CRISPR/Cas9 à large échelle est qu'il est impossible de distinguer des facteurs présentant des redondances fonctionnelles, ce qui rend l'identification de certains facteurs inapplicable.

Ces trois dernières limitations sont semblables à celles décrites lors de la réalisation de cribles basés sur l'utilisation de shARNs. N'ayant pas réalisé de crible fonctionnel à l'aide de shARNs, il m'est donc impossible de comparer ces deux approches. Cependant, certaines études qui comparent ces deux approches montrent que le criblage CRISPR/Cas9 reste plus efficace et présente un taux plus faible de faux positifs que le criblage basé sur l'utilisation de shARNs (Peng et al., 2015; Smith et al., 2017).

Malgré les limitations de notre crible, nous avons identifiés des gènes candidats n'ayant aucune fonction connue dans la CI qui peuvent être requis pour une CI efficace. Dans le but de valider ces gènes candidats, nous avons décidé de réaliser un crible secondaire ciblant les 1 000 gènes les plus enrichis identifiés à partir des deux cribles. Afin de limiter les effets hors cible, nous avons décidé dans un premier

temps de générer une lignée cellulaire CH12 exprimant la nucléase Cas9-HF (Haute-Fidélité) (Kleinstiver et al., 2016). Malheureusement, la lignée générée n'a pas montré la capacité d'induire des DSBs lors de la transfection d'ARNg connus pour être efficace avec cette nouvelle Cas9-HF. Nous pensons donc réaliser ce crible dans la lignée cellulaire CH12^{Cas9} avec un pool d'ARNg provenant de la librairie Sabatini (Wang et al., 2014a). La librairie comptera donc 11 500 ARNg (6 ARNg/gène) dont 128 ARNg contrôles ainsi que des ARNg ciblant des gènes connus pour être requis pour une CI efficace (Aicda, Batf, Ung, ...).

Il est intéressant de noter que l'équipe de K. Rajewsky a établi une méthode de criblage CRISPR/Cas9 à petite échelle applicable dans des cellules B spléniques. En effet, cette stratégie repose sur la génération d'un modèle murin exprimant la nucléase Cas9 fusionnée à la Cas9 et l'infection rétrovirale de cellules B spléniques Cas9iGFP/+ et WT avec des ARNg clonés individuellement. Afin d'avoir un contrôle interne, les cellules B spléniques Cas9iGFP/+ et WT sont mélangées dans un rapport 1 :4 (Chu et al., 2016; Graf et al., 2019). Cette méthode pourrait permettre de valider plus rapidement les gènes candidats obtenus d'un premier crible.

En conclusion, nous avons identifié avec succès 1232 gènes candidats engagés dans les mécanismes de CI grâce à deux cribles d'inactivation par la technologie CRISPR/Cas9. Les limitations dues au système CRISPR/Cas9 donnent une liste imparfaite de gènes candidats qui peut être améliorée en la recoupant avec d'autres cribles ou listes de gènes exprimés spécifiquement dans les cellules B. Etant donné la nouveauté des cribles CRISPR/Cas9, des points sont à améliorer mais beaucoup d'études récentes s'attèlent à la tâche. Cette technologie est prometteuse et puissante pour étudier les phénomènes biologiques complexes dépendants de nombreux facteurs et mécanismes moléculaires, tels que les mécanismes de diversification des anticorps.

CHAPITRE 2 :

Etude du rôle de Fam72a dans la diversification des immunoglobulines.

Contexte

La réalisation d'un crible fonctionnel CRISPR/Cas9 visant à identifier les acteurs moléculaires impliqués lors de la CI nous a permis d'identifier le gène Fam72a, un gène codant pour une protéine mal caractérisée qui se lie à UNG2 (Guo et al., 2008). Les travaux réalisés sur la caractérisation du facteur Fam72a dans les mécanismes de diversification des anticorps sont inclus dans l'ébauche du papier suivant.

- I- 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^{1,2,3,4,*}, Jacques Moritz^{1,2,3,4,*}, Isabelle Robert^{1,2,3,4}, Chloé Lescale⁵, Vincent Heyer^{1,2,3,4}, Anne-Sophie Thomas-Claudepierre^{1,2,3,4}, Arthur Abello^{1,2,3,4}, Ludovic Deriano⁵ and Bernardo Reina-San-Martin^{1,2,3,4,#}.

¹ Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Illkirch, France.

² Institut National de la Santé et de la Recherche Médicale (INSERM), U1258, Illkirch, France.

³ Centre National de la Recherche Scientifique (CNRS), UMR7104, Illkirch, France.

⁴ Université de Strasbourg, Illkirch, France.

⁵ 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

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)¹ and by the subsequent excision of the resulting uracils by Uracil DNA glycosylase (UNG) and by mismatch repair (MMR) proteins²⁻⁴. Although uracils arising in DNA are faithfully repaired²⁻⁷, 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^{2,4,8}. 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)⁹ and that is overexpressed in several cancers⁹. We show that the FAM72A-UNG2 interaction controls the protein levels of UNG2 and that CSR is defective in *Fam72a*^{-/-} B cells due to the specific upregulation of UNG2. Moreover, we show that in *Fam72a*^{-/-} 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^{2,3}. 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^{2,3}. CSR modulates antibody effector functions by replacing the isotype expressed from IgM/IgD to IgG, IgA or IgE¹⁰. 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)¹⁰ 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 isotype¹⁰. Mechanistically, both SHM and CSR are triggered by activation induced cytidine deaminase (AID), an enzyme which deaminates cytosines into uracils in DNA¹¹. AID-generated U:G mismatches are processed mainly by the base-excision repair (BER) enzyme uracil-DNA-glycosylase (UNG)^{4,12} but also by mismatch repair (MMR)¹³ proteins to introduce mutations or double-stranded DNA breaks (DSBs) during SHM and CSR, respectively²⁻⁴. 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 DSBs¹⁰, which are a necessary intermediates for recombination³. While uracils in DNA are usually faithfully repaired²⁻⁷, 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 unknown²⁻⁴.

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)¹⁴ gRNA library and using the IgM⁺ 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- β , IL-4 and an anti-CD40 antibody¹⁵. To conduct the screen, we established a CH12 cell line stably expressing Cas9 (CH12^{Cas9}; **Fig. S1A**) and sub-cloned the gRNA library into a retroviral vector (pMX-mBrie; **Fig. S1B**), which efficiently transduces CH12 cells. CH12^{Cas9} 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 sequencing¹⁶ (**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- β , 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 manner⁹ and which is overexpressed in multiple cancers⁹. To validate the screen and determine whether *Fam72a* is required for CSR, we generated *Fam72a*^{-/-} CH12 cells using the high-fidelity Cas9 (Cas9-HF1)¹⁷ (**Fig. S2A and S2B**). Interestingly, we found that the efficiency of CSR was reduced by 50-60 % in four independent *Fam72a*^{-/-} CH12 clones when compared to *Fam72a*^{+/+} 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*^{-/-} mouse model (**Fig. S3A**) generated by the knockout mouse program (KOMP)¹⁸. *Fam72a*^{-/-} mice gave offspring at Mendelian ratios and showed no obvious deleterious phenotype. *Fam72a*^{-/-} 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*^{-/-} 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*^{-/-} 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^{W125R} or FAM72A^{W125A} UNG2-binding defective mutants⁹ in *Fam72a*^{-/-} 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^{W125R} or FAM72A^{W125A} mutants were unable to rescue the CSR defect (**Fig. 3A and 3C**). Furthermore, we found that the residual CSR activity observed in *Fam72a*^{-/-} B cells is dependent on the catalytic activity of UNG, as expression of an UNG-inhibitor (Ugi)¹⁹ 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) isoform²⁰ (**Fig. S2F**), both of which have been shown to be able to sustain CSR²¹. To examine the relative contribution of both these isoforms relative to FAM72A, we generated *Ung1*^{-/-}, *Ung2*^{-/-}, *Ung*^{-/-}, *Fam72a*^{-/-} *Ung1*^{-/-} and *Fam72a*^{-/-} *Ung2*^{-/-} CH12 cell clones using CRISPR/Cas9-HF1 (**Fig. S2F**) and tested their ability to undergo CSR. As shown before^{12,21}, while *Ung1*^{-/-} and *Ung2*^{-/-} cells underwent CSR at wildtype levels (**Fig. 3D and S2G**), CSR in *Ung*^{-/-} CH12 cells was completely abolished (**Fig. 3D and S2G**). Surprisingly, while *Fam72a*^{-/-} *Ung2*^{-/-} rescued the CSR defect observed in *Fam72a*^{-/-} CH12 cells, *Fam72a*^{-/-} *Ung1*^{-/-} 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*^{-/-} primary B cells (**Fig. 3G**). Furthermore, when we overexpressed FAM72A in *Fam72a*^{-/-} CH12 cells (**Fig. S2B**), the up-regulation of UNG2 was almost abolished (**Fig. 3E**). Significantly, overexpression of the FAM72A^{W125A} mutant (**Fig. S2B**), which does not interact with UNG2⁹, 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*^{-/-} and *Fam72a*^{-/-} *Ung1*^{-/-} 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*^{-/-} B cells²²⁻²⁴. 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 CSR²⁵ and at a region spanning immunoglobulin variable regions during SHM³. 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 μ switch regions in *Fam72a*^{-/-} 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*^{-/-} 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_H4 intron (J_H4i), a sequence which gets heavily mutated by AID during SHM *in vivo*²⁶, in germinal center B cells isolated from the Peyer's patches of unimmunized *Fam72a*^{+/+} and *Fam72a*^{-/-} mice (**Fig. 4A**). Interestingly, we found that the mutation frequency at the J_H4i sequence was drastically reduced (five-fold) in *Fam72a*^{-/-} 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*^{+/+} to 16 % in *Fam72a*^{-/-} 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*^{+/+} to 84 % in *Fam72a*^{-/-} B cells (**Fig. 4C and**

S4B). This phenotype is diametrically opposed to *Ung*^{-/-} B cells¹², in which transition mutations at C:G base pairs reaches 95 % and transversion mutations at C:G base pairs are suppressed to 5 %¹². 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^{3,10}. Indeed, AID is haplo-insufficient and CSR is reduced in *AID*^{+/-} B cells²⁷. The CSR defect observed in *Fam72a*^{-/-} 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 translocations²⁸. 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*^{-/-}, *Ung*^{-/-}, *Fam72a*^{+/+} and *Fam72a*^{-/-} cells reconstituted (or not) with FAM72A or FAM72A^{W125R} and quantified aberrant metaphases in these cells (**Fig. 4D, 4E and Table S1**). We found lower levels of aberrant metaphases in *Fam72a*^{-/-} cells or *Fam72a*^{-/-} cells reconstituted with FAM72A^{W125R} as compared to wild type and *Fam72a*^{-/-} 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 β ^{5,6,29,30}. 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 lines⁹ and as *Ung*^{-/-}

mice develop B cell lymphomas³¹⁻³³, 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> (
- 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^{Cas9} 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 described¹⁶ 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 described¹⁶.

CRISPR/Cas9 knockout screen.

CH12^{Cas9} 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 described¹⁶. 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 described¹⁶. Ranking of candidate genes was performed using MaGeCK³⁴. 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 β -mercaptoethanol (50 μ M).

CSR assays.

CH12 cells were cultured for 72h in the presence of TGF- β (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 μ g/ml, Sigma-Aldrich), IL-4 (25 μ g/ml, Preprotech), anti-IgD-Dextran (6 ng/ml, Fina Biosolutions), IFN- γ (100 ng/ml, Preprotech), IL-5 (5 ng/ml, R&D Systems), TGF- β (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 β 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^{W125R} (pMX-mFam72a^{W125R}), mFam72a^{W125A} (pMX-Fam72a^{W125A}), mUng2 (pMX-mUng2) or Ugi (pMX-Ugi). Transduced cells were then selected with puromycin (1 μ 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 nuclease¹⁷ 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 μ and JH4i somatic hypermutation analysis.

For S μ mutation analysis, genomic DNA was extracted from CH12 cells cultured for 3 days 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 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 μ 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 MEGAWHOP³⁵ method. Inserts were sequenced (Sanger sequencing) using the M13 Forward universal primer. Sequences were aligned with Lasergene (DNASTAR) and analyzed with the SHMTool³⁶ server.

Cloning.

All gRNA/Cas9HF1 plasmids were generated through golden gate cloning³⁷. cDNAs were cloned into the pMX-PIE retrovirus through SLICE³⁸.

Mice.

Fam72a^{-/-} mice (C57BL6) were generated by the knockout mouse program (KOMP)¹⁸. 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 procedures³⁹. DNA FISH on metaphases spreads was performed as previously described³⁹ 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-AID⁴⁰, anti-β-Actin (Sigma), anti-UNG1/2 (Gift from B. Kavli) antibodies. Cell fractionation experiments were performed as previously described⁴¹.

B cell development.

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

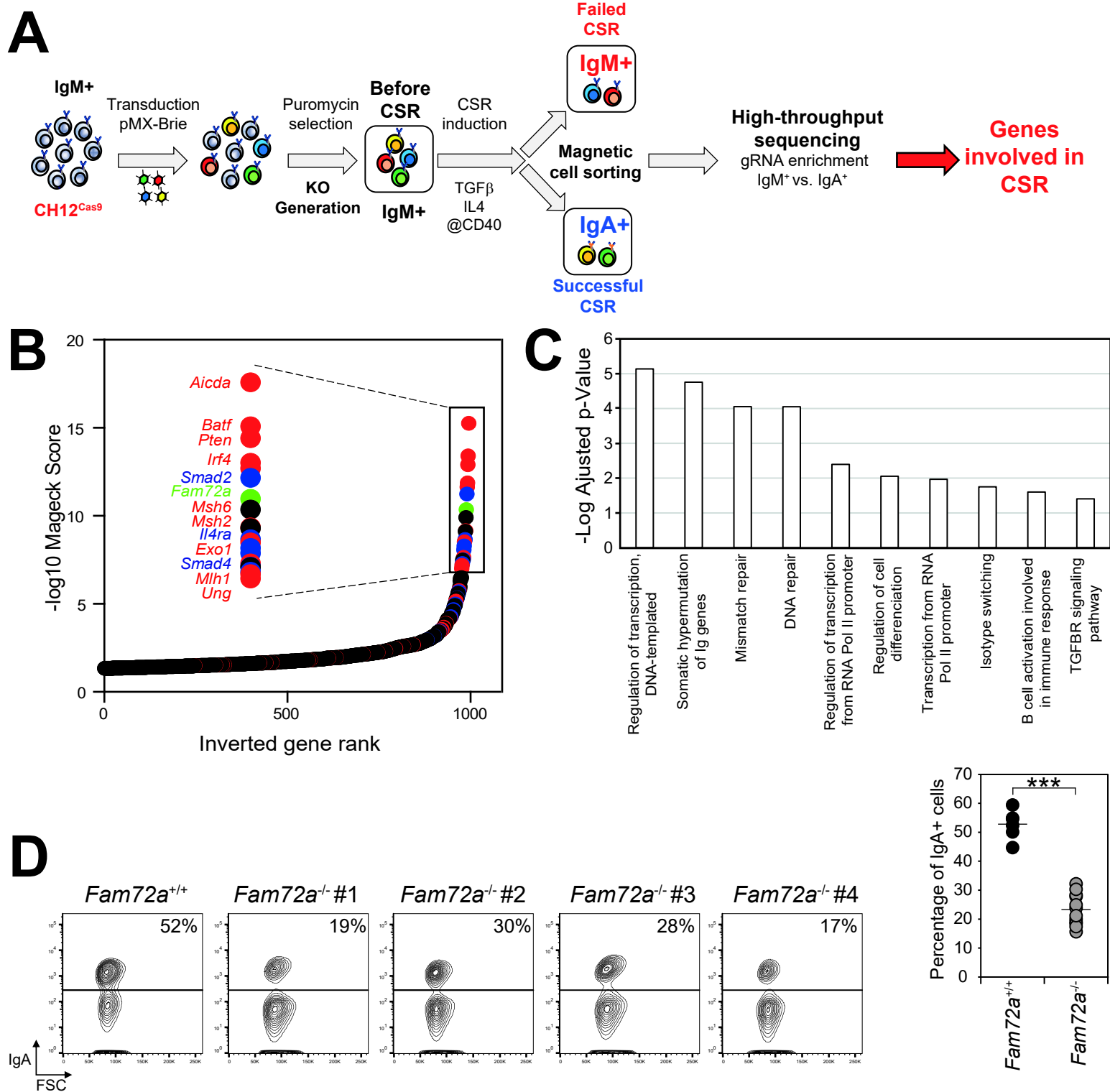


Figure 1. A genome-wide CRISPR/Cas9 knockout screen for genes involved in CSR identifies *Fam72a*. (A) Schematic overview of the CRISPR/Cas9 screen. (B) Graph depicting the gene significance score of the ranking ($-\log_{10}$ Score Mageck) calculated with the MAGeCK algorithm, plotted against the logarithmic 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- β , 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*^{+/+} and four independent *Fam72a*^{-/-} CH12 cell clones cultured for 3 days with TGF- β , 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.

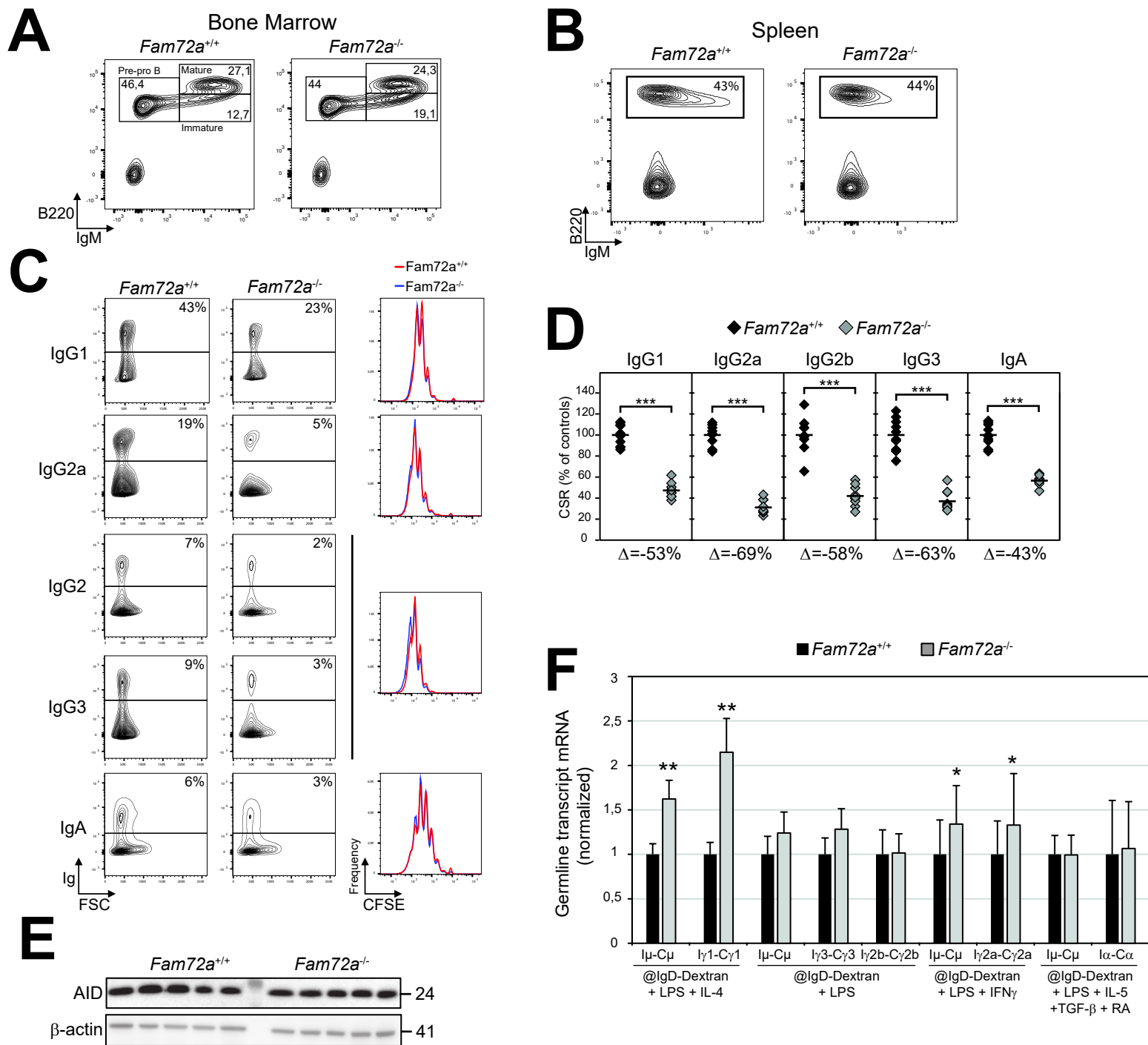


Figure 2. CSR is defective in B cells from *Fam72a*^{-/-} mice. Flow cytometry analysis of bone marrow (**A**) and spleen (**B**) cells from *Fam72a*^{+/+} and *Fam72a*^{-/-} mice. Staining antibodies and percentage of cells within gates are indicated. (**C**) Flow cytometry analysis of Ig expression in *Fam72a*^{+/+} and *Fam72a*^{-/-} splenic B cells cultured for 96h with either LPS + IL-4 + anti-IgD-Dextran (CSR to IgG1), LPS + IFN-γ + anti-IgD-Dextran (CSR to IgG2a) or LPS + anti-IgD-Dextran (CSR to IgG2b and IgG3) or LPS + IL-5 + TGF-β + 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*^{-/-} B cells relative to *Fam72a*^{+/+} B cells for the different isotypes from 10 independent mice is shown on the right. A horizontal line indicates mean values. The difference (Δ) in the percentage of CSR for each isotype between *Fam72a*^{+/+} and *Fam72a*^{-/-} 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 β-Actin in *Fam72a*^{+/+} and *Fam72a*^{-/-} 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_μ) and acceptor switch regions (GLT_{γ3}, GLT_{γ1}, GLT_{γ2b}, GLT_{γ2a} and GLT_α) in *Fam72a*^{+/+} and *Fam72a*^{-/-} splenic B cells cultured for 96h as in (C). Expression is normalized to Igβ and is presented relative to expression in *Fam72a*^{+/+} 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).

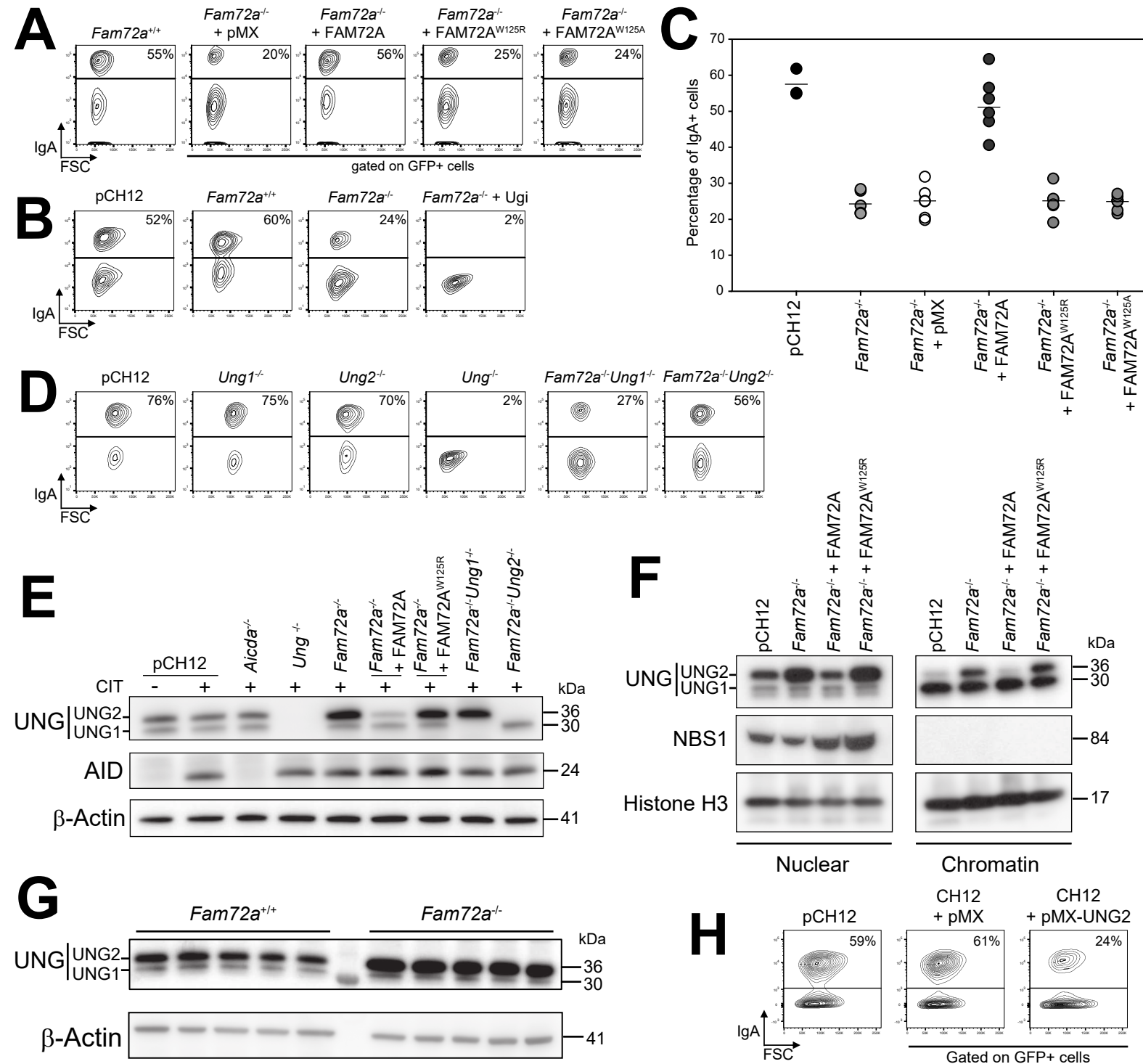


Figure 3. The CSR defect in *Fam72a*^{-/-} B cells is due to the specific up-regulation of *Ung2*. (A) Flow cytometry analysis of IgA expression in *Fam72a*^{-/-} CH12 cells transduced with an empty retrovirus (pMX) or expressing FAM72A, FAM72A^{W125A} or FAM72A^{W125R} and cultured for 3 days with TGF- β , 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*^{+/+} and *Fam72a*^{-/-} CH12 cells expressing (or not) an UNG inhibitor (Ugi) and cultured for 3 days with TGF- β , 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*^{-/-}, *Ung1*^{-/-}, *Ung2*^{-/-}, *Ung*^{-/-}, *Fam72a*^{-/-} *Ung1*^{-/-} and *Fam72a*^{-/-} *Ung2*^{-/-} CH12 cells cultured for 72h with TGF- β , 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 β -Actin in *Fam72a*^{-/-}, *Ung1*^{-/-}, *Ung2*^{-/-}, *Ung*^{-/-}, *Fam72a*^{-/-} *Ung1*^{-/-}, *Fam72a*^{-/-} *Ung2*^{-/-} and *Fam72a*^{-/-} CH12 cells transduced with a retrovirus expressing FAM72A or FAM72A^{W125R} and cultured for 72h with TGF- β , 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*^{-/-} B cells expressing FAM72A or FAM72A^{W125R}. (G) Western blot analysis for UNG (UNG1 and UNG2) and β -Actin protein expression levels in *Fam72a*^{+/+} and *Fam72a*^{-/-} 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.

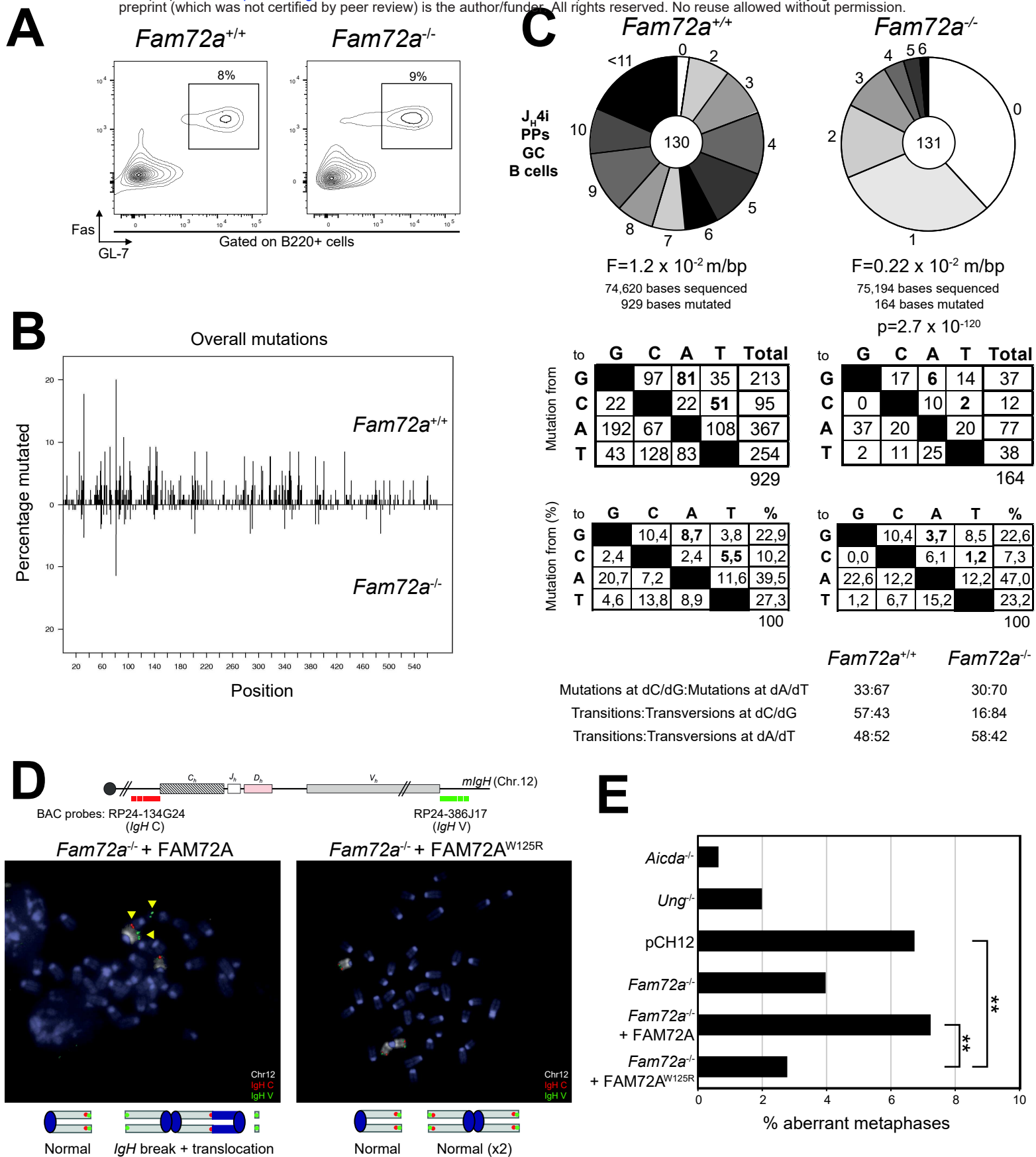


Figure 4. Deficiency in *Fam72a* shifts the balance toward error-free DNA repair. (A) Flow cytometry analysis of germinal center B cells (B220+ Fas+ GL-7+) isolated from the Peyer's patches of unimmunized *Fam72a*^{+/+} and *Fam72a*^{-/-} mice. Plots are gated on B220+ cells. (B) Distribution of mutations at the J_H4i intron (J_H4i) in *Fam72a*^{+/+} (top) and *Fam72a*^{-/-} (bottom) sequences obtained from germinal center B cells (B220+ Fas+ GL-7+) isolated from the Peyer's patches (PPs) of unimmunized *Fam72a*^{+/+} and *Fam72a*^{-/-} mice. (C) Pie charts depict the proportion of J_H4i 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*^{-/-} cells complemented with FAM72A or FAM72A^{W125R} 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.

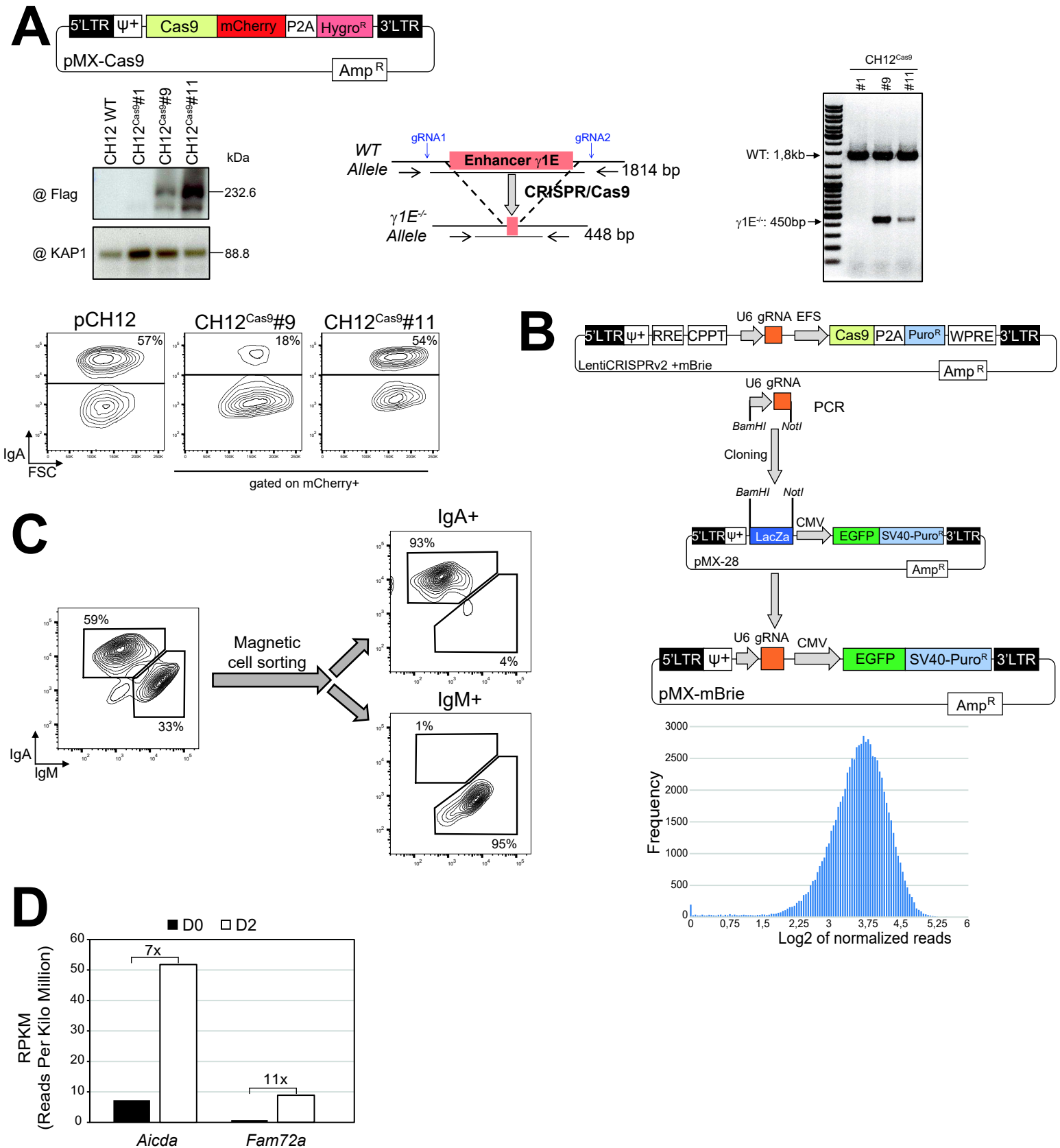


Figure S1. (A) Generation and validation of Cas9-expressing (CH12Cas9) cells. Clones were verified by Western blot for Cas9 expression, by PCR for Cas9-induced deletion of the enhancer $\gamma 1$ and by flow cytometry for CSR. Clone #11 was chosen for the screen. **(B)** The U6-gRNA sequence was amplified 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⁺ and IgA⁺ 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.

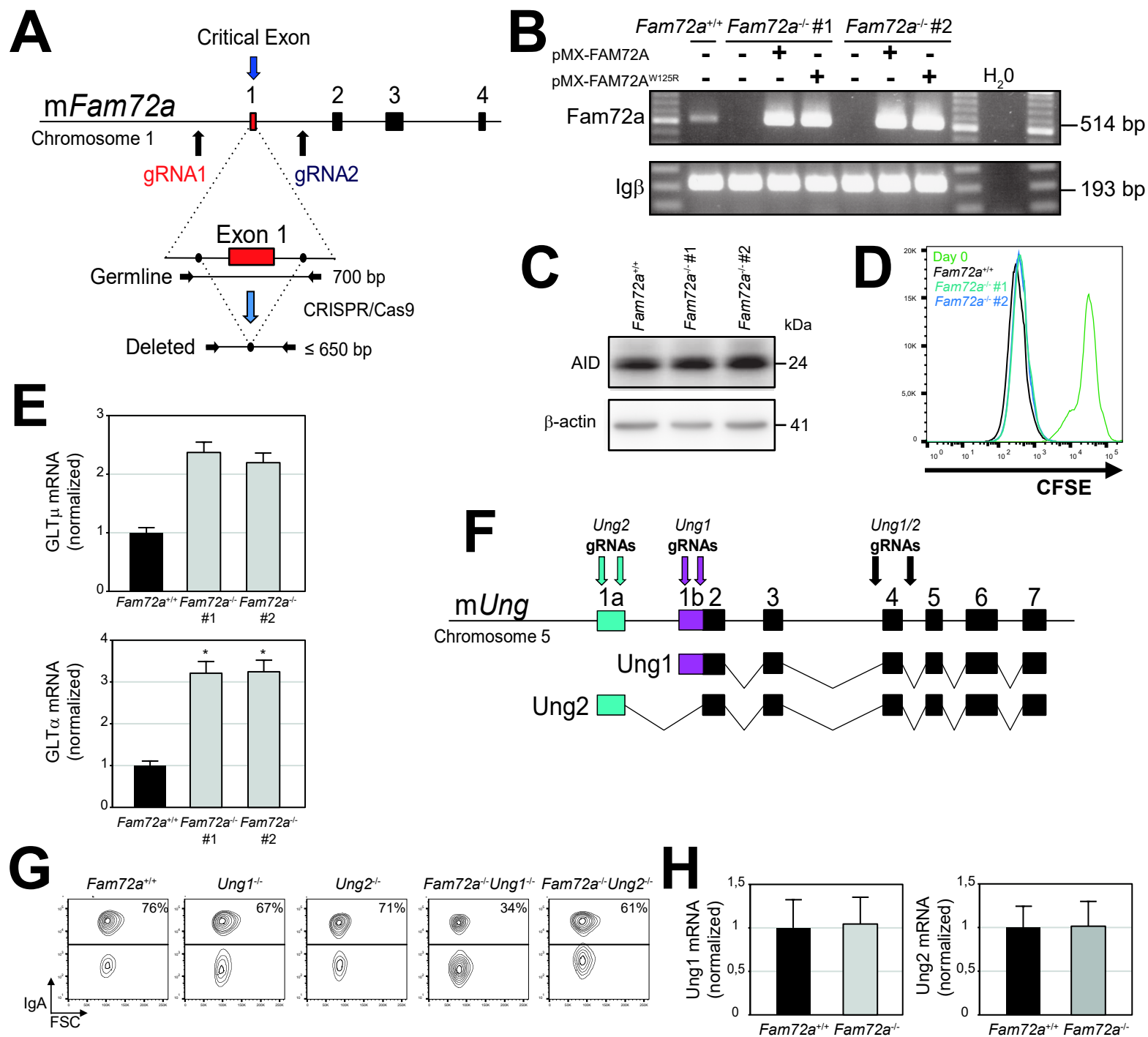


Figure S2. (A) Schematic representation of the murine *Fam72a* locus and location of the gRNAs used to generate *Fam72a*^{-/-} clones using CRISPR/Cas9-HF1. (B) RT-PCR analysis for *Fam72a* and *Igβ* in *Fam72a*^{+/+} and *Fam72a*^{-/-} CH12 cells transduced (or not) with a retrovirus expressing FAM72A, FAM72A^{W125R} or FAM72A^{W125A}. (C) Western blot analysis for AID and β -actin in *Fam72a*^{+/+} and two independent *Fam72a*^{-/-} 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*^{+/+} and two independent *Fam72a*^{-/-} CH12 cell clones cultured for 3 days with TGF- β , IL-4 and anti-CD40 antibody. (E) RT-qPCR analysis for *GLT μ* and *GLT α* in *Fam72a*^{+/+} and *Fam72a*^{-/-} 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*^{+/+} 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*^{-/-}, *Ung2*^{-/-}, *Ung*^{-/-}, *Fam72a*^{-/-} *Ung1*^{-/-} and *Fam72a*^{-/-} *Ung2*^{-/-} CH12 cell clones using CRISPR/Cas9-HF1. (G) Flow cytometry analysis of *IgA* expression in additional *Ung1*^{-/-}, *Ung2*^{-/-}, *Fam72a*^{-/-} *Ung1*^{-/-} and *Fam72a*^{-/-} *Ung2*^{-/-} 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*^{+/+} and *Fam72a*^{-/-} 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*^{+/+} controls.

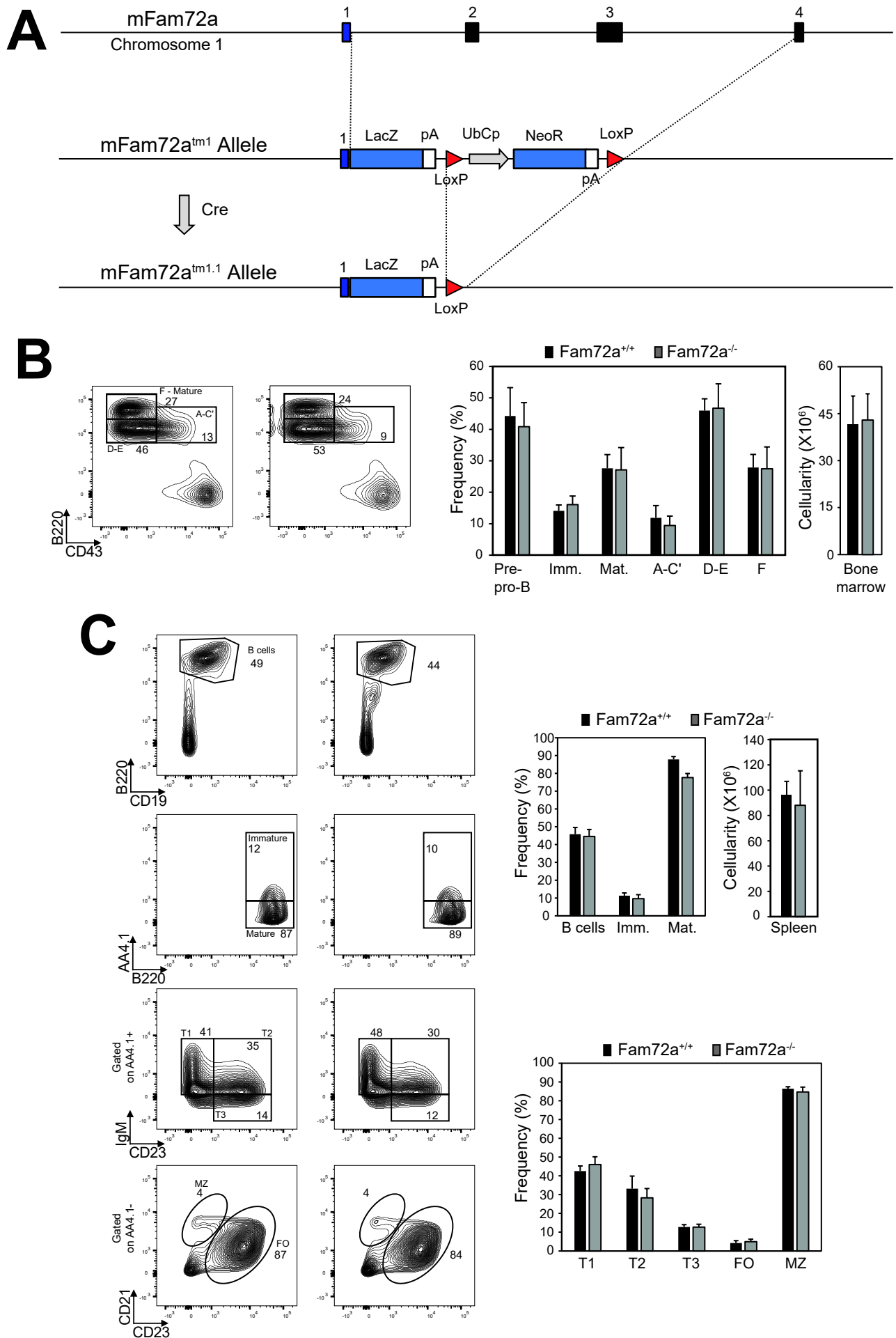
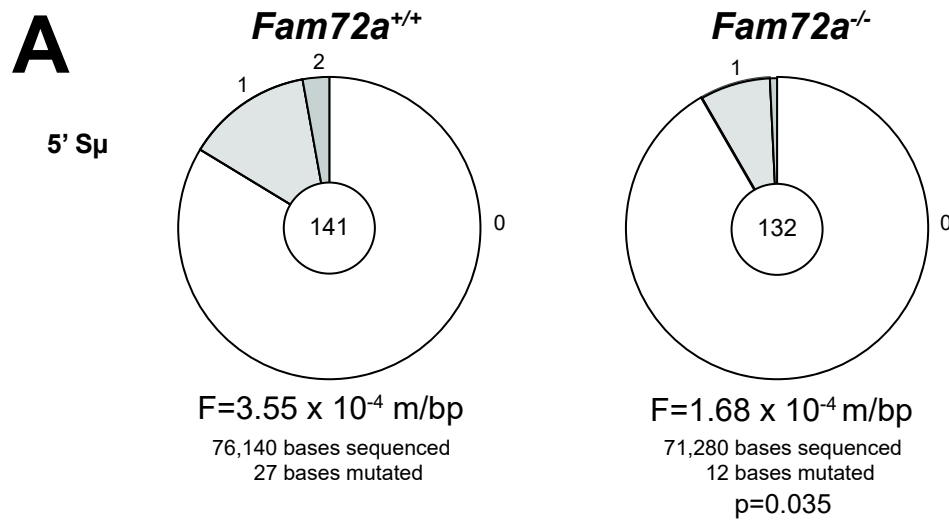


Figure S3. (A) Schematic representation of the murine *Fam72a* locus and strategy for the generation of *Fam72a*^{-/-} mouse model. Flow cytometry and cellular analysis of B cell populations in the bone marrow **(B)** or in the spleen **(C)** from *Fam72a*^{+/+} and *Fam72a*^{-/-} mice, using the indicated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



B

ALL MUTATIONS	WT	<i>Fam72a</i> ^{-/-}	SITES	bp WT	bp <i>Fam72a</i> ^{-/-}	Frequency WT	Frequency <i>Fam72a</i> ^{-/-}	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 μ mutation analysis in *Fam72a*^{+/+} and *Fam72a*^{-/-} CH12 cells cultured for 3 days with TGF- β , IL-4 and anti-CD40 antibody. Pie charts depict the proportion of S μ 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_H4i 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> ^{-/-}	1	2	0	2	248	250	0,8
	2	1	0	1	219	220	0,5
	TOTAL	3	0	3	467	470	0,6
<i>Ung</i> ^{-/-}	1	4	0	4	235	239	1,7
	2	4	1	5	209	213	2,3
	TOTAL	8	1	9	444	452	2,0
pCH12	1	11	6	16	233	249	6,4
	2	17	2	19	252	271	7,0
	TOTAL	28	8	35	485	520	6,7
<i>Fam72a</i> ^{-/-}	1	6	4	10	242	252	4,0
	TOTAL	6	4	10	242	252	4,0
<i>Fam72a</i> ^{-/-} +FAM72A	1	10	3	13	241	254	5,1
	2	15	0	15	118	133	11,3
	TOTAL	25	3	28	359	387	7,2
<i>Fam72a</i> ^{-/-} +FAM72A ^{W125R}	1	3	0	3	157	160	1,9
	2	7	1	8	228	236	3,4
	TOTAL	10	1	11	385	396	2,8

Statistical analysis (Fisher exact test)			p value	Level
<i>Fam72a</i> ^{+/+}	vs	<i>Fam72a</i> ^{-/-}	0.14188	ns
<i>Fam72a</i> ^{+/+}	vs	<i>Fam72a</i> ^{-/-} + FAM72A	0.79264	ns
<i>Fam72a</i> ^{+/+}	vs	<i>Fam72a</i> ^{-/-} + FAM72A ^{W125R}	0.00873	**
<i>Fam72a</i> ^{-/-}	vs	<i>Fam72a</i> ^{-/-} + FAM72A	0.12234	ns
<i>Fam72a</i> ^{-/-}	vs	<i>Fam72a</i> ^{-/-} + FAM72A ^{W125R}	0.49575	ns
<i>Fam72a</i> ^{-/-} + FAM72A	vs	<i>Fam72a</i> ^{-/-} + FAM72A ^{W125R}	0.00481	**

Table S2. Primers and gRNAs.

Primers	Orientation	Sequence	UPL
Transcripts			
Ig β	Fwd	TGGTGCTGTCTTCCATGC	18
	Rev	TTGCTGGTACCGGCTCAC	
I μ -C μ	Fwd	CCCAGACCTGGGAATGTATG	29
	Rev	GGAAGACATTTGGGAAGGACT	
I μ -C μ 2	Fwd	ACCTGGGAATGTATGGTTGTGGCTT	
	Rev	TCTGAACCTTCAAGGATGCTCTTG	
I γ 3-C γ 3	Fwd	GCAGAAATCTGCAGGACTAACA	71
	Rev	ACCGAGGATCCAGATGTGTC	
I γ 2b-C γ 2b	Fwd	TGGGCCTTCCAGACCTAAT	88
	Rev	GGGCTGATCTGTCAACTCCT	
I γ 2a-C γ 2a	Fwd	CAGCCTGGGATCAAGCAG	109
	Rev	TGGGGCTGTTGTTTTGGT	
I γ 1-C γ 1	Fwd	GGCCCTTCCAGATCTTTGAG	
	Rev	ATGGAGTTAGTTTGGGCAGCA	
I α -C α	Fwd	GGAGACTCCCAGGCTAGACA	27
	Rev	CGGAAGGGAAGTAATCGTGA	
Ung1	Fwd	CTGCTCGGCTGGACCAT	
	Rev	GCGCCAACCGCAAAGAC	
Ung2	Fwd	AGTGGCGGCCGAGATC	
	Rev	CCACCCGGCCTTCTTG	
Sub-cloning mBrie library			
5'-BamHI site	Fwd	CGAATGCATCTAGATATCGGATCCCTTCCCATGATTCCTTCATATTTGC	
3'-NotI site	Rev	GCAGGCCTCTGCAGTCGACGGGCCCGCGCCGCAAAAAAGCACCGACTC GGTGCCAC	
High-throughput sequencing			
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	CAAGCAGAAGACGGCATAACGAGATGCCTAAGTGACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
	4	CAAGCAGAAGACGGCATAACGAGATTGGTCAGTGACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
	5	CAAGCAGAAGACGGCATAACGAGATCACTGTGTGACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
	6	CAAGCAGAAGACGGCATAACGAGATATTGGCGTGACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
	7	CAAGCAGAAGACGGCATAACGAGATGATCTGGTGACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
	8	CAAGCAGAAGACGGCATAACGAGATTCAAGTGTGACTGGAGTTCAGACGT GTGCTCTTCCGATCAATGGAAAGTCCCTATTGGCGTTAC	
Genotyping of CH12 KO clones			
<i>Fam72a</i> ^{-/-}	Fwd	CGGGGTCGGAGCTGAAGAGTAGAAG	
	Rev	TCAACATAAGGGATGCACAGACACATG	
<i>Ung1</i> ^{-/-}	Fwd	CAATCAGAAGCGGCAAGGGGCGGGAGGTG	
	Rev	CTTCCCGAACTCCCCGCACAGC	
<i>Ung2</i> ^{-/-}	Fwd	CTCAGCCCCTCCACACATG	
	Rev	CTAAGGTGTCCCAGTCTGTCCG	
<i>Ung</i> ^{-/-}	Fwd	GAGGCCGGATGTGGGGTGGGTGAGAC	
	Rev	TGAAACAGGAGCGGCCAAGGATAACAG	
Sμ mutations			
Sμ	Fwd	CGAATGCATCTAGATATCGGATCCCAATGGATACCTCAGTGGTTTTTAATG GTGGGTTTA	
	Rev	GCAGGCCTCTGCAGTCGACGGGCCAGCGGCCAGCTCATTCCAGTTCAT TACAG	
SHM			
VH588/FR3	Fwd	CGAATGCATCTAGATATCGGATCCCGCCTGACATCTGAGGACTCTGC	
J _H 4 intron	Rev	GCAGGCCTCTGCAGTCGACGGGCCGACTTTTGCAGGCTCCACCAGACC	
gRNAs			
<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')	TGCGTTGGCGCGGAGAGCG	
<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	

II- Résultats supplémentaires :

1. Reconstitution de l'expression de mFam72a et mFam72^{W125R} dans les cellules CH12 Fam72a^{-/-}

Afin de montrer que le défaut de CI décrit dans les cellules CH12 Fam72a^{-/-} 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^{W125R} 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^{-/-} à l'aide de rétrovirus exprimant Flag-mFam72a ou Flag-mFam72a^{W125R}. Ces cellules ont ensuite été sélectionnées puis stimulées pendant 72h avec du TGFβ, 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% avec une absence de reconstitution pour la version totale Flag-mFam72a (27 à 30%) ainsi que Flag-mFam72a^{W125R} (25 à 32%) comparée aux cellules Fam72a^{-/-} 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^{W125R} 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^{W125R} 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^{W125R} comme décrit dans le papier.

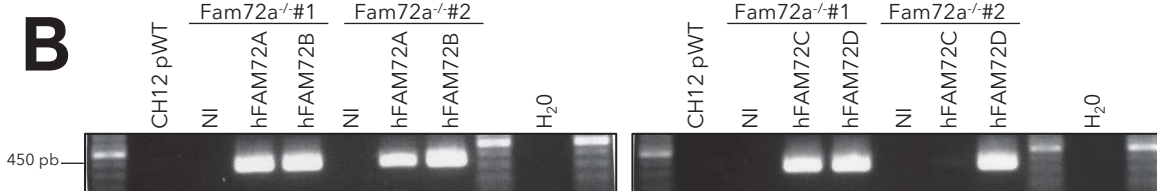
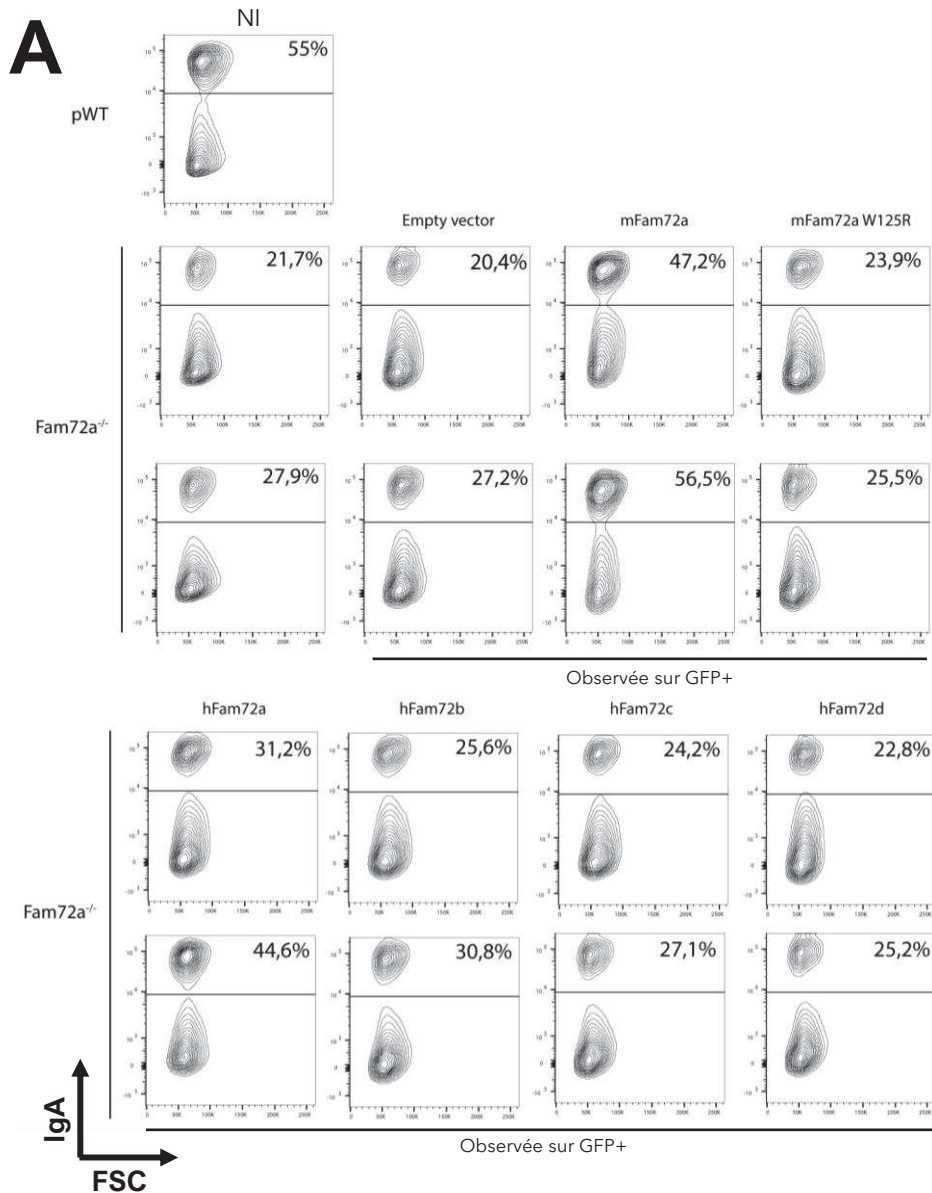


Figure 27: Reconstitution des cellules Fam72a^{-/-} avec les formes souris, mFam72a et mFam72a^{W125R} et les formes humaines hFAM72A-D. **A.** Les lignées clonales CH12 Fam72a^{-/-} ont été infectées à l'aide de rétrovirus exprimant la version totale de mFam72a, le mutant mFam72a^{W125R} et les versions 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^{-/-} non infectées (NI) ont été ajoutées comme contrôle lors de la CI. Dans les lignées clonales CH12 Fam72a^{-/-}, 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

L'expression de mFam72a et mFam72a^{W125R} 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^{-/-}

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 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^{-/-} à l'aide de rétrovirus exprimant les différentes formes hFAM72A-D ainsi que les versions mFam72a et mFam72a^{W125R} en contrôle. Ces cellules ont ensuite été sélectionnées puis stimulées pendant 72h avec du TGFβ, 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^{-/-} #2 qui est due à un problème d'extraction d'ARN. Nous en concluons qu'il y a une

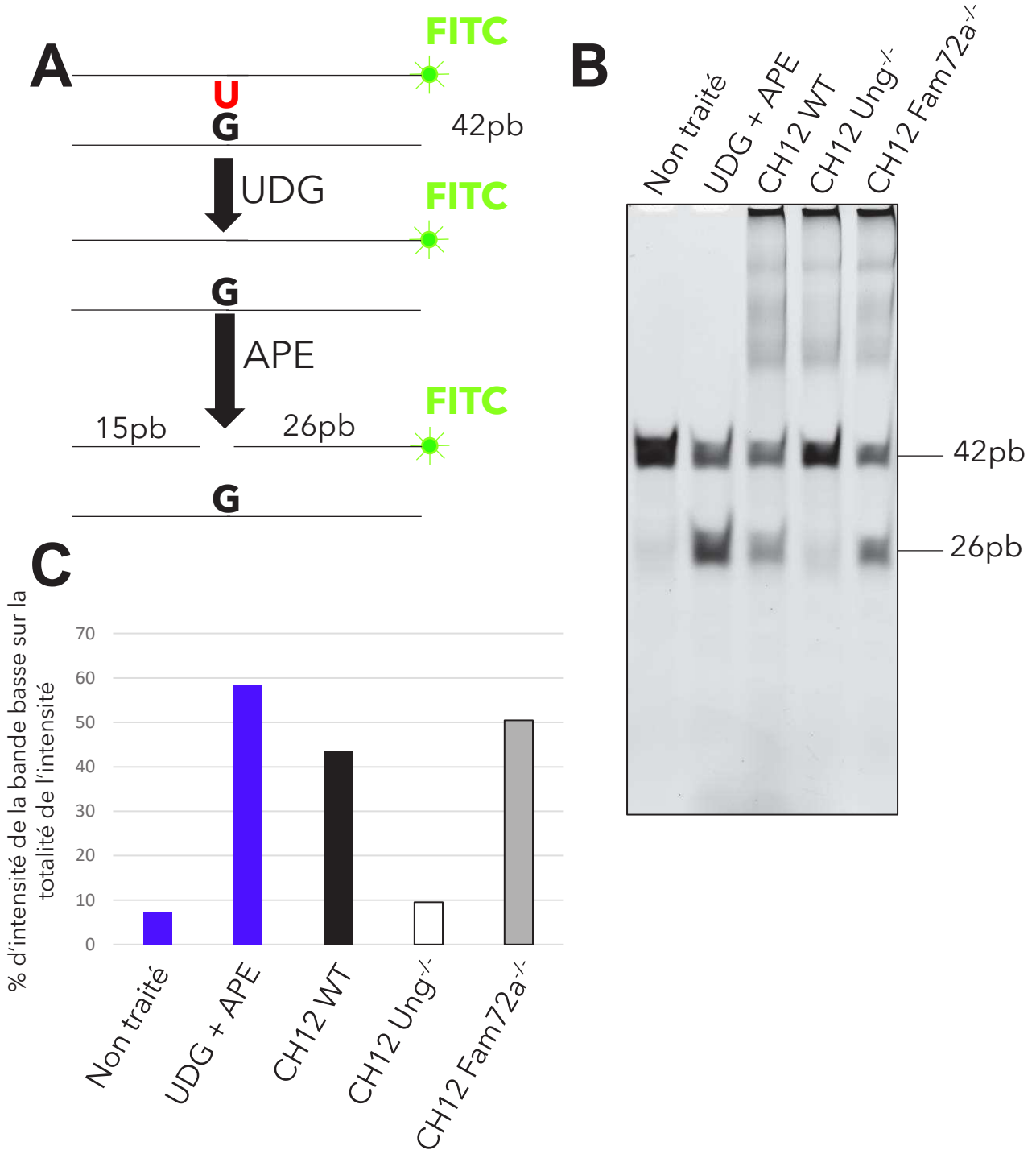


Figure 28: Test d'activité d'Ung en absence de Fam72a. **A.** Schéma explicatif du test d'activité d'UNG. Deux oligonucléotides sont hybridés afin de générer un mésappariement U:G (42 pb) qui peut être reconnu par UNG pour créer un site abasique pris en charge par APE qui va générer une cassure simple 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.

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., 2016; 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^{-/-} et Fam72a^{-/-} qui ont été stimulées avec du TGFβ, 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^{-/-} 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 glycosylation d'Ung abondante (43,6%). En revanche, l'extrait de cellules Fam72a^{-/-} 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

Durant les mécanismes de diversification des anticorps, le détournement de la réparation fidèle des voies du BER et du MMR permet l'apparition de mutations et de DSBs nécessaires lors de l'HMS et de la CI respectivement (Methot and Di Noia, 2017). Même si des facteurs principaux impliqués dans cette réparation non-fidèle ont été identifiés, les mécanismes conduisant au choix entre une réparation fidèle ou mutagène restent aujourd'hui encore mal compris. Notre étude nous a permis l'identification de Fam72a en tant que facteur faisant pencher la balance en faveur d'une réparation mutagène dans les loci Ig. Fam72a est alors requis pour réaliser efficacement la CI et l'HMS. Nos données montrent que l'interaction entre Fam72a et Ung2 est nécessaire pour cet effet, et que Fam72a régule directement la quantité d'Ung2 dans les cellules B activées. Effectivement, nous avons montré que l'expression de Fam72a semble être négativement corrélée à la quantité d'Ung2 dans les cellules B. Également, la suppression de Fam72a mène à un défaut important de la commutation isotypique et une baisse de la fréquence de mutations dans les loci Ig. Par conséquent, notre hypothèse est que Fam72a entraîne la dégradation d'Ung2 pour promouvoir une réparation non-fidèle requise pour accomplir efficacement l'HMS et la CI.

1. Ung2 et la CI : une histoire de quantité ?

1.1. Une trop grande quantité d'Ung2 porte-t-elle atteinte à la CI ?

A première vue, il peut être difficile d'expliquer en quoi une diminution de la quantité d'Ung2 favorise l'efficacité de CI. En effet, on pourrait émettre l'hypothèse qu'avoir plus d'Ung2 permettrait 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, qui sont nécessaires à la CI. Cette hypothèse est incompatible avec ce que nous observons dans nos expériences, où une surexpression d'Ung2 par infection rétrovirale inhibe la CI dans les CH12 (Figure 3D du manuscrit). Cependant, il est à

noter que nos résultats vont à l'encontre de ce qui a été observé dans de précédentes 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^{-/-} (Begum et al., 2004; Di Noia et al., 2007a).

L'hypothèse selon laquelle l'infection 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 est peu probable. Effectivement, 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^{+/+} 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 infectées et les cellules Ung^{+/+}.

Enfin, il est possible que l'effet d'une baisse d'efficacité de CI en cas de surexpression d'Ung2 soit spécifique à la lignée CH12, qui posséderait alors des particularités que ne possèdent pas les cellules B primaires pour expliquer cet effet. Par conséquent, nous prévoyons ainsi de surexprimer 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 ?

Au contraire, 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, celle-ci 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). Cependant, après stimulation de ces cellules, une perte d'environ 50% de l'efficacité de CI est observée dans les cellules exprimant Vpr, ainsi qu'une augmentation de la quantité d'uraciles présents dans l'ADN (Eldin et al., 2020). De plus, la baisse de l'efficacité de CI observée est dépendante de la dose de virus utilisée pour infecter les cellules, et donc dépendante de la quantité de Vpr exprimée. Enfin, cette étude confirme ce même effet dans des cellules B primaires,

excluant la possibilité d'un mécanisme propre aux CH12. Contrairement à ce que montre cette étude, nous n'observons pas de relation dose-effet entre une faible quantité d'Ung2 et une diminution d'efficacité de la CI. Effectivement, nous montrons que la surexpression de Fam72a par infection rétrovirale de cellules Fam72a^{-/-} provoque une plus forte dégradation d'Ung2, mais permet une restauration complète de la déficience en CI (Figures 2E et 3B du manuscrit). En conclusion, cela indique que soit Vpr a un rôle inhibiteur de la CI indépendamment du fait qu'elle dégrade Ung2, soit c'est Fam72a qui favorise une CI efficace de manière indépendante de son action de dégradation d'Ung2.

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

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

Parmi les autres actions que Fam72a pourrait jouer sur Ung2, il est possible qu'elle régule l'interaction entre Ung2 et Rev1, ADN polymérase translésionnelle impliquée dans le ciblage d'Ung2 aux régions S et dans l'augmentation de son activité catalytique (Zan et al., 2012).

Une autre hypothèse est que Fam72a peut réguler la présence et/ou l'activité d'Ung2 en fonction du cycle cellulaire. Effectivement, 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 celle-ci (Le and Maizels, 2015; Sharbeen et al., 2012). Afin d'aboutir à une réparation non-fidèle dans les régions lors de l'HMS, cette réparation en G1 est essentielle (Sharbeen et al., 2012), ce qui peut être aussi le cas dans les régions S afin de promouvoir la réparation des cassures par NHEJ.

L'expression et l'activité d'Ung2 sont augmentées lors de l'activation des lymphocytes B, autant chez l'humain (Imai et al., 2003c) 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 obscurs.

Il a cependant été démontré que la version humaine d'UNG2 est régulée par un ensemble de modifications post-traductionnelles telles que des phosphorylations, des ubiquitinations et de la sumoylation (Hagen et al., 2008; Lu et al., 2004; Ma et al., 2009). Ainsi, nous pensons que Fam72a peut avoir un rôle dans la présence ou l'absence de ces modifications. En effet, Fam72a pourrait permettre à Ung2 de persévérer durant la phase G1 du cycle cellulaire et de réaliser une réparation mutagène des uraciles.

Pour répondre à cette question, il est possible d'arrêter des cellules CH12 Fam72a^{+/+} et Fam72a^{-/-} en G1 à l'aide de thymidine, et d'observer s'il existe des différences de quantité et/ou d'activité d'Ung2.

3. Et Ung1 dans tout ça ?

Les différentes expériences que nous avons conduites ont permis de confirmer une étude récente montrant qu'Ung1 peut soutenir la CI à des niveaux normaux en absence d'Ung2. Ceci mène à vouloir savoir la contribution d'Ung1 dans les mécanismes de CI dans des conditions normales. Sa suppression ne provoque pas de diminution d'efficacité de CI dans les cellules CH12, cela ne veut pas dire qu'elle n'est pas impliquée dans le mécanisme normal de CI, puisqu'il est probable qu'Ung2 soit capable de compenser la perte d'Ung1.

Il a été montré que le domaine d'interaction de Fam72a dans Ung2 est dans la partie N-terminale qui la différencie d'Ung1, il est donc peu probable que Fam72a régule l'activité d'Ung1. Si Ung1 est capable à elle seule d'initier une CI efficace, pour quelle raison l'expression de Fam72a et son interaction avec Ung2 sont-elles nécessaires pour obtenir une CI avec la même efficacité ? Il a été montré qu'une faible proportion d'Ung1 est présente dans le noyau à un moment donné (Sarno et al., 2019), et ceci pourrait être la raison pour laquelle Ung1 soutient des niveaux normaux de CI. Cette idée est cohérente avec le fait qu'une surexpression d'Ung1 entraîne également une diminution de l'efficacité de CI (FIGURE 3D du manuscrit). Une fois de plus, cela renforce le modèle qui affirme que c'est la quantité d'Ung1 ou

d'Ung2 qui dicte l'efficacité de la CI. Ung1 et Ung2 pourraient avoir un fonctionnement similaire, et l'augmentation du niveau de l'un ou de l'autre aboutirait à la diminution de l'efficacité de CI.

4. Fam72a et Ung2 dans l'HMS

De manière semblable à ce que nous observons dans les cellules Fam72a^{-/-}, il a été observé que la surexpression d'Ung2 mène à une baisse relativement importante du nombre de transitions à G:C et des mutations à A:T, ainsi qu'une réduction plus faible des transversions en G:C (Pérez-Durán et al., 2012; Thientosapol et al., 2017). Ces résultats renforcent 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 en observant une baisse du nombre de mutations à A:T. Cette réduction pourrait être due soit par une inhibition de la réparation des uraciles par le MMR seul, soit par 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). Néanmoins, il reste encore à expliquer en quoi une quantité supérieure d'Ung2 empêche cette collaboration ainsi que pourquoi cela entraîne une réparation plus fidèle. Ung2 n'est effectivement pas le seul facteur 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). En revanche, une étude a montré que l'ADN polymérase β est le facteur limitant dans le BER classique (Nakamura and Swenberg, 1999). De plus, une nouvelle étude met en cause une réduction *in vivo* du niveau d'ADN polymérase β dans des cellules B pour permettre la mutagenèse à A:T pendant l'HMS (Bahjat et al., 2020), montrant ainsi que la quantité d'Ung2 n'est pas le seul paramètre qui peut faire pencher la réparation du côté mutagène ou fidèle. Il est alors probable qu'une quantité supplémentaire d'Ung2 change la manière dont l'ADN polymérase β est recrutée aux loci Ig.

5. Perspectives

Afin de confirmer les résultats obtenus et d'apporter une meilleure compréhension de ceux-ci, plusieurs expériences restent encore à être réalisées. 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 ?

Suite à l'obtention de résultats contradictoires de ceux obtenus par d'autres groupes, nous devons confirmer nos observations et valider nos hypothèses en utilisant plusieurs techniques. Tout d'abord, nous souhaitons 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^{Flag-Biotag} Fam72a^{-/-}). 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^{Flag-Biotag} Fam72a^{+/+} et Ung2^{Flag-Biotag} Fam72a^{-/-}) 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^{+/+} par rapport à des cellules Fam72a^{-/-}. 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 ?

Ne possédant pas d'anticorps spécifique anti-Ung2, avec les 2 lignées KI précédemment citées, nous pourrions effectuer 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. Enfin, afin de mettre en évidence les partenaires d'Ung2 spécifiques à la présence ou non de Fam72a, nous envisageons d'effectuer une co-immunoprécipitation (co-IP), en utilisant les deux lignées KI Ung2^{Flag-Biotag} Fam72a^{+/+} et Ung2^{Flag-Biotag} Fam72a^{-/-}, couplée à la spectrométrie de

masse. Cette expérience nous permettra de savoir si Fam72a est impliquée dans la régulation de l'interaction d'Ung2 avec d'autres partenaires protéiques.

De plus, nous souhaitons également regarder si la déficience de CI causée par la surexpression d'Ung2 peut être renversée en abaissant de manière artificielle le niveau d'expression d'Ung2. Pour cela, plusieurs alternatives existent, comme, par exemple, l'expression d'Ung2 à partir d'un promoteur inductible à la doxycycline (système Tet-On) ou bien le couplage de la surexpression d'Ung2 avec l'utilisation d'un système de dégradation inductible à l'auxine (Nishimura et al., 2009). Alors, une expression faible d'Ung2 dans des cellules Fam72a^{-/-} nous permettra de savoir si le rôle de Fam72a est réduit à limiter l'expression d'Ung2 ou pas.

5.3. Est-ce que le recrutement d'AID à S μ est altéré en absence de Fam72a ?

L'observation d'une baisse de la fréquence de mutations aux loci Ig en absence de Fam72a nous mène à vouloir 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 μ en présence et en absence de Fam72a. De plus, des contrôles pWT et AID^{-/-} seront ajoutés.

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

Des données préliminaires présentées dans les résultats supplémentaires nous permettent d'évaluer que l'activité d'Ung2 est peu affectée par la présence de Fam72a (Figure 28).

Néanmoins, il est 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 (Zan et al., 2012).

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

Une autre manière de confirmer que la surexpression d'Ung2 est couplée à une réparation plus efficace des uraciles dans l'ADN est la quantification globale de la quantité d'uraciles dans l'ADN en absence de Fam72a. Toutefois, il est possible que nous rencontrions des problèmes techniques car 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., 2019)).

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

Il est probable que Fam72a est à l'origine de la dégradation active d'Ung2 par le protéasome. Pour le confirmer, nous souhaitons réaliser la mise en culture des lignées CH12 pWT, Ung^{-/-}, Fam72a^{-/-}, Fam72a^{-/-} reconstituées avec mFam72a et mFam72a^{W125R} en présence ou en absence de l'inhibiteur du protéasome MG132. Dans le but de quantifier une possible différence entre les conditions en présence ou en absence de MG132, un western blot pour Ung2 sera effectué.

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

La CI nécessite la génération de DSBs dans les régions S pour pouvoir être réalisée (Petersen et al., 2001). Comme vu précédemment dans l'introduction, ces DSBs sont le fruit d'une réparation mutagène initiée majoritairement par l'action d'Ung. Ainsi, une quantité accrue d'Ung2 pourrait faire pencher la balance vers une réparation fidèle, ce qui entrainerait une réduction de la génération de DSBs dans les régions S et donc une perte d'efficacité de la CI.

Afin d'étudier ce point, une quantification du nombre de DSBs induites par AID dans les régions S des lignées cellulaires CH12 Fam72a^{+/+}, AID^{-/-}, Fam72a^{-/-} et les cellules CH12 Fam72a^{-/-} reconstituées avec mFam72a et mFam72a^{W125R} stimulées par la technique 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 favorisant 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^{-/-} comme contrôles positifs.

5.9. Est-ce que la transcription des régions S est affectée dans les cellules B primaires Fam72a^{-/-} ?

D'après les résultats obtenus dans les cellules CH12 Fam72a^{-/-}, la transcription des régions S acceptrice S μ et donneuse S α 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 μ , S γ 3, S γ 2b, S γ 2a et S α , ainsi que pour Fam72a dans les cellules B primaires Fam72a^{-/-} et Fam72a^{+/+}.

6. Conclusion

En résumé, une caractérisation plus profonde de ce facteur prometteur mènera à une meilleure compréhension des mécanismes impliqués lors de la réparation non fidèle aux loci Ig lors de la diversification des anticorps. De plus, celle-ci permettra de connaître par quels mécanismes Ung2 favorise une réparation fidèle ou mutagène. La compréhension de ces mécanismes est importante pour comprendre le rôle protecteur d'Ung2 au niveau de certains loci ciblés par l'action d'AID.

CHAPITRE 3 :

Etude de l'impact de Zbtb24 dans la commutation isotypique.

I- La perte de ZBTB24 compromet le NHEJ et la CI chez les patients atteints d'ICF

ZBTB24 (zinc-finger-and BTB (bric-a-brac, tramtrack, broad complex) domain containing 24) est un membre de la famille des protéines ZBTB, dont plusieurs d'entre elles jouent un rôle dans le développement des lymphocytes et la lymphomagenèse, principalement par une répression transcriptionnelle (Lee and Maeda, 2012).

La présence de mutations dans le gène ZBTB24 chez l'homme provoque le syndrome de l'immunodéficience combinée avec une instabilité centromérique et une dysmorphie faciale de type 2, appelée ICF2 (Immunodeficiency with centromeric instability and facial anomalies) (Chouery et al., 2012; de Greef et al., 2011; Nitta et al., 2013). Le syndrome ICF est un trouble autosomique récessif rare caractérisé par des infections respiratoires et gastro-intestinales récurrentes et souvent mortelles associées à une hypogammaglobulinémie et une immunodéficience. Trois autres gènes sont actuellement connus pour provoquer le syndrome de l'ICF en présence de mutations, dont l'ADN méthyltransférase (DNMT3B) (ICF1) (Hansen et al., 1999; Xu et al., 1999), le gène CDCA7 (cell division cycle associated 7) (ICF3) (Thijssen et al., 2015) et le gène HELLS (Helicase, lymphoid-specific) (ICF4) (Thijssen et al., 2015). Une classe additionnelle, appelée ICFX, comprend des cas où la mutation sous-jacente reste non identifiée (Thijssen et al., 2015).

Environ 30% des cas d'ICF comportent des mutations dans le gène ZBTB24. Cependant, malgré la connaissance des mutations dans le gène ZBTB24 associées au syndrome de l'ICF, le mécanisme menant à l'immunodéficience est encore inconnu. De plus, il est important de noter que le nombre de cellules B circulant chez les patients atteints d'ICF est normal mais un manque de cellules B mémoires commutées et une proportion augmentée de cellules B immatures ont été reportés (Blanco-Betancourt, 2004), suggérant la présence d'un défaut dans le stade final de la différenciation des cellules B. De plus, les similitudes entre l'immunodéficience chez les patients atteints d'ICF et les individus présentant des mutations dans les

gènes impliqués dans le NHEJ suggèrent que ZBTB24 est impliqué dans la CI et plus particulièrement dans le NHEJ.

Le laboratoire du Dr Haico van Attikum a établi une collaboration avec notre laboratoire afin que nous évaluions l'impact de Zbtb24 dans la commutation isotypique.

Afin d'évaluer l'implication de Zbtb24 dans la commutation isotypique, nous avons inactivé Zbtb24 en utilisant le système CRISPR/Cas9 dans la lignée cellulaire CH12, dérivée d'un lymphome B murin dans laquelle l'expression d'AID et la commutation isotypique peuvent être efficacement induites. Nous avons observé un défaut d'efficacité de CI d'environ 40% dans les différentes lignées clonales Zbtb24^{-/-} générées et montré que celui-ci n'est pas lié à un défaut d'expression d'AID et qu'il est bien dû à l'absence de Zbtb24.

ARTICLE

Loss of ZBTB24 impairs nonhomologous end-joining and class-switch recombination in patients with ICF syndrome

Angela Helfricht^{1*}, Peter E. Thijssen^{1*}, Magdalena B. Rother^{1*}, Rashmi G. Shah², Likun Du³, Sanami Takada⁴, Mélanie Rogier^{5,6,7,8}, Jacques Moritz^{5,6,7,8}, Hanna Ijspeert⁹, Chantal Stoeper¹, Monique M. van Ostaijen-ten Dam⁴, Vincent Heyer^{5,6,7,8}, Martijn S. Luijsterburg¹, Anton de Groot¹, Rianca Jak¹, Gwendolynn Grootaers¹, Jun Wang¹, Pooja Rao¹⁰, Alfred C.O. Vertegaal¹¹, Maarten J.D. van Tol⁴, Qiang Pan-Hammarström³, Bernardo Reina-San-Martin^{5,6,7,8}, Girish M. Shah², Mirjam van der Burg⁴, Silvère M. van der Maarel¹, and Haico van Attikum¹

The autosomal recessive immunodeficiency, centromeric instability, and facial anomalies (ICF) syndrome is a genetically heterogeneous disorder. Despite the identification of the underlying gene defects, it is unclear how mutations in any of the four known ICF genes cause a primary immunodeficiency. Here we demonstrate that loss of ZBTB24 in B cells from mice and ICF2 patients affects nonhomologous end-joining (NHEJ) during immunoglobulin class-switch recombination and consequently impairs immunoglobulin production and isotype balance. Mechanistically, we found that ZBTB24 associates with poly(ADP-ribose) polymerase 1 (PARP1) and stimulates its auto-poly(ADP-ribosylation). The zinc-finger in ZBTB24 binds PARP1-associated poly(ADP-ribose) chains and mediates the PARP1-dependent recruitment of ZBTB24 to DNA breaks. Moreover, through its association with poly(ADP-ribose) chains, ZBTB24 protects them from degradation by poly(ADP-ribose) glycohydrolase (PARG). This facilitates the poly(ADP-ribose)-dependent assembly of the LIG4/XRCC4 complex at DNA breaks, thereby promoting error-free NHEJ. Thus, we uncover ZBTB24 as a regulator of PARP1-dependent NHEJ and class-switch recombination, providing a molecular basis for the immunodeficiency in ICF2 syndrome.

Introduction

Immunodeficiency with centromeric instability and facial anomalies (ICF) syndrome (OMIM 242860; 614069) is a rare autosomal recessive disorder characterized by a triad of phenotypes (Hagleitner et al., 2008; Weemaes et al., 2013). Patients suffer from a variable immunodeficiency, mainly characterized by hypo- or agammaglobulinemia in the presence of B cells, resulting in recurrent and often fatal respiratory and gastrointestinal infections. Furthermore, patients often present with a distinct set of facial anomalies, including a flat nasal bridge, hypertelorism, and epicanthal folds. The cytogenetic hallmark of the disease is centromeric instability, specifically at chromosomes 1, 9, and 16, which is associated with CpG hypomethylation of the pericentromeric satellite II and III repeats.

ICF syndrome is genetically heterogeneous and can be subdivided into five different groups (ICF1-4 and ICFX) based on the

genetic defect underlying the phenotype (Thijssen et al., 2015; Weemaes et al., 2013). ICF1 patients, comprising ~50% of the total patient population, carry mutations in the de novo DNA methyltransferase 3B gene (*DNMT3B*, ICF1; Hansen et al., 1999; Xu et al., 1999). Approximately 30% of the cases have mutations in the zinc-finger and BTB (bric-a-bric, tramtrack, broad complex)-containing 24 gene (*ZBTB24*, ICF2; Chouery et al., 2012; de Greef et al., 2011; Nitta et al., 2013). Finally, mutations in the cell division cycle-associated protein 7 (*CDC47*, ICF3) or helicase, lymphoid-specific (*HELLS*, ICF4) were also reported in patients (~20% of the total patient population), leaving only a few cases genetically unaccounted for (ICFX; Thijssen et al., 2015). Remarkably, however, although the genetic defects underlying ICF syndrome have been mostly elucidated, it remains largely unclear how these defects lead

¹Department of Human Genetics, Leiden University Medical Center, Leiden, Netherlands; ²CHU de Québec Research Centre (site CHUL) and Laboratory for Skin Cancer Research and Axe Neuroscience, Université Laval, Québec, Canada; ³Department of Biosciences and Nutrition, Karolinska Institute, Solna, Sweden; ⁴Laboratory for Pediatric Immunology, Department of Pediatrics, Willem Alexander Children's Hospital, Leiden University Medical Center, Leiden, Netherlands; ⁵Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch, France; ⁶Centre National de la Recherche Scientifique, UMR7104, Illkirch, France; ⁷Institut National de la Santé et de la Recherche Médicale, U1258, Illkirch, France; ⁸Université de Strasbourg, Illkirch, France; ⁹Department of Immunology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, Netherlands; ¹⁰ServiceXS B.V., Leiden, Netherlands; ¹¹Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, Netherlands.

*A. Helfricht, P.E. Thijssen, and M.B. Rother contributed equally to this paper; Correspondence to Haico van Attikum: h.van.attikum@lumc.nl; J. Wang's present address is Institutes of Biology and Medical Sciences, Soochow University, Suzhou, China.

© 2020 Helfricht et al. This article is available under a Creative Commons License (Attribution 4.0 International, as described at <https://creativecommons.org/licenses/by/4.0/>).

to ICF syndrome, in particular the characteristic life-threatening immunodeficiency.

Interestingly, the number of circulating B lymphocytes in ICF patients is normal, but a lack of switched memory B cells and an increased proportion of immature B cells have been reported (Blanco-Betancourt et al., 2004), suggesting a defect in the final stages of B cell differentiation. A key step in B cell maturation is isotype switching of Igs through class-switch recombination (CSR). Effective CSR heavily relies on the controlled formation and correct repair of DNA double-strand breaks (DSBs) induced by activation-induced (cytidine) deaminase (AID) at conserved motifs within the switch (S) regions, which are upstream from gene segments that encode distinct constant regions of antibody heavy chains (Alt et al., 2013). Upon break formation, two S regions are rejoined by nonhomologous end-joining (NHEJ), the main cellular pathway to repair DSBs (Alt et al., 2013). This leads to loss of the intervening DNA between the S regions, removal of μ and δ heavy chain constant regions, substitution by a γ , α , or ϵ constant region, and consequently a change in the class of immunoglobulins that is expressed by a B cell.

NHEJ is performed by the concerted action of the DNA-dependent protein-kinase complex (DNA-PK), comprised of the KU70/KU80 heterodimer and the DNA-PK catalytic subunit (DNA-PKcs), and the downstream effector proteins x-ray repair cross-complementing protein 4 (XRCC4), DNA ligase 4 (LIG4), and nonhomologous end-joining factor 1 (NHEJ1; Alt et al., 2013). In the absence of this canonical NHEJ (c-NHEJ) mechanism, effective CSR is significantly impaired but not absent, as DSB repair is performed by alternative NHEJ (a-NHEJ). a-NHEJ is a poorly characterized process dependent on poly(ADP-ribose) polymerase 1 (PARP1), XRCC1, and DNA ligases 1 and 3 (LIG1 and LIG3; Audebert et al., 2004; Lu et al., 2016; Paul et al., 2013). Recent studies have also revealed a role for PARP1 in c-NHEJ (Luijsterburg et al., 2016).

Mutations in NHEJ genes (e.g., DNA-PKcs and LIG4) are increasingly recognized as the primary cause of immunodeficiency (Woodbine et al., 2014). Considering the similarities between the immunodeficiency in ICF patients and individuals with defective NHEJ, the question arises whether loss of NHEJ might explain the compromised immune system in ICF patients. Here we demonstrate that ICF2 patient-derived B cells are defective in NHEJ during CSR. Mechanistically, we uncover a regulatory function for ZBTB24 in NHEJ by cooperating with PARP1 and XRCC4/LIG4 during this repair process. This provides a molecular basis for the humoral immunodeficiency in ICF2 patients.

Results

ICF2 patients display features of defective CSR

The immunodeficiency in ICF2 syndrome is characterized by a reduction or even an absence of Igs (hypo- or agammaglobulinemia) and decreased numbers of switched memory B cells, while normal levels of total B cells are observed (de Greef et al., 2011; Weemaes et al., 2013). We corroborated these findings by showing hypogammaglobulinemia in sera of four independent ICF2 patients but normal serum levels in age-matched controls

(Table S1). Moreover, we characterized peripheral blood lymphocytes by immunophenotyping and found a decrease in the number of switched memory B cells, whereas numbers of total B cells, naive B cells, and unswitched memory B cells were unaffected (Fig. 1 A). Of note, total numbers of CD4⁺ T cells, as well as naive, central memory, and CD27⁺CD28⁺ early antigen experienced CD4⁺ T cells were increased compared with age-matched controls, whereas those for CD8⁺ T cells were normal (Fig. S1, A and B).

These findings could suggest a defect in V(D)J recombination or CSR, which are processes that are critical for B cell development and ultimately define antibody production and diversification. We therefore first examined the combinatorial diversity of VDJ usage and composition of the junctional region during V(D)J recombination by sequencing Ig heavy chain gene rearrangements in B cells derived from peripheral blood mononuclear cells (PBMCs) of three ICF2 patients. The usage of V, D, and J gene segments, as well as the composition of the junctional regions, meaning the number of nucleotide deletions and insertions of nontemplated nucleotides by terminal deoxynucleotidyl transferase (TdT; N-nucleotides), in these patients resembled that of controls (Fig. S1, C-E). This suggests that ICF2 patients do not suffer from major defects in V(D)J recombination.

To examine impairment in CSR, we first stimulated PBMCs in vitro and measured the production of total IgA and IgG. For all patients analyzed, the capacity to produce IgA and IgG in vitro was significantly impaired compared with healthy controls (Fig. 1 B), which is in line with the hypo- or agammaglobulinemia and reduced numbers of switched memory B cells observed in these patients (Fig. 1 A; de Greef et al., 2011; Weemaes et al., 2013). We then performed sequencing analysis of IgG transcripts in patient-derived PBMCs and determined the relative abundance of IgG subclasses. When comparing relative abundance of IgG₁₋₄ to age-matched controls, we observed a decrease in the relative expression of IgG₁, accompanied by an increase in relative IgG₃ expression in ICF2 patients (Fig. 1 C), which is indicative of impaired CSR.

To investigate how CSR is affected in ICF2, we isolated naive B cells from two ICF2 patients and three healthy donors and stimulated these cells with CD40L, IL21, IL10, and algM for 6 d. To determine whether this treatment could initiate the CSR process, we measured AID expression and germline transcripts after 6 d of culture. In both ICF2 patients and controls, AID expression (Fig. 1 D), as well germline transcription (Fig. 1, E and F), was induced upon stimulation. However, despite the induction of AID and germline transcription, the formation of IgG⁺ and IgA⁺ B cells was severely hampered in ICF2 patients (Fig. 1 G), which indicates an impairment in the final stages of CSR involving the repair of AID-induced DSBs. In line with the reduced number of IgG⁺ and IgA⁺ B cells, the production of IgG, and to a lesser degree IgA, was also severely reduced in ICF2 patients (Fig. 1 H). Together, these data show that although CSR can be initiated normally, it cannot be completed in B cells from ICF2 patients, probably owing to impaired DSB repair.

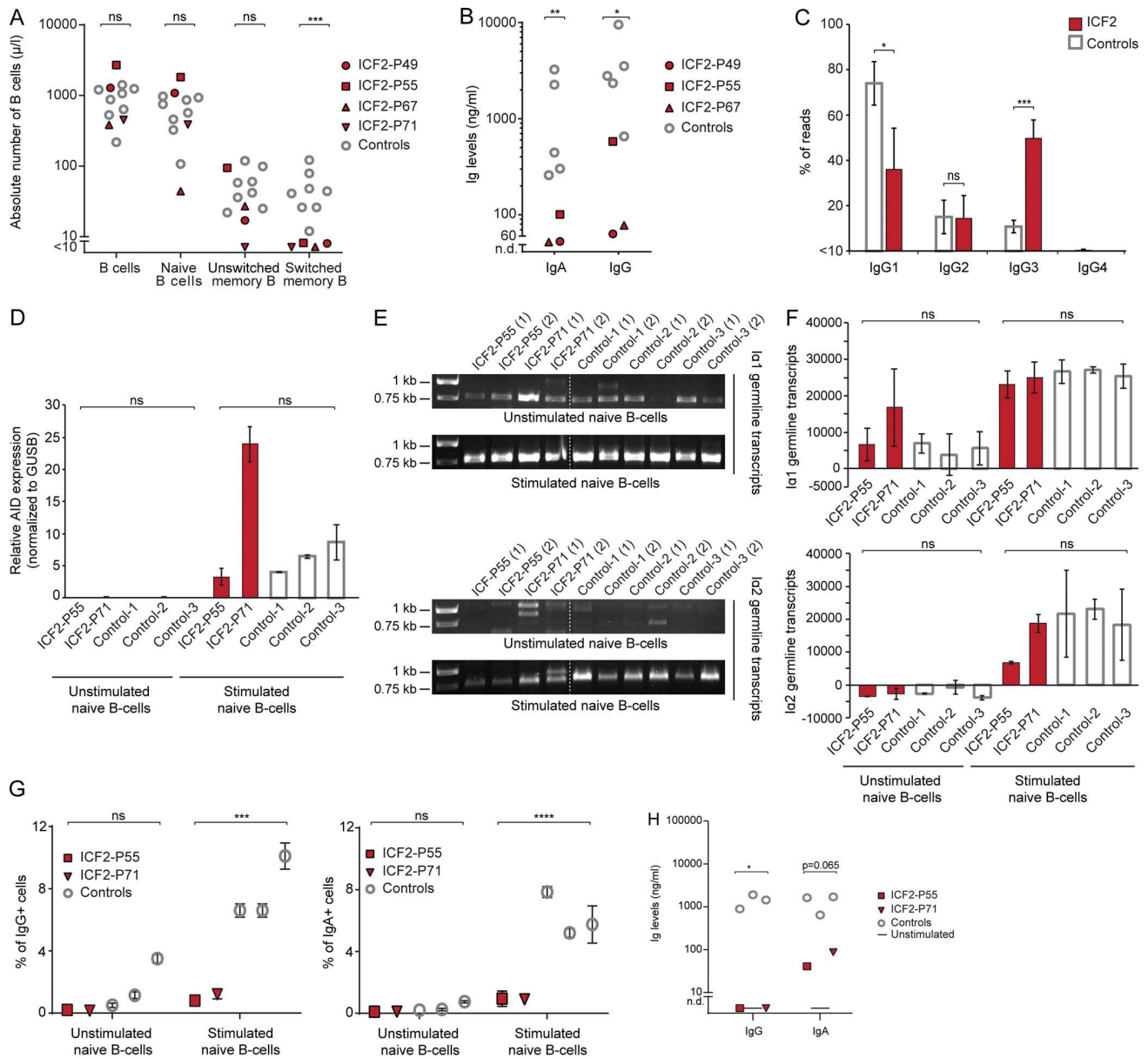


Figure 1. Defective CSR in ICF2 patients due to loss of ZBTB24-dependent NHEJ. (A) The number of cells in the indicated differentiation stages of the total peripheral blood CD19⁺/CD20⁺ B cell population was measured by flow cytometry. Naive B cells, IgM^{dull}, IgD⁺, CD27⁻; unswitched memory B cells, IgM⁺, IgD^{dull}, CD27⁺; switched memory B cells, IgM⁻, IgD⁻, CD27⁺. Four ICF2 patients and eight healthy age-matched controls (age range 0.8–4.3 yr) were analyzed. Statistical significance was calculated using Student's *t* test (***, *P* < 0.001; ns, not significant). (B) PBMCs were stimulated with CD40L, algM, CpG, and IL-21. After 7 d, IgG and IgA concentrations were determined by ELISA. Three ICF2 patients and five adult controls were analyzed. n.d., not detectable. Statistical significance was calculated using Student's *t* test (*, *P* < 0.05; **, *P* < 0.01). (C) Frequency of IgG subclass usage with unique switched *IGG* transcripts. The mean \pm SEM of three ICF2 patients and four healthy controls is shown. Statistical significance was calculated using χ^2 test (*, *P* < 0.05; ***, *P* < 0.001). (D) Naive B cells from two ICF2 patients and three healthy controls were stimulated with CD40L, algM, IL-10, and IL-21 or left unstimulated. After 6 d, relative expression of AID transcripts in unstimulated and stimulated cells was determined using RT-qPCR. The mean \pm SD of two technical replicates is shown. Statistical significance was calculated using Student's *t* test (ns, not significant). (E) As in D, except that *Ia1-Ca1* and *Ia2-Ca2* germline transcripts were amplified by PCR. PCR products were resolved on a 1% agarose gel. (F) Quantification of the intensity of the expected bands from E with ImageJ. The mean \pm SD of two technical replicates is shown. Statistical significance was calculated using Student's *t* test (ns, not significant). (G) As in D, except that the percentage of IgG⁺ and IgA⁺ CD19⁺ B cells was determined with flow cytometry. The mean \pm SD of two technical replicates is shown. Statistical significance was calculated using Student's *t* test (***, *P* < 0.001; ****, *P* < 0.0001; ns, not significant). (H) As in D, except that after 10 d of culture, IgG and IgA concentrations were determined by ELISA. The mean of two technical replicates is shown. Statistical significance was calculated using Student's *t* test (*, *P* < 0.05).

Table 1. Characterization of CSR junctions^a

Study subjects	Perfectly matched short homology				No. of junctions
	0 bp				
	Direct end-joining	Small insertions	1–6 bp	≥7 bp	
Sμ-Sα					
ICF2-deficient ^a	4 (5)* \downarrow	11 (13)* \downarrow	34 (42)	33 (40)** \uparrow	82
Lig4-deficient ^b	1 (3)	0 (0)** \downarrow	11 (37)	18 (60)** \uparrow	30
Artemis-deficient ^c	0 (0)** \downarrow	6 (11)* \downarrow	18 (33)	30 (56)** \uparrow	54
Controls (1–13 yr) ^d	34 (16)	52 (24)	77 (36)	50 (24)	213
Sμ-Sγ					
ICF2-deficient ^a	9 (20)	3 (7)	34 (74)	0 (0)	46
Lig4-deficient ^b	4 (12)	11 (32)	19 (56)	0 (0)	34
Artemis-deficient ^c	5 (21)	4 (17)	15 (63)	0 (0)	24
Controls (1–6 yr) ^e	13 (22)	9 (16)	36 (62)	0 (0)	58

Data are n (%). Statistical analysis was performed by χ^2 test. *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.0001.

^aThe sequences of S μ -S α and S μ -S γ junctions are available in Data S1 and Data S2, respectively.

^bPreviously published CSR junctions from Lig4-deficient patients (Pan-Hammarström et al., 2005).

^cPreviously published CSR junctions from Artemis-deficient patients (Du et al., 2008).

^dNewly acquired and previously published S μ -S α junctions from children controls (Du et al., 2008; Enverald et al., 2013).

^ePreviously published S μ -S γ junctions from children controls (Du et al., 2008).

Loss of ZBTB24 resembles NHEJ deficiency in CSR

CSR heavily relies on the c-NHEJ-mediated repair of AID-induced DSBs upstream of the constant regions of the *IgH* locus (Alt et al., 2013). To study the functional consequences of ZBTB24 mutations in the repair of DSBs during CSR, a PCR-based assay for amplification of S μ -S α junctions (located upstream of the C μ and C α regions of the *IgH* locus, respectively) was performed on ICF2-patient cells. 82 S μ -S α junctions were amplified from these patients and compared with 213 (30 newly generated and 183 previously published) S μ -S α junctions from healthy children who served as controls (Du et al., 2008; Enverald et al., 2013). The junctions from ICF2 patients showed an altered repair pattern with a decrease in direct end-joining (5% versus 16% in controls, χ^2 test, P = 0.0109; Table 1, Data S1, and Data S2), a decrease in small insertions (13% versus 24% in controls, χ^2 test, P = 0.0389; Table 1, Data S1, and Data S2), and an increased usage of long (≥ 7 bp) microhomologies (40% versus 24% in controls, χ^2 test, P = 0.0041; Table 1, Data S1, and Data S2), suggesting a shift to the use of an alternative end-joining pathway in cells from these patients. A similar shift is also apparent in c-NHEJ-deficient cells from patients with mutations in LIG4 or Artemis (Table 1, Data S1, and Data S2), suggesting that the shift to alternative repair may be due to a defect in c-NHEJ. Furthermore, 46 S μ -S γ junctions (located upstream of the C μ and C γ regions of the *IgH* locus, respectively) were isolated from the ICF2-deficient cells and compared with our previously published 58 S μ -S γ junctions from healthy controls (Du et al., 2008). Similar to patients with mutations in LIG4 or Artemis, the repair patterns at the S μ -S γ junctions were largely normal in ICF2 patients (Table 1 and Data S3), although one S μ -S γ junction showed a “footprint” of sequential switching (S μ -S γ 3-S γ 2; 9%), which is rarely observed in controls (2%) but frequently seen in

c-NHEJ-defective cells such as Artemis- or DNA-PKcs-deficient cells (Björkman et al., 2015; Du et al., 2008). Thus, the altered CSR patterns in ICF2 patient cells and their resemblance to those observed in several known c-NHEJ-deficient patients suggest that ZBTB24 might be a novel NHEJ factor involved in CSR.

ZBTB24 promotes DSB repair via c-NHEJ

To assess whether ZBTB24 is involved in NHEJ, which is the dominant pathway for the repair of DSBs in mammalian cells, we made use of the well-established HEK293T EJ5-GFP reporter cell line (Fig. 2 A; Bennardo et al., 2008). Depletion of ZBTB24 by multiple siRNAs resulted in a marked decrease in NHEJ, which was comparable to the impact of depleting XRCC4 (Fig. 2, B and C; and Fig. S2 A). Cell cycle profiles remained unaffected in these cells, ruling out effects of cell cycle misregulation (Fig. S2 B). To corroborate these findings, we also used fibroblasts containing the GC92 reporter (Fig. S2 C; Taty-Taty et al., 2016). Depletion of ZBTB24 by two different siRNAs resulted in a marked decrease in NHEJ, which was comparable to the impact of depleting KU80 and reminiscent of the effect on NHEJ observed in the EJ5-GFP reporter (Fig. 2, B and C; and Fig. S2, D and E).

The two major known pathways for the end-joining-dependent repair of DSBs in mammalian cells are c-NHEJ and a-NHEJ (Alt et al., 2013). Although the EJ5-GFP and GC92 reporters cannot differentiate between these pathways, we observed a remarkably similar phenotype after loss of ZBTB24 and the c-NHEJ factors XRCC4 and KU80 (Fig. 2, B and C; and Fig. S2, D and E). Moreover, ICF2 patient cells showed altered CSR patterns that resembled those observed in patient cells deficient for the c-NHEJ factor LIG4 (Table 1, Data S1, Data S2, and Data S3), suggesting a role for ZBTB24 in c-NHEJ. To provide further support for this, we used a plasmid integration assay

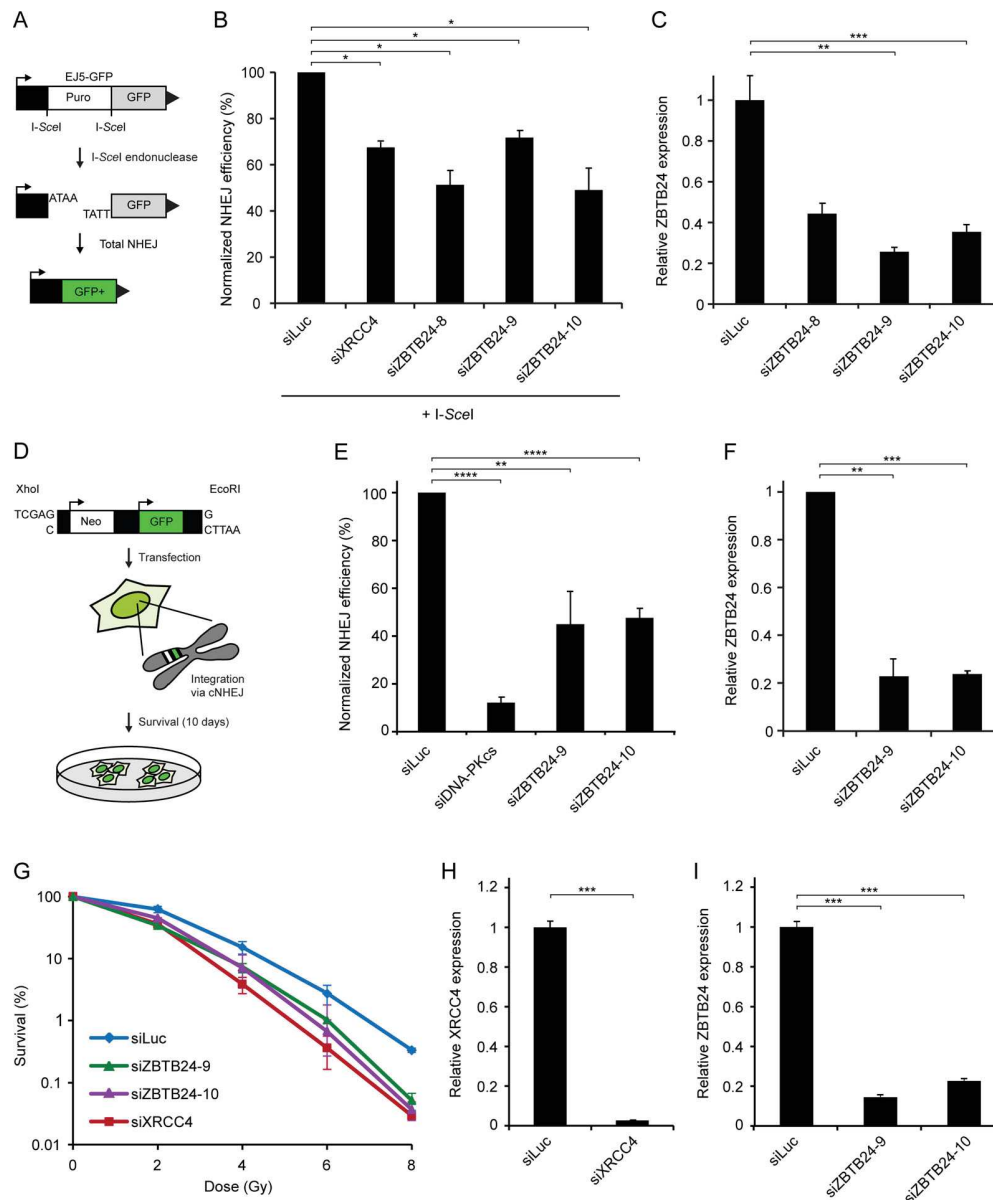


Figure 2. ZBTB24 promotes DSB repair via c-NHEJ. (A) Schematic representation of the EJ5-GFP reporter for NHEJ. (B) HEK293T EJ5-GFP cells were treated with the indicated siRNAs and, 48 h later, cotransfected with I-SceI (pCBASce) and mCherry expression vectors. The ratio of GFP/mCherry-expressing cells was counted by flow cytometry 48 h later. The mean \pm SD of two independent experiments is shown. Statistical significance was calculated using Student's *t* test (*, $P < 0.05$). (C) Cells from B were subjected to RNA extraction. cDNA was synthesized from total RNA samples followed by qPCR to determine the expression levels of ZBTB24. The mean \pm SEM of two independent experiments is shown. Statistical significance was calculated using Student's *t* test (**, $P < 0.01$; ***, $P < 0.001$). (D) Schematic of the plasmid integration assay. pEGFP-C1 plasmid containing Neo and GFP markers is linearized with the indicated restriction enzymes and transfected into U2OS cells. Stable integrants are selected on medium containing G418. GFP was used as a control for transfection efficiency. (E) Plasmid integration assays in U2OS cells transfected with the indicated siRNAs. The mean \pm SEM of two to four independent experiments is shown. Statistical significance was calculated using Student's *t* test (**, $P < 0.01$; ****, $P < 0.0001$). (F) As in C, except that cells from E were used. The mean \pm SEM of two independent experiments is shown. Statistical significance was calculated using Student's *t* test (**, $P < 0.01$; ***, $P < 0.001$). (G) VH10-SV40 cells were treated with the indicated siRNAs for 48 h, exposed to different doses of IR, and scored for clonogenic survival. The mean \pm SEM of two independent experiments is shown. (H) As in C, except that cells from G were used to monitor XRCC4 expression. The mean \pm SEM of two independent experiments is shown. Statistical significance was calculated using Student's *t* test (***, $P < 0.001$). (I) As in C, except that cells from G were used. The mean \pm SEM of two independent experiments is shown. Statistical significance was calculated using Student's *t* test (***, $P < 0.001$).

to specifically study the role of ZBTB24 in c-NHEJ (Fig. 2 D). Depletion of DNA-PKcs (catalytic subunit of DNA-PK complex) resulted in an 80–90% decrease in cell survival (Fig. 2 E and Fig. S2 F), indicating that the assay provides a readout for c-NHEJ as

reported previously (Caron et al., 2019; Luijsterburg et al., 2016). Moreover, knockdown of ZBTB24 caused a ~50% reduction in c-NHEJ efficiency compared with control cells (Fig. 2, E and F; and Fig. S2 A).

To rule out that ZBTB24 regulates NHEJ indirectly through transcriptional regulation of DSB repair factors, we depleted ZBTB24 and performed whole-transcriptome analysis using RNA sequencing in HEK293T cells (Fig. S2 G). In total, we found 158 differentially expressed genes (false discovery rate < 0.05), of which 90 were up-regulated and 68 were down-regulated (Table S2). We compared the list of deregulated genes with 66 unique genes in gene ontology term 0006302 (DSB repair), but we did not find any overlapping genes (Fig. S2 G). This suggests that ZBTB24 does not affect NHEJ through transcription regulation of DSB repair genes.

To assess the functional relevance of ZBTB24 in NHEJ, we investigated its ability to protect cells against DNA breaks induced by ionizing radiation (IR). To this end, clonogenic survival of VH10-SV40 cells depleted for ZBTB24 or XRCC4 was determined after exposure to IR. This showed a similar dose-dependent decrease in the survival capacity of ZBTB24-depleted and XRCC4-depleted cells compared with control cells (siLuc; Fig. 2, G-I). Surprisingly, however, ICF2 patient-derived fibroblast cells did not show sensitivity to IR (Fig. S2 H). To corroborate these findings, we generated ZBTB24 KO U2OS cells using CRISPR/Cas9-based genome editing. Two independent ZBTB24 KO clones also did not display IR sensitivity, although these clones showed the previously reported reduction in CDCA7 expression (Fig. S2, I-K; Wu et al., 2016). Thus, the NHEJ phenotype is specifically observed in B cells from ICF2 patients and after short-term loss of ZBTB24 in differentiated human cells. These results underscore the functional importance of ZBTB24 in the protection of cells against DNA breaks and implicate a role for ZBTB24 in DSB repair by NHEJ.

ZBTB24 interacts with PARP1 in a PARylation-dependent manner

To assess how ZBTB24 affects NHEJ, we aimed to identify its interaction partners using an unbiased, quantitative proteomics approach. To this end, we expressed GFP-ZBTB24 or GFP (control) in U2OS cells and performed GFP-trap-based immunoprecipitations (IPs) followed by mass spectrometry (MS) after stable isotope labeling of amino acids in culture (SILAC; Fig. 3 A). Our screen identified 110 proteins that were at least fourfold enriched over control cells (Table S3). Interestingly, besides all core histones, PARP1, an enzyme implicated in NHEJ and other DNA repair mechanisms (Ray Chaudhuri and Nussenzweig, 2017), was among the potential interactors of ZBTB24 (Fig. 3 A and Table S3). To explore this further, we performed the reciprocal experiment using cells expressing GFP-PARP1. This screen identified 21 proteins that were at least twofold enriched over control cells (Fig. 3 B and Table S4). Remarkably, not only did we find several known PARP1 interactors such as XRCC1, LIG3, and DNA polymerase β (POLB; Pines et al., 2013), but ZBTB24 was also among the top hits of this screen (Fig. 3 B and Table S4). To confirm the ZBTB24-PARP1 interaction, we performed pull-down IP experiments followed by Western blot analysis. Endogenous PARP1, as well as histone H3, were detected after IP of GFP-ZBTB24, whereas GFP-PARP1 efficiently precipitated Myc-ZBTB24 in the

reciprocal IP (Fig. 3 C and Fig. S3 A). Moreover, using coimmunoprecipitation experiments, we also confirmed that endogenous PARP1 interacts with endogenous ZBTB24 (Fig. 3 D).

PARP1 can covalently link negatively charged ADP-ribose units to itself or to other target proteins, forming poly(ADP-ribose) (PAR) chains through a process known as PARylation (Pines et al., 2013). Upon addition of PARP inhibitor (PARPi), PARylation was efficiently inhibited, and the interaction between ZBTB24 and PARP1 was lost (Fig. 3 C and Fig. S3 B). Together, these results suggest that ZBTB24 and PARP1 interact in a PARylation-dependent manner.

PARP1 recruits ZBTB24 to sites of DNA damage

PARP1 binds to both single-strand breaks and DSBs, where it promotes the assembly of chromatin remodelers and DNA repair proteins (Pines et al., 2013). Given the interaction between ZBTB24 and PARP1, we tested whether ZBTB24 is recruited to sites of DNA damage. We found that both N- and C-terminally tagged ZBTB24 localize at laser microirradiation-induced tracks containing γ H2AX, a known marker of DNA damage (Fig. 3, E and F). Importantly, ZBTB24 recruitment, but not that of the DNA damage sensor NBS1, to such DNA damage tracks was completely abrogated after treatment with PARPi (Fig. 3, G and H), demonstrating its dependence on PARylation. Furthermore, the accumulation of ZBTB24 at DNA damage tracks was rapid but transient, reaching maximum levels at \sim 100 s after DNA damage induction (Fig. 3 H) and greatly resembling the reported dynamics of PARP1 accrual and PARylation at sites of DNA damage (Mortusewicz et al., 2007). Importantly, siRNA-mediated depletion of PARP1, but not PARP2, abrogated ZBTB24 accumulation in laser tracks (Fig. 3, I and J). These results show that ZBTB24 is rapidly recruited to sites of DNA damage in a PARP1- and PARylation dependent manner.

PAR chains are rapidly hydrolyzed by the activity of poly(ADP-ribose) glycohydrolase (PARG), which explains the rapid turnover of PAR chains at sites of DNA damage (Pines et al., 2013). To prevent this rapid turnover, we increased the steady-state levels of PAR chains by siRNA-mediated depletion of PARG (Fig. S3 C). Under these conditions, we observed enhanced and more persistent accumulation of ZBTB24 at sites of damage (Fig. S3, D and E). In contrast, overexpression of mCherry-tagged PARG resulted in a dramatic decrease in the total levels of PARylation and abrogated recruitment of ZBTB24 to sites of damage (Fig. S3, F-H), phenocopying the effect observed after loss of PARP1 activity (Fig. 3, F and G). Thus, the PARP1- and PARG-dependent turnover of PAR chains at DNA lesions is a critical determinant of the rapid and transient accumulation of ZBTB24.

The zinc-finger (ZNF) of ZBTB24 binds PAR to promote PARP1-dependent ZBTB24 recruitment

Three conserved domains can be identified in ZBTB24: an N-terminal BTB domain (aa 9-132), a small AT-hook DNA-binding domain (aa 159-171), and eight tandem C2H2 ZNF motifs (aa 294-512; Fig. 4 A). To dissect the relevance of these domains for ZBTB24's interaction with PARP1 and localization to DNA damage sites, we generated and expressed GFP-fusion

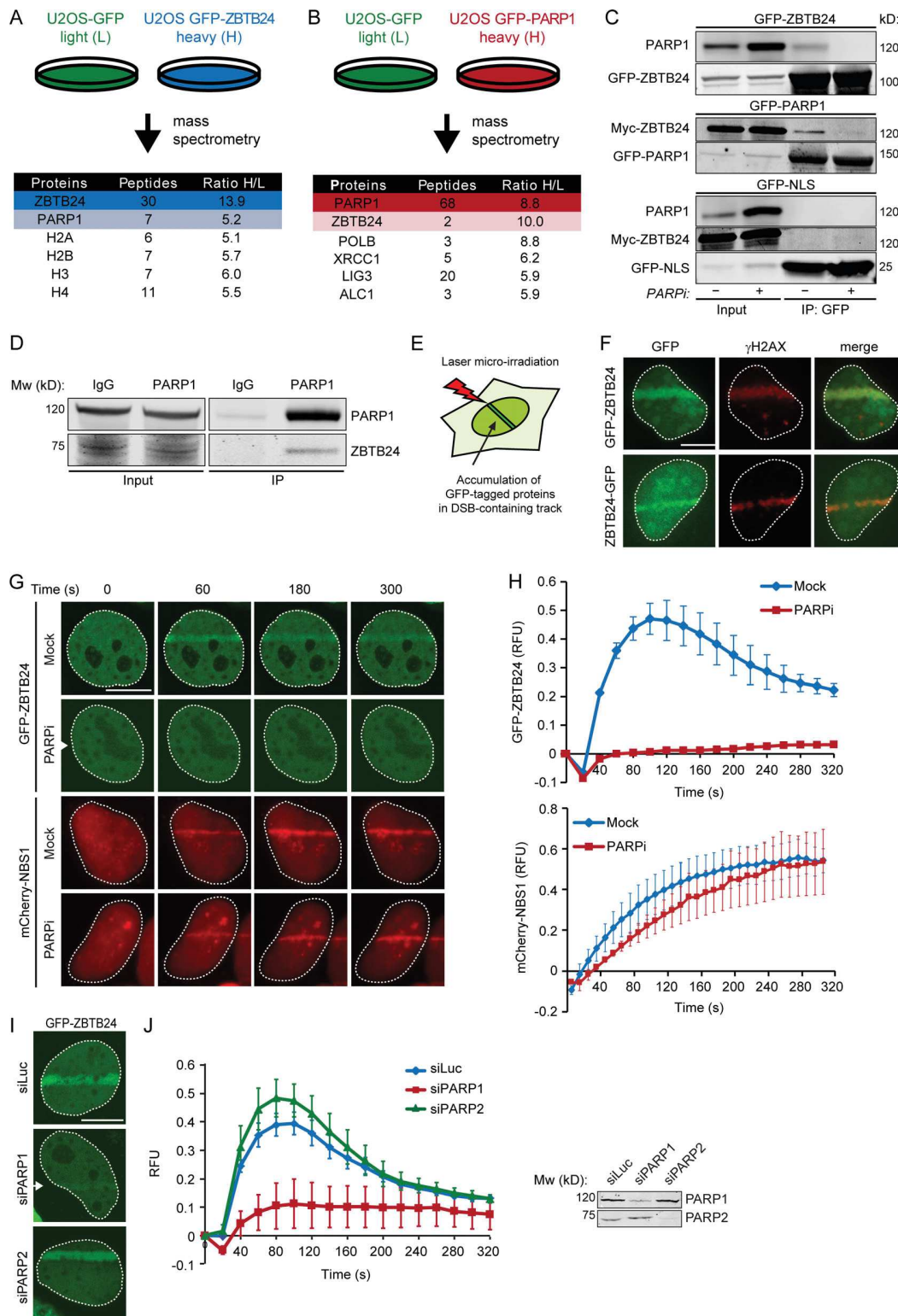


Figure 3. PARP1 interacts with ZBTB24 in a PARylation-dependent manner and recruits ZBTB24 to sites of DNA damage. (A) Schematic representation of SILAC-based MS approach. GFP- or GFP-ZBTB24-expressing U2OS cells were labeled with Lys0 and Arg0 (L) or Lys8 and Arg10 (H), respectively. Lysates were subjected to GFP IP, and equal amounts of both IP fractions were mixed. Proteins in the IP fractions were digested by trypsin and subjected to MS analysis. A list of ZBTB24-interacting proteins, including the number of peptides and the interaction ratio from heavy (H)- over light (L)-labeled cell extracts as revealed by MS, is shown. **(B)** As in A, but with GFP- and GFP-PARP1-expressing U2OS cells. **(C)** Cells expressing GFP-ZBTB24, GFP-PARP1, and Myc-ZBTB24 or GFP-NLS and Myc-ZBTB24 were treated with either DMSO (Mock) or PARP inhibitor (PARPi). Whole-cell extracts (WCEs) were subjected to GFP IP followed by Western blot analysis of the indicated proteins. **(D)** IP of endogenous PARP1 in U2OS cells. IgG is a negative control. Blots were probed for ZBTB24 and

PARP1. **(E)** Schematic representation of the laser microirradiation approach. **(F)** GFP-ZBTB24 or ZBTB24-GFP accumulate at γ H2AX-decorated DNA damage tracks after transient expression and laser microirradiation in U2OS cells. Scale bar, 10 μ m. **(G)** As in F, except that cells transiently expressing GFP-ZBTB24 and mCherry-NBS1 were treated with either DMSO (Mock) or PARPi before GFP-ZBTB24 and mCherry-NBS1 accumulation was monitored at the indicated time points after laser microirradiation. Scale bar, 10 μ m. **(H)** Quantification of the results from G. The mean \pm SEM of two to three independent experiments is shown. **(I)** As in G, except that cells were cotransfected with GFP-ZBTB24 and the indicated siRNAs. Scale bar, 10 μ m. **(J)** Quantification of the results from I. The mean \pm SEM of two to three independent experiments is shown (left). Western blot showing the knockdown efficiency of PARP1 and PARP2 (right).

constructs of the different domains (Fig. 4, B-E). Interestingly, GFP-BTB, GFP-BTB-AT, or GFP- Δ ZNF did not accumulate at sites of laser-induced DNA damage, whereas GFP-BTB-AT-ZNF (GFP-BAZ) and GFP-ZNF were recruited with similar kinetics as GFP-ZBTB24 (Fig. 4 C and Fig. S3, I-L) and in a manner dependent on PARP activity as well (Fig. S3, I-L). This suggests that the ZNF domain is required for the PARP1 activity-dependent accumulation of ZBTB24 at sites of DNA damage.

PARP1 is responsible for ~85% of the synthesized PAR chains in cells (Mortusewicz et al., 2007; Shieh et al., 1998). The PAR-dependent accumulation of ZBTB24 could be a consequence of the PARylation of ZBTB24 or the binding of ZBTB24 to PARP1-

associated PAR chains. To examine whether ZBTB24 itself is PARylated, we exposed cells to IR or the DNA-alkylating agent *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG) and compared the PARylation status of ZBTB24 to that of PARP1. We observed a significant increase in PARylated proteins after MNNG treatment, and a modest increase shortly after exposure to IR (Fig. S3 M), indicating that these treatments result in the activation of PARP enzymes. Subsequently, we immunoprecipitated GFP-ZBTB24 or GFP-PARP1 from these cells using stringent, high-salt wash conditions to disrupt all noncovalent protein-protein interactions and examined their PARylation status by Western blot analysis. As expected, PARP1 was strongly PARylated under

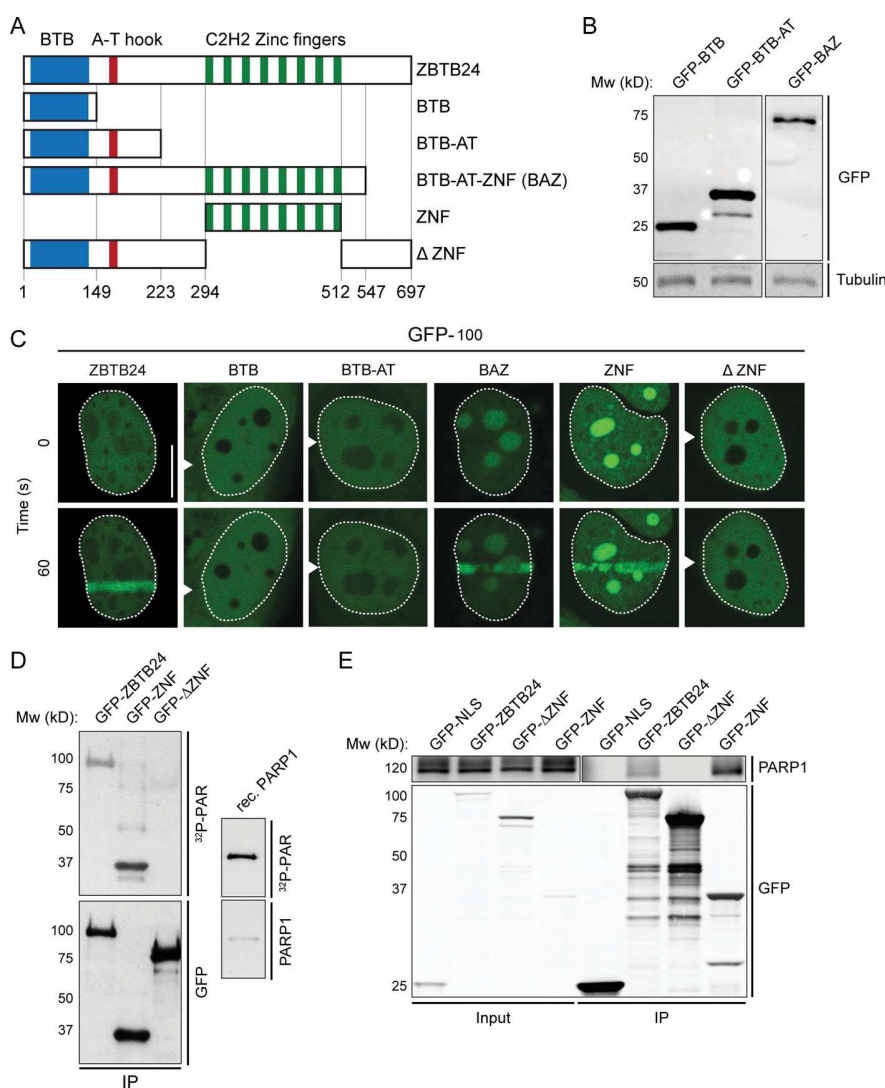


Figure 4. The ZNF domain in ZBTB24 interacts with PAR and mediates its recruitment to sites of DNA damage. **(A)** Schematic representation of isoform 1 of ZBTB24 and its BTB-, DNA-binding AT hook- and $8 \times$ C₂H₂ ZNF domain. Protein domains were separated as indicated and fused to GFP for functional analysis. **(B)** Western blot analysis of WCEs from U2OS cells expressing the indicated GFP-tagged ZBTB24 domains. **(C)** Accumulation of the indicated GFP-tagged ZBTB24 domain in laser microirradiated U2OS cells. Representative images of unirradiated and irradiated cells (taken at the indicated time point after irradiation) are shown. Scale bar, 10 μ m. **(D)** HEK293T cells expressing the indicated GFP-tagged ZBTB24 domains were subjected to GFP IP. GFP-purified proteins were resolved by gel electrophoresis, blotted onto a membrane, renatured, and incubated with radioactive PAR (³²P-PAR). Recombinant (rec.) PARP1 is a positive control. **(E)** Lysates from U2OS cells transiently expressing either GFP-NLS or the indicated GFP-tagged ZBTB24 domains were subjected to GFP IP and Western blot analysis for the indicated proteins.

all conditions (Fig. S3 N), showing that our approach can detect the attachment of PAR chains to proteins. However, we failed to detect PARylation of ZBTB24 under these conditions, suggesting that ZBTB24 is not a preferred target for PARylation by PARP1 (Fig. S3 N).

Next, we examined whether ZBTB24 could physically associate with PAR chains *in vitro* by using Southwestern blotting. GFP-ZBTB24 was immunoprecipitated, transferred to a membrane, and exposed to *in vitro*-generated ³²P-labeled PAR chains. Indeed, GFP-ZBTB24, similar to recombinant PARP1, was able to bind PAR chains efficiently (Fig. 4 D). Because the ZNF domain in ZBTB24 is a key determinant of the PARP1 activity-dependent recruitment of ZBTB24 to sites of DNA damage, we examined whether this domain would mediate the interaction with PAR polymers. Indeed, GFP-ZNF, but not GFP-ΔZNF (full-length ZBTB24 lacking the ZNF domain), was able to bind PAR chains (Fig. 4 D). In concordance, IP experiments revealed an interaction between PARP1 and GFP-ZNF, but not GFP-ΔZNF (Fig. 4 E). Together, these results suggest that the ZNF of ZBTB24 acts as a PAR-binding domain that mediates ZBTB24 recruitment to DNA damage through interactions with PARylated PARP1.

ZBTB24 promotes PAR synthesis and protects PAR chains through its ZNF

Considering that ZBTB24 efficiently associates with PARP1-generated PAR chains, we wondered whether ZBTB24 could be involved in regulating the steady-state levels of such chains in response to DNA damage. To examine this possibility, we monitored global PAR levels by Western blot analysis in cells exposed to IR. Although hardly any PARylation could be observed in mock-treated cells, exposure to IR triggered robust DNA damage-induced PARylation (Fig. 5, A and B), which was largely suppressed (~60–70%) by knockdown of PARP1 (Fig. 5, A and B). Strikingly, knockdown of ZBTB24 also caused a significant reduction (~50%) in PARylation in IR-exposed cells (Fig. 5, A and B), suggesting that ZBTB24 is required to boost the DNA damage-induced PARylation response.

It is feasible that ZBTB24 regulates steady-state PAR levels by either stimulating the synthesis of such chains or preventing their degradation. To examine a potential stimulatory role for ZBTB24 in PAR synthesis, we reconstituted PARP1-dependent synthesis of PAR in an *in vitro* system in the absence or presence of recombinant ZBTB24 or ZBTB24 lacking its ZNF domain (ZBTB24 ΔZNF; Fig. 5 C and Fig. S4 A). In the presence of NAD⁺ and a damaged DNA template, we found that the capacity of recombinant PARP1 to synthesize PAR chains was slightly enhanced by the presence of recombinant ZBTB24, but not ZBTB24 ΔZNF (Fig. 5, D and E), suggesting that ZBTB24 may weakly stimulate PARP1-dependent PAR synthesis in manner dependent on its ZNF.

Another nonmutually exclusive possibility is that ZBTB24 binding to PAR chains protects such chains from efficient hydrolysis by the PARP1 antagonist PARG. To explore this possibility, we allowed PARP1-dependent synthesis of PAR in our *in vitro* system and, after the inactivation of PARP1 by PARP1, added recombinant PARG hydrolase with increasing amounts of

recombinant ZBTB24 or ZBTB24 ΔZNF (Fig. 5 F). We could detect efficient hydrolysis of nearly all PAR chains in the absence of ZBTB24 or ZBTB24 ΔZNF (lane 1 versus 2; Fig. 5 G). Interestingly, ZBTB24 inhibited in a dose-dependent manner the breakdown of PAR products in the hydrolysis reaction (Fig. 5, G and H), whereas ZBTB24 ΔZNF was unable to do so, suggesting that ZBTB24 binds PAR chains through its ZNF to protect them from PARG-dependent degradation (Fig. 5, G and H). In conclusion, we found that ZBTB24 promotes the steady-state levels of DNA damage-induced PAR chains by stimulating the PARP1-dependent synthesis and inhibiting the PARG-dependent hydrolysis of such chains.

ZBTB24 and PARP1 promote c-NHEJ by regulating XRCC4/LIG4 assembly

We then sought to address how ZBTB24's role in PAR synthesis and protection is linked to its involvement in c-NHEJ (Fig. 2 and Fig. S2, C–E). Interestingly, *in vitro* studies demonstrated that the c-NHEJ ligase LIG4 interacts with PAR chains through its C-terminal BRCT domain (Li et al., 2013), providing a possible link between ZBTB24's involvement in PAR stability and NHEJ. To study this further, we first applied laser microirradiation to monitor the recruitment of GFP-XRCC4 to damaged DNA in U2OS cells that were either treated with PARP1 or depleted for PARP1. In line with previously published work (Luijsterburg et al., 2016), the loss of both PARP activity and PARP1 protein markedly impaired the recruitment of GFP-XRCC4 (Fig. S4, B–E), suggesting that PARP1-dependent PARylation regulates the assembly of XRCC4/LIG4 complexes at sites of DNA damage to promote c-NHEJ. To confirm this, we used the plasmid integration assay to specifically examine PARP1's contribution to c-NHEJ. In agreement with our recruitment data and previous findings (Luijsterburg et al., 2016), we found that PARP1 depletion resulted in a ~40% reduction in c-NHEJ efficiency (Fig. S2 F and Fig. 4 F), suggesting that PARP1, similar to ZBTB24 (Fig. 2, E and F), promotes c-NHEJ.

Given ZBTB24's role in NHEJ, its interaction with PARP1 and its stimulatory effect on PARylation, we addressed whether ZBTB24 affects the PARP1-dependent assembly of XRCC4/LIG4 at DSBs. Depletion of ZBTB24, similar to that of PARP1, resulted in a strong reduction in GFP-XRCC4 recruitment at sites of laser-induced DNA damage (Fig. 6, A and B). Moreover, ZBTB24 depletion also reduced the accumulation of endogenous XRCC4, whereas DNA damage levels measured by γH2AX remained unaffected (Fig. S4, G–I). Importantly, the accumulation of GFP-XRCC4 at a stably integrated lactose operator (LacO) array upon tethering of a lactose repressor (LacR)-tagged FokI nuclease in U2OS cells was also strongly reduced in cells depleted for ZBTB24 (Fig. 6, C–F). This indicates that ZBTB24 acts at bona fide DSBs to facilitate the accumulation of functional XRCC4/LIG4 complexes.

Finally, we showed that ZBTB24's ZNF domain is important for the PARP1 activity-dependent recruitment of ZBTB24 to DSBs (Fig. 4 and Fig. S3, I–N). Based on this and the fact that ZBTB24 promotes XRCC4/LIG4 accrual at DSBs, we hypothesized that ZBTB24's ZNF may play an important role in this process. To examine this, we generated HeLa Flp-In/T-REx cells

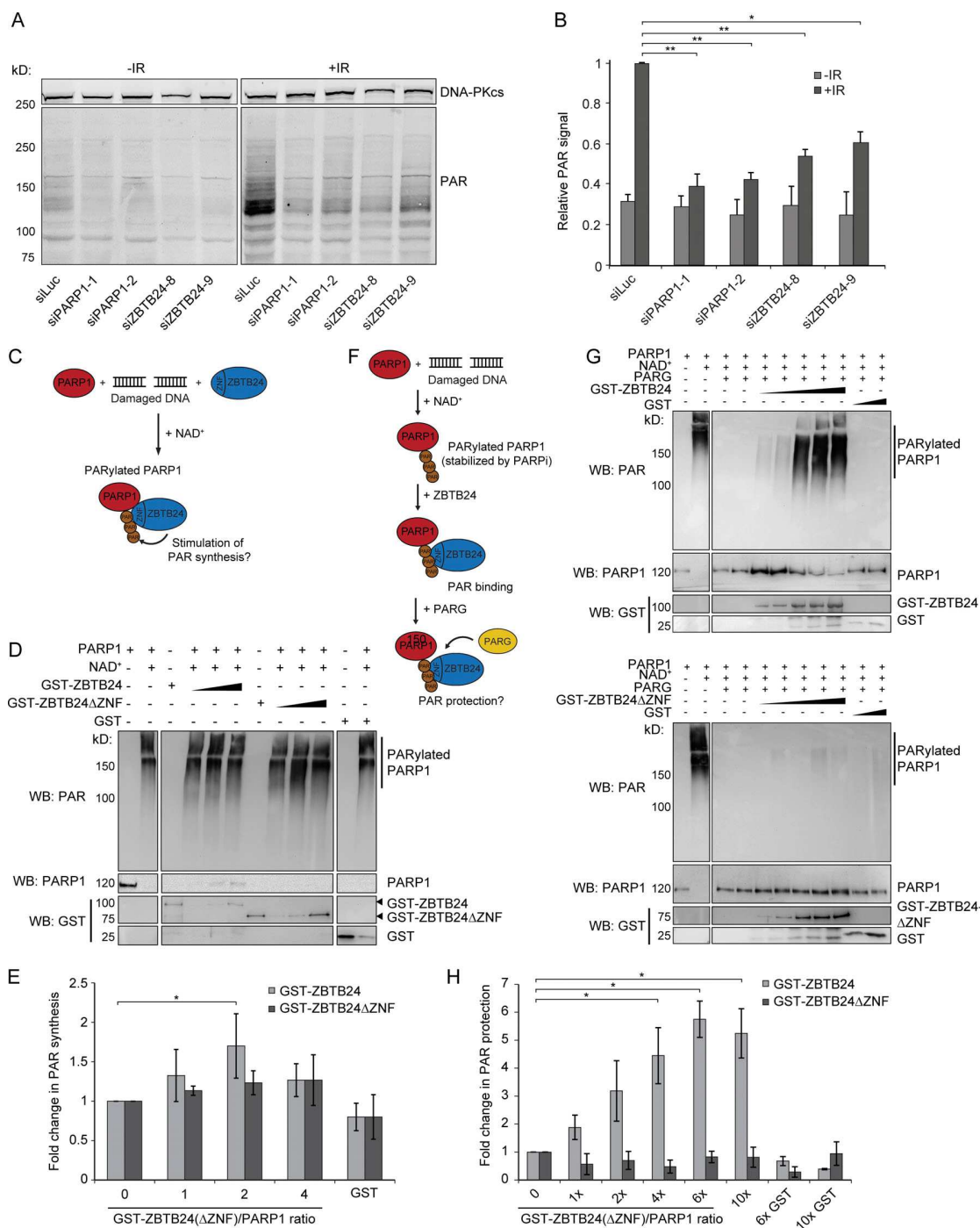


Figure 5. ZBTB24 stimulates PARP1-dependent PAR synthesis and protects PAR chain stability. (A) U2OS cells transfected with the indicated siRNAs were left untreated or exposed to IR. 5 min later, whole-cell extracts (WCEs) were prepared and subjected to Western blot analysis for DNA-PKcs and PAR. DNA-PKcs is a loading control. (B) Quantification of the results from A and a second independent experiment. The mean \pm SEM is shown. The ratio of PAR/loading control signals per sample was normalized to that of the IR-exposed siLuc sample, which was set to 1. Statistical significance was calculated using Student's *t* test (*, $P < 0.05$; **, $P < 0.01$). (C) Schematic of the PAR synthesis assay. (D) Recombinant PARP1 was incubated with a damaged DNA template and activated by NAD^+ in the presence of increasing concentrations of GST-ZBTB24, GST-ZBTB24 Δ ZNF, or GST only. The presence of 10H-PAR chains and recombinant proteins was monitored by Western blot analysis. (E) Quantification of ZBTB24-dependent stimulation of PAR synthesis from D and two other independent experiments. The mean \pm SD is shown. The signal of 10H-PAR for each sample containing GST-ZBTB24 or GST-ZBTB24 Δ ZNF was normalized to that without GST-ZBTB24, which was set to 1. Statistical significance was calculated using Student's *t* test (*, $P < 0.05$). (F) Schematic of the PAR protection assay. (G) Recombinant PARP1 was incubated with a damaged DNA template and activated by NAD^+ to generate PARylated PARP1. Increasing concentrations of GST-ZBTB24, GST-ZBTB24 Δ ZNF, or GST alone were added, followed by incubation with PARG. The presence of 10H-PAR chains and recombinant proteins was monitored by Western blot analysis. (H) As in E, except that PAR protection was measured from G and another independent experiment. The mean \pm SD is shown. Statistical significance was calculated using Student's *t* test (*, $P < 0.05$).

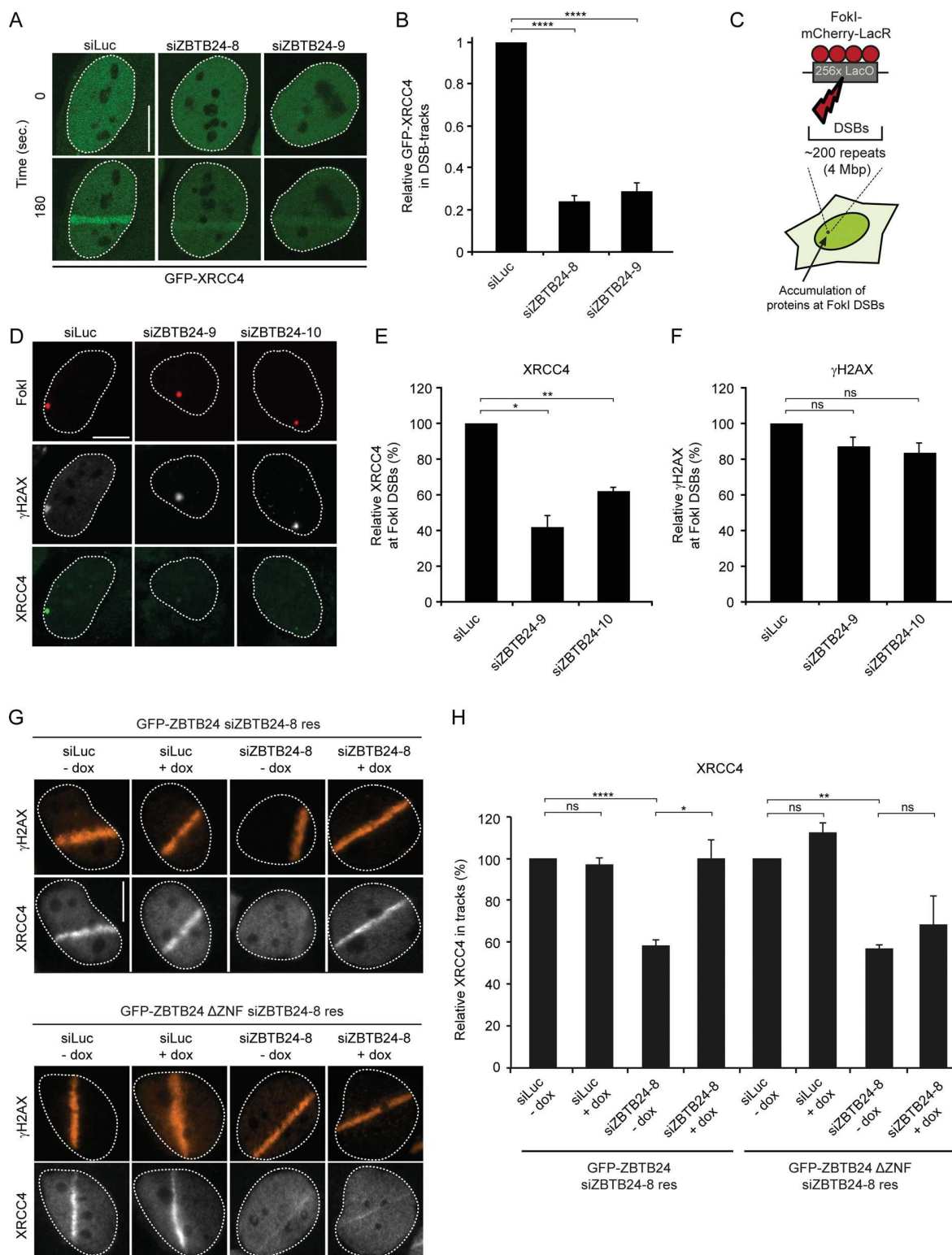


Figure 6. ZBTB24 promotes XRCC4/LIG4 assembly at DNA damage sites. (A) U2OS stably expressing GFP-XRCC4 were transfected with the indicated siRNAs and subjected to laser microirradiation. Representative images of unirradiated and irradiated cells (taken at the indicated time point after irradiation) are shown. Scale bar, 10 μm. (B) Quantification of A. The mean ± SEM of three independent experiments is shown. Statistical significance was calculated using Student's *t* test (****, *P* < 0.0001). (C) Schematic of the system in U2OS 2-6-3 cells used to locally induce multiple DSBs upon tethering of the FokI endonuclease. (D) Accumulation of XRCC4 (green) to γH2AX-marked (white) DSBs induced by FokI-mCherry-LacR at a LacO array (red) in cells transfected with the indicated siRNAs. Scale bar, 10 μm. (E) Quantification of XRCC4 accumulation in D. The mean ± SD of two independent experiments is shown. Statistical significance was calculated using Student's *t* test (*, *P* < 0.05; **, *P* < 0.01). (F) As in E, except for γH2AX. The mean ± SD of two independent experiments is

shown. Statistical significance was calculated using Student's *t* test (ns, not significant). **(G)** Accumulation of endogenous XRCC4 (white) and γ H2AX (orange) in laser microirradiated HeLa Flp-In/T-Rex cells expressing doxycycline (dox)-inducible GFP-ZBTB24 or GFP-ZBTB24 Δ ZNF after transfection with the indicated siRNAs. Cells were fixed and immunostained 10 min after laser microirradiation. Scale bar, 10 μ m. **(H)** Quantification of endogenous XRCC4 levels in laser tracks from G. The mean \pm SEM of two to three independent experiments is shown. Statistical significance was calculated using Student's *t* test (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.0001$; ns, not significant).

stably expressing inducible and siRNA-resistant GFP-tagged ZBTB24 or ZBTB24 Δ ZNF (Fig. S5, A and B). Confirming our previous data (Fig. 4 C and Fig. S3, K–L), we found that GFP-ZBTB24 was recruited to laser-induced DNA damage tracks, whereas GFP-ZBTB24 Δ ZNF failed to do so after expression in ZBTB24-depleted cells (Fig. S5, C and D). Importantly, the expression of GFP-ZBTB24, but not that of GFP-ZBTB24 Δ ZNF, rescued the reduced XRCC4 accumulation in ZBTB24 knock-down cells (Fig. 6, G and H), whereas γ H2AX signals remained unaffected (Fig. S5 E). This indicated that the NHEJ defects observed in ZBTB24-depleted cells were not due to off-target effects of the siRNAs (Figs. 2 and 6). Moreover, these results show that the ZNF in ZBTB24 plays a critical role in regulating the PARP1 activity-dependent assembly of XRCC4/LIG4 at DNA breaks that undergo c-NHEJ.

ZBTB24-deficient cells show a-NHEJ signatures at repair junctions and impaired CSR

ICF2 patients with loss of ZBTB24 suffer from immunodeficiency characterized by defective CSR (Fig. 1). At the molecular level, we found that CSR junctions in B cells from ICF2 patients show an altered repair pattern with a decrease in direct end-joining and an increased usage of long microhomologies, suggesting a shift from the use of c-NHEJ to the use of a-NHEJ similar to that observed in B cells from LIG4- and Artemis-deficient patients (Table 1, Data S1, Data S2, and Data S3). To corroborate these findings, we examined mutational signatures at repair junctions in the GC92-NHEJ reporter (Taty-Taty et al., 2016), in which we observed that loss of ZBTB24 impairs NHEJ (Fig. S2, C–E). Interestingly, compared with control cells, ZBTB24 depletion increased the proportion of larger deletions and use of microhomology during repair (Fig. 7, A–C) to a similar extent as observed after KU80 knockdown (Fig. 7, A–C; Kabotyanski et al., 1998). These repair features in ZBTB24-depleted cells were reminiscent of those observed at CSR junctions in B cells from ICF2 patients (Table 1).

To determine whether ZBTB24 is involved in CSR, we inactivated the *Zbtb24* gene using CRISPR/Cas9-based genome editing in CH12 cells, a murine B cell line that can be induced to express AID and undergo CSR from IgM to IgA in vitro. We obtained one *Zbtb24*^{+/-} and two *Zbtb24*^{-/-} CH12 B cell clones (Fig. 7 D). Upon CSR induction, we found that the *Zbtb24*^{-/-} clones displayed defective CSR compared with *Zbtb24*^{+/+} or *Zbtb24*^{+/-} controls (Fig. 7, E and F). Importantly, the observed CSR defect was independent of defects in AID expression (Fig. 7 D). To rule out potential off-target effects of Cas9-based genome editing and to demonstrate that the CSR defect observed in *Zbtb24*^{-/-} cells is due to the absence of Zbtb24, we reexpressed mouse *Zbtb24* cDNA in these cells (Fig. 7 G). We found that overexpression of mZbtb24 rescued the CSR defect in

Zbtb24^{-/-} cells (Fig. 7, H and I), demonstrating that the loss of *Zbtb24* caused the CSR defect. Taken together, these findings show that ZBTB24 is involved in c-NHEJ during CSR, providing a molecular basis for the immunodeficiency in ZBTB24-deficient ICF2 patients.

Discussion

Mutations in at least four different genes cause the primary immunodeficiency ICF. Approximately 30% of the ICF patients carry causal mutations in the uncharacterized ZBTB24 gene (ICF2; Thijssen et al., 2015; Weemaes et al., 2013). Here, we functionally characterized the role of ZBTB24 in relation to the immunodeficiency by biochemical and cell biological approaches, as well as by functional analysis in patient-derived material. In ICF2 patients, we observed a severe reduction in Ig production and diversification capacity, and a shift toward a-NHEJ events during CSR characterized by larger deletions and more microhomology use in the switch junctions, which is reminiscent of the phenotype observed in cells from c-NHEJ-deficient patients (Du et al., 2008; Pan-Hammarström et al., 2005). Indeed, recent work suggested that in the absence of the KU70/80 complex, Rad52 binds to DSB ends within S regions to modulate CSR by a microhomology-mediated a-NHEJ process (Zan et al., 2017). Our findings provide a plausible molecular explanation for the currently unexplained immunodeficiency in ICF2 and suggest a role for ZBTB24 in c-NHEJ. Indeed, we reveal that ZBTB24 is recruited to sites of DNA damage in a PARP1-dependent manner by associating with PARP1-generated PAR-chains through its ZNF domain. Our biochemical and cellular analyses show that ZBTB24 promotes PARP1-mediated PAR synthesis and acts as a scaffold protein that protects PAR chains from degradation, thereby enhancing the PARP1-dependent recruitment of the LIG4/XRCC4 complex to facilitate efficient DSB repair by c-NHEJ (see model; Fig. 7 J). Consequently, ZBTB24 loss shifts DSB repair from LIG4/XRCC4-dependent c-NHEJ to a-NHEJ, consistent with the CSR phenotype observed in ICF2 patients (Fig. 7 K).

ZBTB24 is required for CSR, a process defective in ICF2 patients

Mutations in ZBTB24 lead to defective CSR in ICF2 patients, whereas V(D)J recombination remains unaffected. This may be unexpected, considering that both processes heavily rely on c-NHEJ. However, mutations in several other DNA damage response (DDR) genes, such as H2AX, NIPBL, and ATM in both mice and humans, cause a remarkably similar defect in CSR without affecting V(D)J recombination (Enervald et al., 2013; Manis et al., 2004; Pan et al., 2002; Reina-San-Martin et al., 2003). It has been suggested that the ends of recombination-

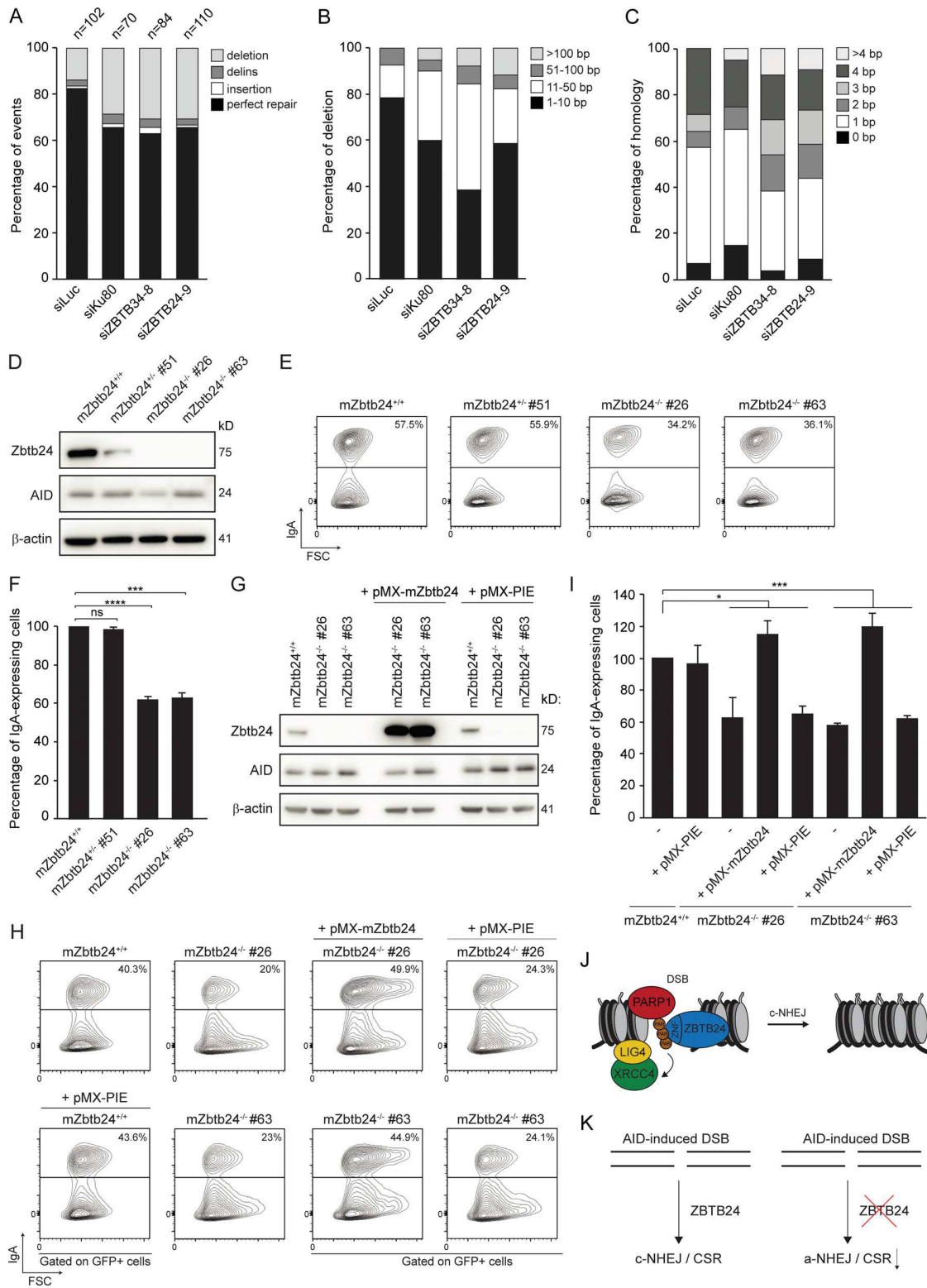


Figure 7. ZBTB24-deficient cells show a-NHEJ signatures at repair junctions. (A–C) Mutational signatures (A), deletion sizes (B), and microhomology usage (in case of deletion formation; C) at repair junctions in the GC92 reporter for NHEJ. GC92 cells were transfected with the indicated siRNAs and I-SceI expression vector. Repair junctions were amplified by PCR and Sanger sequenced. The bars represent data obtained from three independent experiments. **(D)** Western blot analysis of Zbtb24 and AID expression in wild-type CH12 cells (Zbtb24^{+/+}), Zbtb24^{+/-}, and Zbtb24^{-/-} clones stimulated to undergo CSR for 3 d with TGF-β, IL-4, and an anti-CD40 antibody. β-Actin is a loading control. **(E)** Flow cytometry analysis of cells from D. The percentage of IgA-expressing cells is indicated. Representative contour plots of three independent experiments are shown. **(F)** Quantification of cells from E. The mean ± SD from three

independent experiments is shown. Data were normalized to wild-type *Zbtb24*^{+/+} cells, which were set to 100%. Statistical significance was calculated using Student's *t* test (***, *P* < 0.001; ****, *P* < 0.0001; ns, not significant). **(G)** Western blot analysis of *Zbtb24* and AID expression in *Zbtb24*^{-/-} clones transduced with an empty retrovirus (pMX-PIE) or a retrovirus expressing *mZbtb24* and EGFP cDNA (pMX-*mZbtb24*) and stimulated to undergo CSR for 3 d with TGF- β , IL-4, and an anti-CD40 antibody. β -Actin is a loading control. **(H)** Flow cytometry analysis of cells from G. Dot plots are gated on EGFP⁺ cells. The percentage of IgA-expressing cells is indicated. Representative contour plots of two independent experiments are shown. **(I)** Quantification of cells from H. The mean \pm SD from two independent experiments is shown. Data were normalized to uninfected wild-type *Zbtb24*^{+/+} cells, which were set to 100%. -, not infected. Statistical significance was calculated using one-way ANOVA (*, *P* < 0.05; ***, *P* < 0.001). **(J)** Model for the role of ZBTB24 in DSB repair by NHEJ. ZBTB24 accumulates at DSBs, where it functions as a scaffold to protect PARP1-associated PAR-chains, which serve as a docking site for the LIG4-XRCC4 complex, facilitating efficient repair of DSBs via c-NHEJ. **(K)** Schematic illustrating that ZBTB24 loss leads to a shift from c-NHEJ to a-NHEJ and impaired CSR at AID-induced DSBs in B cells.

activating gene 1/2 (*RAG1/2*)-induced DSBs are held together by these enzymes during V(D)J recombination. In contrast, AID-initiated DSBs during CSR are likely held together by factors involved in the signaling of DSB, such as the core chromatin component H2AX and 53BP1 (Manis et al., 2004; Petersen et al., 2001). The role of ZBTB24 may resemble that of the latter DDR components, explaining its specific impact on CSR. Alternatively, *RAG1/2* induces DSBs that are characterized by the production of a hairpin structure at the broken ends. PARP1 swiftly binds to single-strand breaks and DSBs (Eustermann et al., 2011; Langelier et al., 2012), as well as to hairpin structures in vitro (Lonskaya et al., 2005). However, whether it also displays affinity for *RAG1/2*-induced hairpin structures at DSBs in vivo remains to be determined. It is conceivable that these structures are not bound by PARP1 owing to their processing by the structure-specific endonuclease Artemis (Alt et al., 2013), which could rule out a function for PARP1 and most likely ZBTB24 in V(D)J recombination and would be in agreement with our observations. However, PARP1 is activated by and seems to have affinity for AID-induced breaks in mice, where it promotes CSR through a-NHEJ (Robert et al., 2009). Whether it also modulates CSR in humans remains elusive, mainly because patients with loss-of-function mutations in PARP1 have not been reported yet.

Neither ICF2 patient-derived fibroblast cells nor ZBTB24 KO U2OS cells displayed hypersensitivity to IR-induced DNA breaks, suggesting a cell type- and/or context-specific role for ZBTB24 in NHEJ. It was reported that ZBTB24 KO HEK293 cells, which showed reduced CDCA7 expression, were also not sensitive to DNA-damaging agents (Unoki et al., 2019). Additionally, the ZBTB24 KO HEK293 cells displayed a reduced proliferation capacity, but this phenotype could not be rescued by re-expression of ZBTB24, suggesting that irreversible changes have occurred in these KO cells. Such adaptive changes may have also occurred in our ICF2 patient-derived fibroblast and ZBTB24 KO U2OS cells, possibly explaining the lack of a NHEJ phenotype. Consequently, NHEJ defects may "only" be observed after short-term depletion of ZBTB24 in B cells from ICF2 patients or in differentiated human cells. Interestingly, like ICF2, ICF3 and ICF4 patients also suffer from immunodeficiencies associated with hypo- or agammaglobulinemia in the presence of B cells. Moreover, recent work has implied a role for CDCA7 (ICF3) and HELLS (ICF4) in NHEJ (Unoki et al., 2019), although it is unclear whether these ICF proteins participate in the PARP1/ZBTB24-dependent pathway that drives c-NHEJ. These findings may suggest that defects in this process may be a more general

phenomenon associated with ICF syndrome, specifically with regard to the observed immunodeficiency.

ICF syndrome is associated with defects in DNA methylation manifested by hypomethylation of pericentromeric satellite repeats (Vukic and Daxinger, 2019). Thus, besides its direct role in c-NHEJ, ZBTB24 may also regulate CSR indirectly as an epigenetic modifier. ZBTB24 regulates genome-wide DNMT1-dependent DNA methylation, which has been shown to alter transcription programs (Vukic and Daxinger, 2019; Wu et al., 2016). This could potentially affect the expression of genes involved in B cell development and the transcriptional status of CSR machinery (Lee and Maeda, 2012). However, we observed normal naive B cell counts, as well as normal AID and germline transcript levels, within the *C α* part of the *IGH* locus in B cells from ICF2 patients, suggesting that these cells develop normally and can properly initiate CSR. Thus, although we cannot completely rule out epigenetic effects on the immunodeficiency in ICF2, the low switched B cell numbers and impaired production of Igs likely arises from defects in completing the CSR process, i.e., in the c-NHEJ-dependent repair of AID-induced DSBs.

ZBTB24 and PARP1 in NHEJ

The current models for NHEJ distinguish a dominant c-NHEJ pathway that is fully dependent on KU70/KU80 from a PARP1-dependent a-NHEJ pathway that becomes active only in the absence of KU70/KU80 (Wang et al., 2006). However, although PARP1 is required for a-NHEJ, this does not exclude a stimulatory role for PARP1 in c-NHEJ. Indeed, several studies reported that the loss of PARP1 activity modulates the c-NHEJ-dependent rejoining of DSBs in hamster, mouse, and human cells (Luijsterburg et al., 2016; Mitchell et al., 2009; Veuger et al., 2003). Our results corroborate and extend these observations and further support a role for PARP1 in DSB repair through c-NHEJ. However, the c-NHEJ-specific phenotypes, such as impaired random plasmid integration or XRCC4 recruitment to laser/nuclease-induced DSBs (which we observed after knockdown of ZBTB24 or PARP1), were not as strong as seen after depletion of core NHEJ factors, such as DNA-PKcs. This suggests that the PARP1-ZBTB24 axis is not essential for c-NHEJ but stimulates this process in human cells. Moreover, loss of ZBTB24 reduces NHEJ in the EJ5-GFP reporter. Because this reporter cannot discriminate between c-NHEJ and a-NHEJ, we cannot rule out the possibility that ZBTB24 might promote both c-NHEJ and a-NHEJ. An involvement in the latter repair pathway would not be surprising given its interaction with PARP1, which is required for a-NHEJ (Pines et al., 2013).

The C₂H₂ ZNF of ZBTB24 binds PAR chains

Four structurally distinct protein motifs have been characterized to mediate interactions with PAR chains: (1) a consensus of eighth interspersed basic and hydrophobic amino acid residues; (2) macro domains containing a conserved ligand-binding pocket; (3) the WWE domain that recognizes iso-ADP-ribose, which is the smallest internal structural unit of PAR; and (4) the PAR-binding zinc (PBZ) finger (Kalisch et al., 2012). Here we expand the latter category by showing that the C₂H₂ ZNF, as present in ZBTB24, is a new type of motif that mediates PAR binding. Although this motif has been suggested to predominantly bind to DNA (Najafabadi et al., 2015), we demonstrate that the eight C₂H₂ ZNFs within ZBTB24 associate with PAR chains in vitro and mediate the interaction with PARP1 in vivo. Interestingly, a recent screen for DDR factors identified >100 new proteins, many of which were ZNF-containing transcription factors that, similar to ZBTB24, were recruited to sites of laser-induced DNA damage in a PARP/PARYlation-dependent manner (Izhar et al., 2015). Further studies on these DNA damage-associated ZNF-containing proteins may reveal whether they have evolved as general PAR-binding proteins with specialized functions in the PARP-dependent DDRs. Alternatively, part of the ZNF domain in ZBTB24 has been shown to confer specificity of DNA binding (Ren et al., 2019). Thus, we cannot rule out the possibility that both PAR and DNA binding are critical determinants of ZBTB24's function during the PARP-dependent DDR.

ZBTB24 stimulates PAR synthesis and protects PAR chains

Based on its functional domains, ZBTB24 seems to lack enzymatic activity. Indeed, our work suggests that ZBTB24 has at least two noncatalytic roles: (1) it can enhance PAR synthesis by PARP1 and (2) it can bind and protect PAR chains from hydrolysis by PARG. How does ZBTB24 stimulate PAR synthesis by PARP1? Two models exist for the activation of human PARP1: the cis and trans models. In the cis model, a single PARP1 protein binds a DNA end, which triggers intramolecular interactions and conformational changes that enhance the flexibility of the catalytic domain to induce auto-PARYlation (Langelier et al., 2012). One possibility is that ZBTB24, by binding to PARP1, stimulates these intramolecular interactions and conformational changes, resulting in enhanced PARP1 activation. Alternatively, in the trans model, two PARP1 proteins dimerize at a DSB, subsequently enabling one of these PARP1 molecules to modify the catalytic domain of its interaction partner (Ali et al., 2012). BTB domains, such as those found in ZBTB24, are known to mediate dimerization between proteins (Bardwell and Treisman, 1994). It is therefore possible that ZBTB24's interaction with PARP1 and its ability to dimerize could stimulate PARP1 dimerization and its subsequent activation. Additional biochemical work will be required to reveal whether ZBTB24 promotes in cis and/or in trans activation of PARP1.

In contrast to ZBTB24's role in PARP1 activation, its contribution to PAR protection may be easier to explain. We demonstrated that ZBTB24, through its ZNF domain, directly associates with PARP1-associated PAR chains. This may sterically hinder PARG from attacking PAR chains. However, some PAR chains

are digested despite the presence of excess ZBTB24 (Fig. 5, G and H), which could be due to the highly versatile endo- and exoglycosidic activities of PARG toward PAR (Brochu et al., 1994). It may be that additional PAR-binding factors are required to provide full protection against PARG hydrolysis. These factors may for instance include one or more ZNF-containing transcription factors or DDR proteins with intrinsically disordered domains that are recruited to sites of DNA damage in a PAR-dependent manner (Altmeyer et al., 2015; Izhar et al., 2015).

We observed that at concentrations up to two times that of PARP1, ZBTB24 can only activate PARP1, whereas at more than two times the concentration of PARP1, it protects PAR chains rather than that it helps to activate PARP1 (Fig. 5, D, E, G, and H). This suggests that ZBTB24 may switch function depending on its concentration relative to PARP1. Based on this, at sites of DNA damage, we envision a scenario in which ZBTB24, after its initial recruitment, helps with the activation of PARP1 and subsequently protects the synthesized PARP1-associated PAR chains. As such it could facilitate the PARYlation-dependent interaction between the c-NHEJ ligase LIG4 and PARP1, which may either be direct through interaction of the C-terminal BRCT domain of LIG4 with PAR (Li et al., 2013), or indirect through one of known PAR-binding proteins that helps to recruit the XRCC4/LIG4 complex (Ray Chaudhuri and Nussenzweig, 2017; Teloni and Altmeyer, 2016) to promote DSB repair by c-NHEJ (Fig. 7 J).

Materials and methods

Patients

Sera and PBMCs were obtained from four ICF2 patients: p49 and p55 (Weemaes et al., 2013), p67 (Rf1225), and p71 (Rf1461; van den Boogaard et al., 2017). p49, p67, and p71 carry the same recessive mutation (Table S1). ICF2 fibroblasts were from p71. A statement of no objection for the use of anonymized patient material was obtained from the medical review ethics committee of the Leiden University Medical Center. All ICF2 patient and control material was used after informed consent.

Isolation of PBMCs and phenotyping of lymphocytes

PBMCs were obtained from patients and healthy donors by Ficoll density gradient separation. PBMCs were stored in liquid nitrogen until analysis. Thawed PBMCs were stained with the following fluorochrome-labeled antibodies against the indicated cell surface antigen: CD3 (clone UCHT1) and CD4 (13B8.2; Beckman-Coulter); CD8 (SK1), CD19 (SJ25C1), CD20 (L27), CD27 (L128), CD28 (L293), and IgM (G20-127; BD Biosciences); CCR7 (150503; R&D Systems); IgD (rabbit F(ab')₂; Dako); CD45RA (MEM-56; Invitrogen Life Technologies). DAPI was added to discriminate between live and dead cells. Samples were analyzed on a BD Biosciences LSR II flow cytometer with DIVA software.

In vitro B cell stimulation and analysis of IgG and IgA production

PBMCs (0.25 × 10⁶/well) were cultured in a flat-bottom 96-well plate in AIM-V medium supplemented with 5% FCS ultra-low IgG, penicillin/streptomycin (100 IU/ml and 100 µg/ml; Life

Technologies), 0.05 mg/ml transferrin (BioChemika), and 5 μ g/ml insulin (Sigma-Aldrich). Added stimuli were MAB89 (aCD40; 0.5 μ g/ml; Beckman-Coulter), aIgM (1 μ g/ml; Jackson Immunoresearch), CpG (ODN2006; 1 μ g/ml; InvivoGen), and IL-21 (20 ng/ml; Peprotech). Supernatants were harvested at day 7 and analyzed for IgG and IgA levels by sandwich ELISA using goat anti-human IgG or IgA (Life Technologies) for coating of the 96-well microtiter plates and alkaline phosphatase-conjugated goat anti-human IgG or IgA (Life Technologies) for detection.

In vitro naive B cell stimulation and analysis of class switching

Naive B cells were magnetically sorted from PBMCs by negative selection using the Naive B cell Isolation Kit II (Miltenyi). The purity of sorted naive B cells was >95% as assessed by flow cytometry (CD19⁺CD27⁻). Because of the sorting limitation of the kit, the sorted cells contained small CD27-IgG⁺ (0.2–3.5%) or CD27-IgA⁺ (0–0.8%) populations. The sorted cells were stimulated with CD40 ligand (MEGACD40L; 100 ng/ml; Enzo), IL-21 (100 ng/ml), IL-10 (50 ng/ml; Peprotech), and anti-IgM (10 μ g/ml). For analysis of IgA and IgG production, an anti-IgM concentration of 1–10 μ g/ml (10 μ g/ml for flow cytometry and RNA analysis, 1 μ g/ml for ELISA) was used. The cells were cultured for 6–10 d in round-bottom 96-well plates (5–10 \times 10⁴/well) in RPMI 1640 supplemented with 10% FCS, penicillin/streptomycin, 50 μ M β -mercaptoethanol, 2.5 μ g/ml transferrin (BioChemika), 1 μ g/ml insulin, and nonessential amino acids (Gibco). After 6 d, class switching of the stimulated cells was analyzed by flow cytometry with antibodies against CD19 (Beckman Coulter), IgG (BD Biosciences), and IgA (Miltenyi). DAPI was added to discriminate between live and dead cells. Samples were analyzed on a BD Biosciences Canto II flow cytometer with DIVA software. On day 6, RNA samples were extracted from the cultured cells (RNeasy Micro Kit; Qiagen) and subjected to reverse transcription with SuperScript II Reverse Transcription (Invitrogen). On day 10, supernatants were collected, and IgG and IgA levels were analyzed by ELISA as described above.

Expression of AID by real-time quantitative RT-PCR (RT-qPCR)

AID expression in sorted naive B cells was performed as described previously (Cagigi et al., 2009). Briefly, RNA was isolated from unstimulated and stimulated naive B cells after 6 d of culture using the RNeasy Micro Kit (Qiagen) and subjected to reverse transcription with SuperScript II Reverse transcription (Invitrogen). 2 \times GoTaq qPCR Master Mix (Promega) was used together with the previously described primers for amplification of AID transcripts (Cagigi et al., 2009). AID expression was normalized to the housekeeping gene GUSB (Table S5).

Amplification of $I\alpha$ -Ca germline transcripts

$I\alpha$ -Ca germline transcripts were assessed as described previously (Lin et al., 2014). Briefly, RNA was isolated from unstimulated and stimulated sorted naive B cells after 6 d of culture using the RNeasy Micro Kit (Qiagen) and subjected to reverse transcription with SuperScript II Reverse transcriptase (Invitrogen). KAPA HiFi HotStart ReadyMix (Roche) together with previously described $I\alpha$ -consensus, Ca1-specific, and Ca2-

specific primers (Lin et al., 2014) were used to amplify the $I\alpha$ 1-Ca1 and $I\alpha$ 2-Ca2 germline transcripts. PCR amplification was performed using 40 cycles at 95°C for 30 s, 68°C for 30 s, and 72°C for 1 min.

Sequencing of switch recombination junctions

Amplification, cloning, and sequencing of the S μ -S α or S μ -S γ fragments derived from PBMCs was performed using a previously described PCR strategy (Pan-Hammarström et al., 2005). The CSR junctions were determined by aligning the switch fragment sequences with the reference S μ , S α , or S γ sequences. Analysis of the repair pattern of the CSR junctions was performed based on the suggested guidelines (Stavnezer et al., 2010).

Ig heavy chain (IgH) repertoire analysis using next-generation sequencing

The VH-JH rearrangements and Ca and C γ transcripts were amplified from post-Ficoll PBMCs in a multiplex PCR using the VH1-6 FR1 and JH consensus BIOMED-2 primers (van Dongen et al., 2003) or a consensus Ca (IGHA-R; 5'-CTTTCGCTCCAGGTCACACTGAG-3') and C γ primer (3'-C γ -CHI; (Tiller et al., 2008). The primers were adapted for 454 sequencing by adding the forward A or reverse B adaptor, the TCAG key and multiplex identifier (MID) adaptor. PCR products were purified by gel extraction (Qiagen) and Agencourt AMPure XP beads (Beckman Coulter). DNA concentration was measured using the Quant-it Picogreen dsDNA assay (Invitrogen). Purified PCR products were sequenced on the 454 GS junior instrument (Roche) according to the manufacturer's recommendations, using the GS Junior Titanium emPCR (Lib-A), GS Junior Titanium sequencing, and PicoTiterPlate kits for the VH-JH rearrangements, and the GS Junior+ emPCR (Lib-A), GS Junior sequencing XL+, and PicoTiterPlate kits for the Ca and C γ transcripts. Using the IGGalaxy Tool (Moorhouse et al., 2014), sequences were demultiplexed based on their MID sequence and quality checked. FASTA files were uploaded in IMGT HighV-Quest (<http://www.imgt.org>). Further analysis of the data was done using the IGGalaxy tool. Uniqueness of sequences was defined by V, D, and J gene usage and nucleotide sequence of the CDR3 region for the VH-JH rearrangements, and V gene usage, amino acid sequence of the CDR3 region, and C gene usage for the Ca and C γ transcripts. Only unique, productive sequences were used for the analysis, and the frequency of mutated nucleotides in the V_H gene was calculated from CDR1 until FR3.

Cell culture

U2OS, HEK293, HEK293T, HeLa Flp-In/T-Rex, VH10-SV40-immortalized fibroblasts, and SV40 T-transformed GM639 human fibroblasts were grown in DMEM (Gibco) containing 10% FCS (Bodinco BV) and 1% penicillin/streptomycin unless stated otherwise, whereas CH12 cells were grown in RPMI 1640 (Gibco) supplemented with 10% FCS. U2OS 2-6-3 cells containing 200 copies of a LacO-containing cassette (~4 Mbp) were gifts from Dr. J. Lukas (University of Copenhagen, Copenhagen, Denmark) and Dr. S. Janicki (The Wistar Institute, Philadelphia, PA; Doil et al., 2009; Shanbhag et al., 2010) and were used to establish

U2OS 2-6-3 cell lines stably expressing GFP-tagged XRCC4. Single U2OS clones stably expressing GFP-XRCC4 were isolated after selection on puromycin (1 mg/ml). Immunoblotting with anti-GFP antibody showed that the XRCC4 fusion proteins were expressed at the expected molecular weight. U2OS 2-6-3 cells stably expressing ER-mCherry-LacR-FokI-DD, which were a gift from Dr. R. Greenberg (University of Pennsylvania, Philadelphia, PA; [Tang et al., 2013](#)), were induced for 5 h by 1 μ M Shield-1 (Clontech) and 1 μ M 4-OHT (Sigma-Aldrich). SV40 T-transformed GM639 human fibroblasts with a stably integrated GC92 reporter (GC92 cells) were a gift from Bernard Lopez (Université de Paris, Paris, France; [Taty-Taty et al., 2016](#)) and were used to study mutational signatures at repair junctions. ZBTB24 KO U2OS cells were generated by transfection of pSpCas9(BB)-2A-GFP (PX458; Addgene 48138) containing Cas9 and a gRNA against ZBTB24 (5'-AGATCCTCTTGCTGAACA-3'), which was cloned into the BbsI site. 48 h after transfection, cells were sorted by flow cytometry for GFP expression and seeded at low density, after which individual clones were isolated. Knockout of ZBTB24 in U2OS cells was first verified by Sanger sequencing and TIDE analysis (<https://tide.nki.nl>). Clones harboring out-of-frame deletions were further verified by Western blot analysis. HeLa Flp-In/T-REx cells, which were generated using the Flp-In/T-REx system (Thermo Fisher Scientific), were a gift of Geert Kops (University Medical Centre Utrecht, Utrecht, Netherlands) and Stephen Taylor (Washington University, St. Louis, MO). These cells were used to generate stable cells expressing inducible and siZBTB24-8-resistant versions of GFP-ZBTB24 and GFP-ZBTB24 Δ ZNF by cotransfection of pCDNA5/FRT/TO-Puro plasmid encoding GFP-ZBTB24 siZBTB24-8-res or GFP-ZBTB24 Δ ZNF siZBTB24-8-res, together with pOG44 plasmid encoding the Flp recombinase. After selection on 1 μ g/ml puromycin, single clones were isolated and expanded. Stable HeLa Flp-In/T-REx clones were incubated with 2 μ g/ml doxycycline for 24 h to induce expression of cDNAs. Additionally, GFP-tagged ZBTB24 Δ ZNF expression was reduced to endogenous ZBTB24 levels by repeated washout during 146 h. To generate Zbtb24^{-/-} CH12 clones, cells were transfected by electroporation using the Neon transfection System (Thermo Fisher Scientific) with a plasmid expressing a gRNA targeting the first exon of mZbtb24 (5'-AAGCTGCCACAAGGCTCCG-3') and coexpressing the high-fidelity Cas9 nuclease ([Kleinstiver et al., 2016](#)) fused to EGFP. 24 h after transfection, individual EGFP-positive cells were sorted in 96-well plates and cultured for 10 d. Clones were then genotyped by PCR, sequencing, and Western blot.

Plasmids

The full-length cDNA of human ZBTB24 was obtained by RT-PCR and flanking restriction sites for conventional cloning (BglII/Sall) were introduced using a nested PCR on the cDNA. The obtained PCR product was subsequently cloned into pEGFP-C1 and pEGFP-N1 (both Clontech) using the BglII and Sall restriction sites. The GST-ZBTB24 expression vector was generated by cloning the ZBTB24 ORF from pEGFP-C1-ZBTB24 as a BglII/EcoRI fragment into BamHI/EcoRI-digested pGEX-6p-3 (GE Healthcare). The Myc-ZBTB24 expression vector was obtained by exchanging GFP, using the AgeI and BglII restriction

sites, for a single Myc tag (EQKLISEEDL) by oligo annealing in the pEGFP-ZBTB24 construct. Deletion constructs were generated by amplifying the specified regions using internal primers containing BglII (forward) or EcoRI (reverse) restriction sites and subsequent exchange of the deletion fragments for the full-length cDNA. pCDNA5/FRT/TO-Puro plasmids encoding GFP-ZBTB24 siZBTB24-8-res or GFP-ZBTB24 Δ ZNF siZBTB24-8-res were generated by cloning GFP-ZBTB24 or GFP-ZBTB24 Δ ZNF fragments into pCDNA5/FRT/TO-Puro. The underlined mutations 5'-CGAAAAGAGCACCGAGCAA-3' were introduced by PCR to generate resistance against siZBT24-8: 5'-UGAGAAAAGUACAGAACAA-3'. All ZBTB24 expression constructs were verified using Sanger sequencing. The murine Zbtb24 cDNA was amplified by PCR from a cDNA library prepared from CH12 cells using standard techniques and cloned into the pMX-PIE plasmid ([Barreto et al., 2003](#)) using BamHI and NotI restriction enzymes. mCherry-PARG wt/cd were kindly provided by Michael Hendzel ([Ismail et al., 2012](#)) and GFP-PARP1 was obtained from Valerie Schreiber (Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch, France; [Mortusewicz et al., 2007](#)). The XRCC4 cDNA, a generous gift of P. Jeggo (School of Life Sciences, East Sussex, UK; [Girard et al., 2004](#)), was inserted into EGFP-C1-IRES-Puro.

Transfections, RNA interference, and retroviral transductions

siRNA and plasmid transfections were performed using Lipofectamine RNAiMAX (Invitrogen), Lipofectamine 2000 (Invitrogen), and JetPEI (Polyplus Transfection), respectively, according to the manufacturer's instructions. siRNA sequences are listed in Table S5. Cells were transfected twice with siRNAs (40 or 80 nM) within 24 h and examined further 48 h after the second transfection unless stated otherwise. PARP inhibitor (KU-0058948) was a gift from Mark O'Connor (AstraZeneca, Cambridge, UK) and was used at a concentration of 10 μ M. The DNA-PK inhibitor (NU7026; EMD Biosciences) was used at a concentration of 10 μ M. CH12 cells were transduced with retroviral supernatants obtained by transfecting Bosc23 cells with an empty retrovirus (pMX-PIE; Puromycin-IRES-EGFP) or a retrovirus expressing mZbtb24 and EGFP cDNA (pMX-mZbtb24) as described previously ([Barreto et al., 2003](#)). Transduced cells were then selected with puromycin (1 μ g/ml) for 10 d.

NHEJ reporter assays

HEK293 cell lines containing a stably integrated copy of the EJ5-GFP reporter or SV40 T-transformed GM639 human fibroblasts containing a stably integrated copy of the GC92 reporter were used to measure the repair of I-SceI-induced DSBs or NHEJ ([Bennardo et al., 2008](#); [Pierce et al., 1999](#); [Taty-Taty et al., 2016](#)). Briefly, 48 h after siRNA transfection, cells were transfected with the I-SceI expression vector pCBASce and an mCherry expression vector. 48 h later, the fraction of GFP-positive cells or CD4-FITC-positive cells among the mCherry-positive cells was determined by FACS on a BD LSRII flow cytometer (BD Bioscience) using FACSDiva software version 5.0.3. Quantifications were performed using Flowing software 2.5.1 (by Perttu Terho in collaboration with Turku Bioimaging).

Analysis of repair junctions in the GC92 reporter

Sequence analysis of repair junctions in the GC92 reporter was performed as described (Taty-Taty et al., 2016). Briefly, GC92-containing fibroblasts (GC92 cells) were first transfected with siRNAs and 48 h later with the I-SceI expression vector pCBASce (Pierce et al., 1999). 48 h later, genomic DNA was extracted using phenol:chloroform:isoamyl alcohol (25:24:1 vol/vol; Invitrogen). PCR was performed on the genomic DNA using the CMV1 (5'-TGGCCGCCTGGCATTATGCC-3') and CD4int (5'-GCTGCCCA GAATCTTCCTCT-3') primers to amplify repair junctions. PCR products were cloned into pGEM-T easy vector (Promega). Colony PCR was performed using M13 primers (M13 FW 5'-GTA AAACGACGGCCAGT-3' and M13 RV 5'-CAGGAAACAGCTATG AC-3') on individual bacterial colonies to amplify repair junctions, which were subjected to Sanger sequencing using the M13 FW primer. Sequences were analyzed using a custom Sanger sequence analyzer as described previously (Schimmel et al., 2017).

Plasmid integration assay

Upon siRNA-mediated knockdown of the indicated genes, U2OS cells were transfected with XhoI/BamHI-linearized pEGFP-C1 plasmid DNA. After overnight transfection, a fraction of the cells was used to determine transfection efficiency, which was measured by the amount of GFP-positive cells using the Array-Scan high content analysis reader (Thermo Fisher Scientific) with the target activation protocol. In parallel, cells were seeded on 14-cm plates at a density of 10,000 and 2,000 cells per plate for determination of the cloning efficiency with and without G418 (0.5 mg/ml; Gibco) selection, respectively. After 10 d, cells were washed in 0.9% NaCl and stained with methylene blue. NHEJ efficiency was calculated as follows: (cloning efficiency G418 selection)/[(cloning efficiency without selection) × (transfection efficiency)] and subsequently normalized to the luciferase control.

CSR assay in CH12 cells

CH12 cells were cultured for 72 h in the presence of TGF- β (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 efficiency by flow cytometry. Before analysis, DAPI was added to discriminate dead cells. Samples were analyzed using an LSR flow cytometer (BD Biosciences) and FlowJo software.

Cell survival assay

VH10-SV40 cells were transfected with siRNAs, trypsinized, seeded at low density, and exposed to IR at indicated doses. 7 d later, cells were washed with 0.9% NaCl and stained with methylene blue. Colonies of >10 cells were counted, and relative survival compared with the untreated sample was calculated.

Cell cycle profiling

For cell cycle analysis, cells were treated as described in the figure legends and fixed in 70% ethanol, followed by DNA staining with 50 μ g/ml propidium iodide in the presence of RNase A (0.1 mg/ml). Cell sorting was performed on a BD LSRII flow cytometer (BD

Biosciences) using FACSDiva software (version 5.0.3; BD Biosciences). Quantification was performed using Flowing software 2.5.1.

RNA expression analysis by RT-qPCR and RNA sequencing

Gene expression analysis using RT-qPCR was performed as described before (Helfricht et al., 2013). Briefly, RNA isolation was done using the miRNeasy minikit (Qiagen), and subsequently polydT-primed cDNA was generated using the RevertAid first strand cDNA synthesis kit (Thermo Fisher Scientific) according to the manufacturer's instructions. RT-qPCR was performed in duplicate on the CFX96/384 system using SYBR green master mix (Bio-Rad). Primers, which are listed in Table S5, were designed using Primer3Plus software (<http://primer3plus.com>). Relative expression levels were obtained with the CFX manager (version 3.0), correcting for primer efficiencies and using GAPDH and GUSB as reference genes. For RNA sequencing, the RNA 6000 Nano kit (Agilent Technologies) was used to confirm RNA integrity before the RNA was subjected to poly(A) enrichment. cDNA synthesis, library preparation, and sequencing were performed using the Ion Total RNA-Seq kit v2, the Ion PI Template OT2 200 Kit v3, and the Ion Sequencing 200 kit v3, respectively, according to the manufacturer's instructions (Thermo Fisher Scientific). RNA was sequenced on an Ion Proton System at a depth of ~20 million reads per sample, with a median read length of 90 bp. Sequence files obtained in the bam format were converted to fastq using the bam2fastq conversion utility from the bedtools package. Reads were aligned to the human genome build GRCh37 - Ensembl using Tophat2 (version 2.0.10). In a second alignment step, Bowtie2 (version 2-2.10) was used in the local, very sensitive mode to align remaining unaligned reads. HTSeq-Count (version 0.6.1) was used with default settings to quantify gene expression. Finally, DESeq (version 1.2.10) was used to generate a list of genes differentially expressed between ZBTB24-depleted and control cells (Table S2). The data have been deposited to the SRA database with the accession number PRJNA556576.

Sample preparation and MS

For SILAC, U2OS cells were cultured for 14 d in light (L; [$^{12}\text{C}_6,^{14}\text{N}_2$]lysine/[$^{12}\text{C}_6,^{14}\text{N}_4$]arginine) or heavy (H; [$^{13}\text{C}_6,^{15}\text{N}_2$]lysine/[$^{13}\text{C}_6,^{15}\text{N}_4$]arginine) SILAC medium. SILAC-labeled cells were transiently transfected with either GFP-PARP1 or GFP-ZBTB24 (H) and an empty vector (L). Equal amounts of H and L cells were lysed separately in EBC-150 buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.5% NP-40, and 1 mM EDTA) supplemented with protease and phosphatase inhibitor cocktails. The lysed cell suspension was sonicated six times for 10 s on ice and subsequently incubated with 500 U Benzonase for 1 h under rotation. The NaCl concentration was increased to 300 mM, and the cleared lysates were subjected to GFP IP with GFP Trap beads (Chromotek). The beads were then washed twice with EBC-300 buffer (50 mM Tris, pH 7.5, 300 mM NaCl, 0.5% NP-40, and 1 mM EDTA) and twice with 50 mM $(\text{NH}_4)_2\text{CO}_3$ followed by overnight digestion using 2.5 μ g trypsin at 37°C under constant shaking. Peptides of the H and L precipitates were mixed and desalted using a Sep-Pak tC18 cartridge by washing with 0.1%

acetic acid. Finally, peptides were eluted with 0.1% acetic acid/60% acetonitrile and lyophilized. Samples were analyzed by nanoscale liquid chromatography–MS/MS using an EASY-nLC system (Proxeon) connected to a Q-Exactive Orbitrap (Thermo Fisher Scientific). Peptides were separated in a 13-cm analytical column with inner diameter of 75 μm , in-house packed with 1.8 μm C18 beads (ReproSpher; Dr. Maisch). The gradient length was 120 min with a flow rate of 200 nl/min. Data-dependent acquisition was used with a top 10 method. Full-scan MS spectra were acquired at a target value of 3×10^6 and a resolution of 70,000, and the higher-collisional dissociation tandem mass spectra (MS/MS) were recorded at a target value of 10^5 and with resolution of 17,500 with a normalized collision energy of 25%. The precursor ion masses of scanned ions were dynamically excluded from MS/MS analysis for 60 s. Ions with charge 1 and >6 were excluded from triggering MS2 events (Hendriks et al., 2014). Analysis of raw data was performed using MaxQuant software version 1.4.1.2 (Cox and Mann, 2008). The data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD014741.

Protein interaction studies

To study ZBTB24 interactions, cells expressing the indicated GFP fusion proteins were lysed in 1 ml EBC buffer (50 mM Tris, pH 7.3, 150 mM NaCl, 0.5% NP-40, and 2.5 mM MgCl) supplemented with protease and phosphatase inhibitor cocktails (Roche). Lysis and protein extraction were enhanced by 6×10 -s sonication in a sonicator bath (Bioruptor UCD-20; Diagenode) followed by 1-h incubation with 500 units benzonase (Novagen) on ice. Upon centrifugation, cleared lysates were subjected to IP with GFP Trap beads (Chromotek) for 1.5 h at 4°C top over top. Beads were washed six times with cold EBC buffer and boiled in Laemmli buffer, and interacting proteins were visualized using Western blot analysis.

Western blot analysis

Protein extracts were generated by direct lysis of cells in $2 \times$ Laemmli buffer and boiled for 10 min at 95°C. Proteins were size separated using Novex 4–12% Bis-Tris mini gels (Invitrogen) or 4–12% Criterion XT Bis-Tris gels (Bio-Rad) in $1 \times$ MOPS buffer (Invitrogen) and transferred to PVDF membranes, which were blocked in 4% milk for ≥ 30 min and incubated with the indicated antibodies overnight. Several wash steps before and after 1-h incubation with secondary antibodies rabbit-anti-700 and mouse-anti-800 (Sigma-Aldrich) were executed. Protein bands were visualized using the Odyssey infrared imaging system or the C-Digit blot scanner (both Licor) according to the manufacturer's instructions. Representative Western blot images of two to five independent experiments are shown.

Laser microirradiation

Multiphoton laser microirradiation was performed with a Leica SP5 confocal microscope equipped with an environmental chamber set to 37°C and 5% CO₂ as described (Helfricht et al., 2013). Briefly, U2OS or HeLa Flp-In/T-Rex cells were grown on 18-mm glass coverslips, and medium was replaced with colorless DMEM or CO₂-independent Leibovitz L15 medium, both

supplemented with 10% FCS and penicillin/streptomycin. Cells were placed in a ChamSlide TC-A live-cell imaging chamber before imaging and were kept at 37°C. DSB-containing tracks (1- or 1.5- μm width) were generated with a Mira modelocked Ti:Sapphire laser ($\lambda = 800$ nm, pulselength = 200 fs, repetition rate = 76 MHz, and output power = 80 mW). Typically, cells were microirradiated with 1 iteration/pixel using LAS-AF software. For live-cell imaging, confocal images were recorded before and after laser irradiation at different time intervals. For UV-A laser microirradiation, U2OS or HeLa Flp-In/T-Rex cells were sensitized with 10 μM BrdU for 24 h, as described (Helfricht et al., 2013). For microirradiation, the cells were placed on the stage of a Leica DM IRBE wide-field microscope stand (Leica) integrated with a pulsed nitrogen laser (Micro-point Ablation Laser System; Photonic Instruments; 16 Hz, 364 nm), which was directly coupled to the epifluorescence path of the microscope and focused through a Leica 40 \times HCX Plan Apo 1.25–0.75 oil-immersion objective. The laser output power was set to 78 to generate strictly localized subnuclear DNA damage, and images were taken before and after microirradiation at the indicated time points or after immunofluorescent labeling using Andor IQ software.

Immunofluorescent labeling

Immunofluorescent labeling of γH2AX , XRCC4, and GFP was performed as described previously (Helfricht et al., 2013). Briefly, cells were grown on glass coverslips and treated as indicated in the figure legends. Subsequently, cells were washed with PBS, fixed with 4% formaldehyde for 15 min and treated with 0.25% Triton X-100 in PBS for 5 min. Cells were rinsed with PBS and equilibrated in WB (PBS containing 5 g BSA/liter and 1.5 g glycine/liter) before immunostaining. Detection was done using goat anti-mouse or goat anti-rabbit IgG coupled to Alexa Fluor 488, 555, or 647 (Invitrogen Molecular probes). Samples were incubated with 0.1 $\mu\text{g/ml}$ DAPI and mounted in Polymount.

Microscopy analysis

Images of fixed samples were acquired on a Zeiss AxioImager M2 or D2 wide-field fluorescence microscope equipped with 40 \times , 63 \times , and 100 \times Plan Apo (1.4-NA) oil-immersion objectives (Zeiss) and an HXP 120 metal-halide lamp used for excitation. Fluorescent probes were detected using the following filters: DAPI (excitation filter, 350/50 nm; dichroic mirror, 400 nm; emission filter, 460/50 nm), GFP/Alexa Fluor 488 (excitation filter, 470/40 nm; dichroic mirror, 495 nm; emission filter, 525/50 nm), mCherry (excitation filter, 560/40 nm; dichroic mirror, 585 nm; emission filter, 630/75 nm), Alexa Fluor 555 (excitation filter, 545/25 nm; dichroic mirror, 565 nm; emission filter, 605/70 nm), and Alexa Fluor 647 (excitation filter, 640/30 nm; dichroic mirror, 660 nm; emission filter, 690/50 nm). Images recorded after multiphoton- and UV-A-laser microirradiation and immunofluorescence stainings were analyzed using ImageJ (National Institutes of Health). The average pixel intensity of laser tracks induced by either the multiphoton- or the UV-A laser system was measured within the locally irradiated area (I_{damage}), in the nucleoplasm outside the locally irradiated area

($I_{\text{nucleoplasm}}$), and in a region not containing cells in the same field of view ($I_{\text{background}}$) using ImageJ. The relative level of accumulation expressed relative to the protein level in the nucleoplasm was calculated as $[(I_{\text{damage}} - I_{\text{background}})/(I_{\text{nucleoplasm}} - I_{\text{background}}) - 1]$. The accumulation in the control cells transfected with siLuc within each experiment was normalized to 100%. Images obtained from live-cell imaging after multiphoton microirradiation were analyzed using LAS-AF software. Fluorescence intensities were subtracted by the prebleach values and normalized to the first data point, which was set to 0, to obtain relative fluorescence units. The average reflects the quantification of between 50 and 150 cells from two to three independent experiments.

Antibodies

Immunofluorescence, Western blot, and flow cytometry analysis were performed using antibodies against GFP (1:1,000, 11814460001, Roche; or 1:1,000, ab290, Abcam), PARP1 (1:1,000, 9542, Cell Signaling, Alexis), Myc (1:1,000, 9E10, SC-40, Santa Cruz), γ H2AX (1:1,000, 07-164, Millipore), α -tubulin (Sigma-Aldrich), DNA-PKcs (1:500, ab1832, Abcam), LIG4 (1:1,000, 80514, Abcam), XRCC4 (1:500, gift from Mauro Modesti, Marseille Cancer Research Center, Marseille, France), histone H3 (1:2,000, 1791, Abcam), GST (1:2,000, Amersham), PARP1 (1:1,000, 9542S, Cell Signaling), PARP2 (1:500, C3956, Sigma-Aldrich), ZBTB24 (1:1,000, PM085, MBL), CDCA7 (1:250, ProteinTech), RAD51 (1:2,000, sc-6862, Santa Cruz), CD4-FITC (1:100, 100509, BioLegend), β -actin (1:2,000, AC15, Sigma-Aldrich), PAR (1:1,000, 4336-BPC-100, Trevigen; used in Fig. 5, A and B), PAR monoclonal 10H, which was purified from the culture medium of 10H hybridoma obtained from Dr. Miwa (Nagahama Institute of Bio-Science and Technology, Nagahama, Japan) through the Riken cell ban (Kawamitsu et al., 1984), and custom-made monoclonal AID (Jeevan-Raj et al., 2011).

GST protein purification

For GST purification, 50-ml cultures of *Escherichia coli* BL21 cells containing pGEX-6p-3 or pGEX-6p-3-ZBTB24 plasmid were grown to an OD₆₀₀ of 0.6 absorbance units. 2 mM IPTG was added, and cells were incubated overnight at 20°C. After centrifugation, cell pellets were frozen and stored at -80°C. For protein purification, cell pellets were lysed at room temperature for 30 min in 2.5 ml lysis buffer (125 mM Tris-HCl, pH 8, 150 mM NaCl, 1 mM MgCl₂, 5 mM DTT, 0.1 volume BugBuster 10 \times [Novagen-Merck], 2,500 units rLysozyme [Novagen-Merck], 62.5 units benzonase [Novagen-Merck], and Protease Inhibitor Cocktail EDTA-free [Sigma-Aldrich]). The lysate was centrifuged at 4°C in a table centrifuge for 10 min at full speed. Supernatant was taken and incubated with 500 μ l Glutathione Superflow Agarose beads (Life Technologies) for 2 h at 4°C. The agarose beads were packed in a column and loaded on an ÄKTA chromatography system (GE Healthcare Biosciences). The column was rinsed using wash buffer (125 mM Tris-HCl, pH 8, 150 mM NaCl, and 10 mM β -mercaptoethanol) and eluted using wash buffer supplemented with 10 mM reduced glutathione (Sigma-Aldrich). Fractions with purified protein were collected and concentrated using 50-kD Vivaspun ultrafiltration cups

(Sartorius). Finally, the buffer was changed in ultrafiltration cups to 125 mM Tris-HCl, pH 8, 150 mM NaCl, and 10% glycerol, and purified proteins were frozen in liquid nitrogen and stored at -80°C.

Analysis of protein PARylation

Cells were washed with ice-cold PBS supplemented with PARG inhibitor (PARGi; 400 nM tannic acid), scraped in a small volume of PBS with PARGi, and transferred to low binding tubes, followed by high-speed centrifugation at 4°C. Cells were lysed in RIPA buffer (10 mM Tris-HCl, pH 8, 1% Triton X-100, 0.1% deoxycholate, 0.1% SDS, and 100 μ M tannic acid) supplemented with protease and phosphatase inhibitor cocktails (Roche) comprising a NaCl-concentration of 450 mM. After centrifugation, cleared lysates were subjected to IP with GFP Trap beads (Chromotek) for 2 h on a rotating wheel in the presence of 150 mM NaCl. Beads were washed six times with RIPA buffer containing increasing NaCl concentrations (150 mM and 1 M) followed by two washes with TBS-T buffer (20 \times TBS, 0.1% Tween, and 100 μ M tannic acid). After boiling in Laemmli buffer, the interacting proteins were visualized using Western blot analysis.

Production of radiolabeled PAR

PARP1 activation assays were performed as described earlier (Shah et al., 2011) with minor modifications. To prepare radiolabeled PAR, purified bovine PARP1 was activated at 30°C for 30 min in 900 μ l reaction mix (100 mM Tris-HCl, pH 8.0, 10 mM MgCl₂, 10% glycerol, 10 mM DTT, 500 μ M cold NAD, 250 μ Ci of ³²P-NAD [350 nM], 10% ethanol, and 23 μ g activated calf thymus DNA). Auto-PARYlated PARP1 was precipitated on ice for \geq 30 min by addition of 100 μ l of 3 M Na-acetate, pH 5.2, and 700 μ l isopropanol. After centrifugation, the pellet was washed twice with ethanol, air-dried, and dissolved (1 M KOH and 50 mM EDTA), while heating at 60°C for 1 h. Upon addition of AAGE9 (250 mM NH₄OAc, 6 M guanidine-HCl, and 10 mM EDTA), pH was adjusted to 9.0, and solution was loaded onto DHBB resin in Econocolumns (Bio-Rad). Resin was washed with AAGE9 and NH₄-acetate, pH 9.0. The polymer was eluted with water at 37°C in separate fractions and stored at -30°C until usage in Southwestern assays.

Southwestern assay

The Southwestern assay was performed as described (Robu et al., 2013). Briefly, IP samples were resolved on 8% denaturing PAGE gels along with purified human PARP1 (Aparptosis) as a positive control. Gels were incubated for 1 h with gentle agitation in SDS-PAGE running buffer (20–30 ml of 25 mM Tris, pH 7.5, 192 mM glycine, 5% β -mercaptoethanol, and 0.1% SDS), followed by protein transfer to a nitrocellulose membrane at 4°C. Membrane were rinsed three times with TST buffer (10 mM Tris, pH 7.5, 150 mM NaCl, and 0.05% Tween) and incubated in 20 ml TST buffer supplemented with 250 nM radioactive PAR polymer on a shaker at room temperature for 1 h, followed by three washes with TST and one wash with TST buffer containing 500 mM NaCl. After a final wash with regular TST, membranes were dried and exposed to either a film or a phosphorimager

screen to detect radioactivity. Afterward, membranes were blocked in 5% milk containing 0.1% Tween and probed for PARP and GFP with the indicated antibodies.

PARP1 activation assays

To examine the stimulatory effect of ZBTB24 on the catalytic activity of PARP1, PARP1 activation reactions were performed in a 20- μ l assay volume with 0.4 pmol of PARP1, 160 ng activated DNA, and 100 μ M unlabeled NAD at 30°C for 10 min with no other protein (control) or varying molar ratios of GST-ZBTB24, GST-ZBTB24 Δ ZNF, or GST over PARP1. The reactions were stopped by the addition of equal volumes of 2 \times Laemmli buffer. Aliquots from each sample were resolved on 6 or 10% SDS-PAGE followed by immunoblotting for PAR, PARP1, and GST.

PAR protection assays

To examine the effect of ZBTB24 on PAR protection, PARP1 activation reactions were performed in a 15- μ l assay volume with 4 pmol of PARP1, 3 μ g of activated and 100 μ M unlabeled NAD at 30°C for 30 min to allow the formation of autoPARylated PARP1. The reaction was stopped by the addition of 1 μ l of 1 mM PARPi (PJ-34). One-tenth of the reaction mixes containing 0.4 pmol of PARP1 were reacted for 15 min with no other protein (control) or varying molar ratios of GST-ZBTB24, GST-ZBTB24 Δ ZNF, or GST over PARP1. All samples were reacted at 30°C for 15 min in the PARG-assay buffer (50 mM Tris-HCl, pH 7.5, containing 50 mM KCl, 1.5 mM DTT, 0.1 mg/ml BSA, and 2.5 mM EDTA) with 5 ng PARG (Sigma-Aldrich), whereas the undigested PAR samples were mock-treated with PARG assay buffer. The reactions were stopped by the addition of equal volumes of 2 \times Laemmli buffer. Aliquots from each sample were resolved on 6 or 10% SDS-PAGE followed by immunoblotting for PAR, PARP1, and GST.

Online supplemental material

Fig. S1 shows the differentiation of T cells from ICF2 patients; combinatorial diversity; and junction characteristics of IgH rearrangements in ICF2 patients. **Fig. S2** shows that knockdown of ZBTB24 affects neither cell cycle progression nor the expression of genes involved in DSB repair; ICF2 patient-derived fibroblasts and ZBTB24 KO U2OS cells are not sensitive to IR. **Fig. S3** shows that PARG-dependent turnover of PAR chains modulates the accumulation of ZBTB24 at sites of DNA damage; the ZNF domain of ZBTB24 accumulates at sites of DNA damage in a PARP activity-dependent manner; ZBTB24 is not PARylated after DNA damage induction. **Fig. S4** shows the purification of recombinant ZBTB24; PARP1 promotes XRCC4/LIG4 assembly and NHEJ at DNA damage sites. **Fig. S5** shows an analysis of HeLa Flp-In/T-REx cells expressing siRNA-resistant GFP-ZBTB24 or GFP-ZBTB24 Δ ZNF; GFP-ZBTB24, but not GFP-ZBTB24 Δ ZNF accumulates at sites of DNA damage. Table S1 shows the serum Ig isotype concentrations of ICF2 patients at first analysis. Table S2 lists ZBTB24-regulated genes identified by RNA sequencing. Table S3 lists proteins identified as ZBTB24 interactors by SILAC-based MS. Table S4 lists proteins identified as PARP1 interactors by SILAC-based MS. Table S5 contains sequences of RT-qPCR primers and siRNAs. Data S1 contains the $\Sigma\mu$ -S α junctions from ICF2 patients. Data S2 contains the $\Sigma\mu$ -S α

junctions from healthy donors. Data S3 contains the $\Sigma\mu$ -S γ junctions from ICF2 patients.

Acknowledgments

The authors thank Nisha Verweij, Jer-gung Chang, Andrea Björkman, and Steve Jackson for help with cloning experiments, MS sample analysis, CSR junction analysis, and plasmid integration assays, respectively, and Robin van Schendel and Marcel Tijsterman for the custom Sanger Sequence analyzer and help with the sequence analysis. Geert Kops, Stephen Taylor, Penny Jeggo, Michael Hendzel, Valerie Schreiber, and Caner Aytekin kindly provided cell lines, plasmids and ICF2 patient cells.

This work was financially supported by grants from the Dutch Scientific Organization (NWO-VENI 863.11.007 to M.S. Luijsterburg; NWO-VIDI 91712323 to M. van der Burg; and NWO-VICI VI.C.182.052 to H. van Attikum), People Program - Marie Curie Actions (317246 to P. Rao), Institut National du Cancer (2019-1-PL BIO-04-1 to B. Reina-San-Martin), Fondation pour la Recherche Médicale (EQU201903007818_REINA SAN MARTIN Equipes FRM 2019 to B. Reina-San-Martin), French state fund managed by the Agence Nationale de la Recherche under the frame program Investissements d'Avenir ANR-10-IDEX-0002-02 (ANR-10-LABX-0030-INRT to B. Reina-San-Martin), La Ligue Contre le Cancer (M. Rogier), French Ministère de l'Enseignement Supérieur et de la Recherche and the Fondation ARC (J. Moritz), Natural Sciences and Engineering Research Council of Canada (RGPIN 2016-05868 to G.M. Shah), Dutch Cancer Society (UL2012-5460 to S.M. van der Maarel and H. van Attikum), European Research Council (Starting grant 310913 to A.C.O. Vertegaal, Starting grant 242551 to Q. Pan-Hammarström, and Consolidator grant 617485 to H. van Attikum).

Author contributions: A. Helfricht performed UV-A micro-irradiation experiments, clonogenic survivals, and EJ5-GFP reporter assays; generated cell cycle profiles; and analyzed data. P.E. Thijssen performed co-IPs, RT-qPCR, plasmid integration assays, RNA sequencing, and survival on ICF2-derived fibroblasts and analyzed data. M.B. Rother performed UV-A micro-irradiation experiments, GC92 reporter assays, Western blot analysis, and AID and germline expression analysis and analyzed data. S. Takada stimulated naive B cells and analyzed class switching in vitro. A. Helfricht, P.E. Thijssen, and J. Wang generated human ZBTB24 constructs. A. Helfricht, R. Jak, and G. Grootaers performed multiphoton laser micro-irradiation experiments. R. Jak performed PARylation assays. A. de Groot purified recombinant GFP-ZBTB24 WT and GFP-ZBTB24 Δ ZNF proteins. R.G. Shah and G.M. Shah performed in vitro PAR binding and synthesis and protection assays. P.E. Thijssen and C. Stoepker examined PAR stability. M. Rogier, J. Moritz, V. Heyer, and B. Reina-San-Martin generated Zbtb24 KO CH12 cells, constructed mouse Zbtb24 retroviral vectors, performed Western blot and CSR assays on CH12 cells, and analyzed the CSR data. C. Stoepker performed FokI assays and generated ZBTB24 KO U2OS cells and HeLa Flp-In/T-REx cells expressing inducible and siRNA-resistant GFP-ZBTB24 and GFP-ZBTB24 Δ ZNF. M.S. Luijsterburg performed IPs for MS and generated stable

GFP-XRCC4 cells. A.C.O. Vertegaal analyzed the MS samples. P. Rao analyzed RNA sequencing data. M.M. van Ostaijen-ten Dam and M.J.D. van Tol analyzed lymphocytes and Ig production. H. IJspeert and M. van der Burg analyzed IgH repertoire and sequenced CSR junctions. L. Du and Q. Pan-Hammarström performed CSR junction analysis. S.M. van der Maarel and H. van Attikum supervised the project. A. Helfricht, P.E. Thijssen, M.B. Rother, and H. van Attikum wrote the paper.

Disclosure: The authors declare no competing interests exist.

Submitted: 9 September 2019

Revised: 6 April 2020

Accepted: 19 May 2020

References

- Ali, A.A.E., G. Timinszky, R. Arribas-Bosacoma, M. Kozłowski, P.O. Hassa, M. Hassler, A.G. Ladurner, L.H. Pearl, and A.W. Oliver. 2012. The zinc-finger domains of PARP1 cooperate to recognize DNA strand breaks. *Nat. Struct. Mol. Biol.* 19:685–692. <https://doi.org/10.1038/nsmb.2335>
- Alt, F.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. <https://doi.org/10.1016/j.cell.2013.01.007>
- Altmeyer, M., K.J. Neelsen, F. Teloni, I. Pozdnyakova, S. Pellegrino, M. Gröfte, M.D. Rask, W. Streicher, S. Jungmichel, M.L. Nielsen, et al. 2015. Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). *Nat. Commun.* 6:8088. <https://doi.org/10.1038/ncomms9088>
- 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. *J. Biol. Chem.* 279:55117–55126. <https://doi.org/10.1074/jbc.M404524200>
- Bardwell, V.J., and R. Treisman. 1994. The POZ domain: a conserved protein-protein interaction motif. *Genes Dev.* 8:1664–1677. <https://doi.org/10.1101/gad.8.14.1664>
- Barreto, V., B. Reina-San-Martin, A.R. Ramiro, K.M. McBride, and M.C. Nussenzweig. 2003. C-terminal deletion of AID uncouples class switch recombination from somatic hypermutation and gene conversion. *Mol. Cell*. 12:501–508. [https://doi.org/10.1016/S1097-2765\(03\)00309-5](https://doi.org/10.1016/S1097-2765(03)00309-5)
- Bennardo, N., A. Cheng, N. Huang, and J.M. Stark. 2008. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet.* 4: e1000110. <https://doi.org/10.1371/journal.pgen.1000110>
- Björkman, A., P. Qvist, L. Du, M. Bartish, A. Zaravinos, K. Georgiou, A.D. Børglum, R.A. Gatti, T. Törngren, and Q. Pan-Hammarström. 2015. Aberrant recombination and repair during immunoglobulin class switching in BRCA1-deficient human B cells. *Proc. Natl. Acad. Sci. USA*. 112:2157–2162. <https://doi.org/10.1073/pnas.1418947112>
- Blanco-Betancourt, C.E., A. Moncla, M. Milili, Y.L. Jiang, E.M. Viegas-Péguignot, B. Roquelaure, I. Thuret, and C. Schiff. 2004. Defective B-cell-negative selection and terminal differentiation in the ICF syndrome. *Blood*. 103:2683–2690. <https://doi.org/10.1182/blood-2003-08-2632>
- van den Boogaard, M.L., P.E. Thijssen, C. Aytikin, F. Licciardi, A.A. Kiykim, L. Sposito, V.A.S.H. Dalm, G.J. Driessen, R. Kersseboom, F. de Vries, et al. 2017. Expanding the mutation spectrum in ICF syndrome: Evidence for a gender bias in ICF2. *Clin. Genet.* 92:380–387. <https://doi.org/10.1111/cge.12979>
- Brochu, G., C. Duchaine, L. Thibeault, J. Lagueux, G.M. Shah, and G.G. Poirier. 1994. Mode of action of poly(ADP-ribose) glycohydrolase. *Biochim. Biophys. Acta*. 1219:342–350. [https://doi.org/10.1016/0167-4781\(94\)90058-2](https://doi.org/10.1016/0167-4781(94)90058-2)
- Cagigi, A., L. Du, L.V. Dang, S. Grutzmeier, A. Atlas, F. Chioldi, Q. Pan-Hammarström, and A. Nilsson. 2009. CD27(-) B-cells produce class switched and somatically hyper-mutated antibodies during chronic HIV-1 infection. *PLoS One*. 4: e5427. <https://doi.org/10.1371/journal.pone.0005427>
- Caron, P., T. Pankotaj, W.W. Wiegant, M.A.X. Tollenaere, A. Furst, C. Bonhomme, A. Helfricht, A. de Groot, A. Pastink, A.C.O. Vertegaal, et al. 2019. WWP2 ubiquitylates RNA polymerase II for DNA-PK-dependent transcription arrest and repair at DNA breaks. *Genes Dev.* 33:684–704. <https://doi.org/10.1101/gad.321943.118>
- Chouery, E., J. Abou-Ghoch, S. Corbani, N. El Ali, R. Korban, N. Salem, C. Castro, S. Klayme, M. Azoury-Abou Rjeily, R. Khoury-Matar, et al. 2012. A novel deletion in ZBTB24 in a Lebanese family with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *Clin. Genet.* 82:489–493. <https://doi.org/10.1111/j.1399-0004.2011.01783.x>
- Cox, J., and M. Mann. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26:1367–1372. <https://doi.org/10.1038/nbt.1511>
- de Greef, J.C., J. Wang, J. Balog, J.T. den Dunnen, R.R. Frants, K.R. Straasheijm, C. Aytikin, M. van der Burg, L. Duprez, A. Ferster, et al. 2011. Mutations in ZBTB24 are associated with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *Am. J. Hum. Genet.* 88: 796–804. <https://doi.org/10.1016/j.ajhg.2011.04.018>
- Doil, C., N. Mailand, S. Bekker-Jensen, P. Menard, D.H. Larsen, R. Pepperkok, J. Ellenberg, S. Panier, D. Durocher, J. Bartek, et al. 2009. RNF168 binds and amplifies ubiquitin conjugates on damaged chromosomes to allow accumulation of repair proteins. *Cell*. 136:435–446. <https://doi.org/10.1016/j.cell.2008.12.041>
- Du, L., M. van der Burg, S.W. Popov, A. Kotnis, J.J. van Dongen, A.R. Gennery, and Q. Pan-Hammarström. 2008. Involvement of Artemis in nonhomologous end-joining during immunoglobulin class switch recombination. *J. Exp. Med.* 205:3031–3040. <https://doi.org/10.1084/jem.20081915>
- Enervald, E., L. Du, T. Visnes, A. Björkman, E. Lindgren, J. Wincent, G. Borck, L. Colleaux, V. Cormier-Daire, D.C. van Gent, et al. 2013. A regulatory role for the cohesin loader NIPBL in nonhomologous end joining during immunoglobulin class switch recombination. *J. Exp. Med.* 210: 2503–2513. <https://doi.org/10.1084/jem.20130168>
- Eustermann, S., H. Videler, J.C. Yang, P.T. Cole, D. Gruszka, D. Veprintsev, and D. Neuhaus. 2011. The DNA-binding domain of human PARP-1 interacts with DNA single-strand breaks as a monomer through its second zinc finger. *J. Mol. Biol.* 407:149–170. <https://doi.org/10.1016/j.jmb.2011.01.034>
- Girard, P.M., B. Kysela, C.J. Härer, A.J. Doherty, and P.A. Jeggo. 2004. Analysis of DNA ligase IV mutations found in LG4 syndrome patients: the impact of two linked polymorphisms. *Hum. Mol. Genet.* 13:2369–2376. <https://doi.org/10.1093/hmg/ddh274>
- Hagleitner, M.M., A. Lankester, P. Maraschio, M. Hultén, J.P. Fryns, C. Schuetz, G. Gimelli, E.G. Davies, A. Gennery, B.H. Belohradsky, et al. 2008. Clinical spectrum of immunodeficiency, centromeric instability and facial dysmorphism (ICF syndrome). *J. Med. Genet.* 45:93–99. <https://doi.org/10.1136/jmg.2007.053397>
- Hansen, R.S., C. Wijmenga, P. Luo, A.M. Stanek, T.K. Canfield, C.M. Weemaes, and S.M. Gartler. 1999. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc. Natl. Acad. Sci. USA*. 96:14412–14417. <https://doi.org/10.1073/pnas.96.25.14412>
- Helfricht, A., W.W. Wiegant, P.E. Thijssen, A.C. Vertegaal, M.S. Luijsterburg, and H. van Attikum. 2013. Remodeling and spacing factor 1 (RSF1) deposits centromere proteins at DNA double-strand breaks to promote non-homologous end-joining. *Cell Cycle*. 12:3070–3082. <https://doi.org/10.4161/cc.26033>
- Hendriks, I.A., R.C. D'Souza, B. Yang, M. Verlaan-de Vries, M. Mann, and A.C. Vertegaal. 2014. Uncovering global SUMOylation signaling networks in a site-specific manner. *Nat. Struct. Mol. Biol.* 21:927–936. <https://doi.org/10.1038/nsmb.2890>
- Ismail, I.H., J-P. Gagné, M-C. Caron, D. McDonald, Z. Xu, J-Y. Masson, G.G. Poirier, and M.J. Hendzel. 2012. CBX4-mediated SUMO modification regulates BMI1 recruitment at sites of DNA damage. *Nucleic Acids Res.* 40:5497–5510. <https://doi.org/10.1093/nar/gks222>
- Izhar, L., B. Adamson, A. Ciccia, J. Lewis, L. Pontano-Vaites, Y. Leng, A.C. Liang, T.F. Westbrook, J.W. Harper, and S.J. Elledge. 2015. A Systematic Analysis of Factors Localized to Damaged Chromatin Reveals PARP-Dependent Recruitment of Transcription Factors. *Cell Rep.* 11:1486–1500. <https://doi.org/10.1016/j.celrep.2015.04.053>
- Jeevan-Raj, B.P., I. Robert, V. Heyer, A. Page, J.H. Wang, F. Cammas, F.W. Alt, R. Lissou, 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. <https://doi.org/10.1084/jem.20110118>
- Kabotyanski, E.B., L. Gomelsky, J.O. Han, T.D. Stamato, and D.B. Roth. 1998. Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Res.* 26:5333–5342. <https://doi.org/10.1093/nar/26.23.5333>
- Kalisch, T., J.C. Amé, F. Dantzer, and V. Schreiber. 2012. New readers and interpretations of poly(ADP-ribosylation). *Trends Biochem. Sci.* 37: 381–390. <https://doi.org/10.1016/j.tibs.2012.06.001>

- Kanariou, M., E. Petridou, M. Liatsis, K. Revinthi, K. Mandalenaki-Lambrou, and D. Trichopoulos. 1995. Age patterns of immunoglobulins G, A & M in healthy children and the influence of breast feeding and vaccination status. *Pediatr Allergy Immunol.* 6:24–29. <https://doi.org/10.1111/j.1399-3038.1995.tb00253.x>
- Kawamitsu, H., H. Hoshino, H. Okada, M. Miwa, H. Momoi, and T. Sugimura. 1984. Monoclonal antibodies to poly(adenosine diphosphate ribose) recognize different structures. *Biochemistry.* 23:3771–3777. <https://doi.org/10.1021/bi00311a032>
- Kleinstiver, B.P., V. Pattanayak, M.S. Prew, S.Q. Tsai, N.T. Nguyen, Z. Zheng, and J.K. Joung. 2016. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature.* 529:490–495. <https://doi.org/10.1038/nature16526>
- Langelier, M.F., J.L. Planck, S. Roy, and J.M. Pascal. 2012. Structural basis for DNA damage-dependent poly(ADP-ribosylation) by human PARP-1. *Science.* 336:728–732. <https://doi.org/10.1126/science.1216338>
- Lee, S.U., and T. Maeda. 2012. POK/ZBTB proteins: an emerging family of proteins that regulate lymphoid development and function. *Immunol. Rev.* 247:107–119. <https://doi.org/10.1111/j.1600-065X.2012.01116.x>
- Li, M., L.Y. Lu, C.Y. Yang, S. Wang, and X. Yu. 2013. The FHA and BRCT domains recognize ADP-ribosylation during DNA damage response. *Genes Dev.* 27:1752–1768. <https://doi.org/10.1101/gad.226357.113>
- Lin, M., L. Du, P. Brandtzaeg, and Q. Pan-Hammarström. 2014. IgA subclass switch recombination in human mucosal and systemic immune compartments. *Mucosal Immunol.* 7:511–520. <https://doi.org/10.1038/mi.2013.68>
- Lonskaya, I., V.N. Potaman, L.S. Shlyakhtenko, E.A. Oussatcheva, Y.L. Lyubchenko, and V.A. Soldatenkov. 2005. Regulation of poly(ADP-ribose) polymerase-1 by DNA structure-specific binding. *J. Biol. Chem.* 280:17076–17083. <https://doi.org/10.1074/jbc.M413483200>
- 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. *Proc. Natl. Acad. Sci. USA.* 113:1256–1260. <https://doi.org/10.1073/pnas.1521597113>
- Luijsterburg, M.S., I. de Krijger, W.W. Wiegant, R.G. Shah, G. Smeenk, A.J.L. de Groot, A. Pines, A.C.O. Vertegaal, J.J.L. Jacobs, G.M. Shah, et al. 2016. PARP1 Links CHD2-Mediated Chromatin Expansion and H3.3 Deposition to DNA Repair by Non-homologous End-Joining. *Mol. Cell.* 61:547–562. <https://doi.org/10.1016/j.molcel.2016.01.019>
- Manis, J.P., J.C. Morales, Z. Xia, J.L. Kutok, F.W. Alt, and P.B. Carpenter. 2004. 53BP1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination. *Nat. Immunol.* 5:481–487. <https://doi.org/10.1038/ni1067>
- Mitchell, J., G.C. Smith, and N.J. Curtin. 2009. Poly(ADP-Ribose) polymerase-1 and DNA-dependent protein kinase have equivalent roles in double strand break repair following ionizing radiation. *Int. J. Radiat. Oncol. Biol. Phys.* 75:1520–1527. <https://doi.org/10.1016/j.ijrobp.2009.07.1722>
- Moorhouse, M.J., D. van Zessen, H. IJspeert, S. Hiltemann, S. Horsman, P.J. van der Spek, M. van der Burg, and A.P. Stubbs. 2014. ImmunoGlobulin galaxy (IGGalaxy) for simple determination and quantitation of immunoglobulin heavy chain rearrangements from NGS. *BMC Immunol.* 15:59. <https://doi.org/10.1186/s12865-014-0059-7>
- Mortusewicz, O., J.C. Amé, V. Schreiber, and H. Leonhardt. 2007. Feedback-regulated poly(ADP-ribosylation) by PARP-1 is required for rapid response to DNA damage in living cells. *Nucleic Acids Res.* 35:7665–7675. <https://doi.org/10.1093/nar/gkm933>
- Murray, J.E., M. van der Burg, H. IJspeert, P. Carroll, Q. Wu, T. Ochi, A. Leitch, E.S. Miller, B. Kysela, A. Jawad, et al. 2015. Mutations in the NHEJ component XRCC4 cause primordial dwarfism. *Am. J. Hum. Genet.* 96:412–424. <https://doi.org/10.1016/j.ajhg.2015.01.013>
- Najafabadi, H.S., S. Mnaimneh, F.W. Schmitges, M. Garton, K.N. Lam, A. Yang, M. Albu, M.T. Weirauch, E. Radovani, P.M. Kim, et al. 2015. C2H2 zinc finger proteins greatly expand the human regulatory lexicon. *Nat. Biotechnol.* 33:555–562. <https://doi.org/10.1038/nbt.3128>
- Nitta, H., M. Unoki, K. Ichiyangi, T. Kosho, T. Shigemura, H. Takahashi, G. Velasco, C. Francastel, C. Picard, T. Kubota, et al. 2013. Three novel ZBTB24 mutations identified in Japanese and Cape Verdean type 2 ICF syndrome patients. *J. Hum. Genet.* 58:455–460. <https://doi.org/10.1038/jhg.2013.56>
- 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. Immunol.* 32:1300–1308. [https://doi.org/10.1002/1521-4141\(200205\)32:5<1300::AID-IMMU1300>3.0.CO;2-L](https://doi.org/10.1002/1521-4141(200205)32:5<1300::AID-IMMU1300>3.0.CO;2-L)
- 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. *J. Exp. Med.* 201:189–194. <https://doi.org/10.1084/jem.20040772>
- 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. <https://doi.org/10.1371/journal.pone.0059505>
- Petersen, S., R. Casellas, B. Reina-San-Martin, H.T. Chen, M.J. Difilippantonio, P.C. Wilson, L. Hanitsch, A. Celeste, M. Muramatsu, D.R. Pilch, et al. 2001. AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. *Nature.* 414:660–665. <https://doi.org/10.1038/414660a>
- Pierce, A.J., R.D. Johnson, L.H. Thompson, and M. Jasin. 1999. XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev.* 13:2633–2638. <https://doi.org/10.1101/gad.13.20.2633>
- Pines, A., L.H. Mullenders, H. van Attikum, and M.S. Luijsterburg. 2013. Touching base with PARPs: moonlighting in the repair of UV lesions and double-strand breaks. *Trends Biochem. Sci.* 38:321–330. <https://doi.org/10.1016/j.tibs.2013.03.002>
- Ray Chaudhuri, A., and A. Nussenzweig. 2017. The multifaceted roles of PARP1 in DNA repair and chromatin remodeling. *Nat. Rev. Mol. Cell Biol.* 18:610–621. <https://doi.org/10.1038/nrm.2017.53>
- 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. *J. Exp. Med.* 197:1767–1778. <https://doi.org/10.1084/jem.20030569>
- Ren, R., S. Hardikar, J.R. Horton, Y. Lu, Y. Zeng, A.K. Singh, K. Lin, L.D. Coletta, J. Shen, C.S. Lin Kong, et al. 2019. Structural basis of specific DNA binding by the transcription factor ZBTB24. *Nucleic Acids Res.* 47:8388–8398. <https://doi.org/10.1093/nar/gkz557>
- 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. *J. Exp. Med.* 206:1047–1056. <https://doi.org/10.1084/jem.20082468>
- Robu, M., R.G. Shah, N. Petitclerc, J. Brind'Amour, F. Kandan-Kulangara, and G.M. Shah. 2013. Role of poly(ADP-ribose) polymerase-1 in the removal of UV-induced DNA lesions by nucleotide excision repair. *Proc. Natl. Acad. Sci. USA.* 110:1658–1663. <https://doi.org/10.1073/pnas.1209507110>
- Schimmel, J., H. Kool, R. van Schendel, and M. Tijsterman. 2017. Mutational signatures of non-homologous and polymerase theta-mediated end-joining in embryonic stem cells. *EMBO J.* 36:3634–3649. <https://doi.org/10.15252/emj.201796948>
- Shah, G.M., F. Kandan-Kulangara, A. Montoni, R.G. Shah, J. Brind'Amour, M.D. Vodenicharov, and B. Affar. 2011. Approaches to detect PARP-1 activation in vivo, in situ, and in vitro. *Methods Mol. Biol.* 780:3–34. https://doi.org/10.1007/978-1-61779-270-0_1
- 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. <https://doi.org/10.1016/j.cell.2010.04.038>
- Shieh, W.M., J.C. Amé, M.V. Wilson, Z.Q. Wang, D.W. Koh, M.K. Jacobson, and E.L. Jacobson. 1998. Poly(ADP-ribose) polymerase null mouse cells synthesize ADP-ribose polymers. *J. Biol. Chem.* 273:30069–30072. <https://doi.org/10.1074/jbc.273.46.30069>
- Stavnezer, J., A. Björkman, L. Du, A. Cagigi, and Q. Pan-Hammarström. 2010. Mapping of switch recombination junctions, a tool for studying DNA repair pathways during immunoglobulin class switching. *Adv. Immunol.* 108:45–109. <https://doi.org/10.1016/B978-0-12-380995-7.00003-3>
- Tang, J., N.W. Cho, G. Cui, E.M. Manion, N.M. Shanbhag, M.V. Botuyan, G. Mer, and R.A. Greenberg. 2013. Acetylation limits 53BP1 association with damaged chromatin to promote homologous recombination. *Nat. Struct. Mol. Biol.* 20:317–325. <https://doi.org/10.1038/nsmb.2499>
- Taty-Taty, G.C., C. Chailleux, M. Quaranta, A. So, J. Guirouilh-Barbat, B.S. Lopez, P. Bertrand, D. Trouche, and Y. Canitrot. 2016. Control of alternative end joining by the chromatin remodeler p400 ATPase. *Nucleic Acids Res.* 44:1657–1668. <https://doi.org/10.1093/nar/gkv1202>
- Teloni, F., and M. Altmeyer. 2016. Readers of poly(ADP-ribose): designed to be fit for purpose. *Nucleic Acids Res.* 44:993–1006. <https://doi.org/10.1093/nar/gkv1383>
- Thijssen, P.E., Y. Ito, G. Grillo, J. Wang, G. Velasco, H. Nitta, M. Unoki, M. Yoshihara, M. Suyama, Y. Sun, et al. 2015. Mutations in CDCA7 and HELLS cause immunodeficiency-centromeric instability-facial anomalies syndrome. *Nat. Commun.* 6:7870. <https://doi.org/10.1038/ncomms8870>

- Tiller, T., E. Meffre, S. Yurasov, M. Tsuiji, M.C. Nussenzweig, and H. Wardemann. 2008. Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J Immunol Methods*. 329:112–124. <https://doi.org/10.1016/j.jim.2007.09.017>
- Unoki, M., H. Funabiki, G. Velasco, C. Francastel, and H. Sasaki. 2019. CDCA7 and HELLS mutations undermine nonhomologous end joining in centromeric instability syndrome. *J. Clin. Invest.* 129:78–92. <https://doi.org/10.1172/JCI99751>
- van Dongen, J.J., A.W. Langerak, M. Brüggemann, P.A. Evans, M. Hummel, F.L. Lavender, E. Delabesse, F. Davi, E. Schuurin, R. García-Sanz, et al. 2003. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia*. 17:2257–2317. <https://doi.org/10.1038/sj.leu.2403202>
- Veuger, S.J., N.J. Curtin, C.J. Richardson, G.C. Smith, and B.W. Durkacz. 2003. Radiosensitization and DNA repair inhibition by the combined use of novel inhibitors of DNA-dependent protein kinase and poly(ADP-ribose) polymerase-1. *Cancer Res.* 63:6008–6015.
- Vukic, M., and L. Daxinger. 2019. DNA methylation in disease: Immunodeficiency, Centromeric instability, Facial anomalies syndrome. *Essays Biochem.* 63:773–783. <https://doi.org/10.1042/EBC20190035>
- 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. *Nucleic Acids Res.* 34:6170–6182. <https://doi.org/10.1093/nar/gkl840>
- Weemaes, C.M., M.J. van Tol, J. Wang, M.M. van Ostaijen-ten Dam, M.C. van Eggermond, P.E. Thijssen, C. Aytekin, N. Brunetti-Pierri, M. van der Burg, E. Graham Davies, et al. 2013. Heterogeneous clinical presentation in ICF syndrome: correlation with underlying gene defects. *Eur. J. Hum. Genet.* 21:1219–1225. <https://doi.org/10.1038/ejhg.2013.40>
- Woodbine, L., A.R. Gennery, and P.A. Jeggo. 2014. The clinical impact of deficiency in DNA non-homologous end-joining. *DNA Repair (Amst.)*. 16: 84–96. <https://doi.org/10.1016/j.dnarep.2014.02.011>
- Wu, H., P.E. Thijssen, E. de Klerk, K.K. Vonk, J. Wang, B. den Hamer, C. Aytekin, S.M. van der Maarel, and L. Daxinger. 2016. Converging disease genes in ICF syndrome: ZBTB24 controls expression of CDCA7 in mammals. *Hum. Mol. Genet.* 25:4041–4051. <https://doi.org/10.1093/hmg/ddw243>
- Xu, G.L., T.H. Bestor, D. Bourc'his, C.L. Hsieh, N. Tommerup, M. Bugge, M. Hulten, X. Qu, J.J. Russo, and E. Viegas-Péquignot. 1999. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature*. 402:187–191. <https://doi.org/10.1038/46052>
- 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. *Nat. Commun.* 8: 14244. <https://doi.org/10.1038/ncomms14244>

Supplemental material

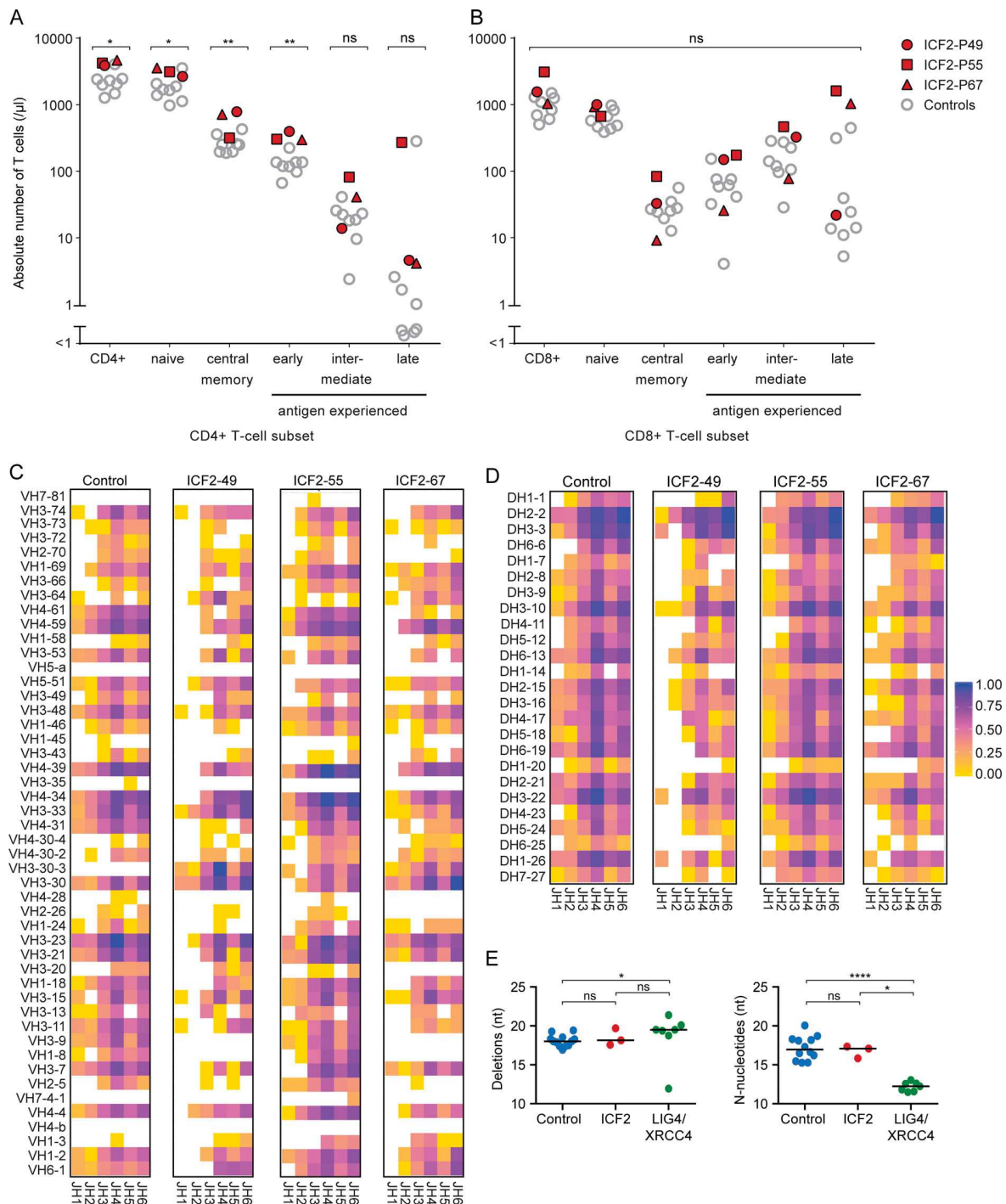


Figure S1. T cell differentiation in ICF2 patients and combinatorial diversity and junction characteristics of IgH rearrangements. Related to Fig. 1. **(A and B)** Absolute numbers (per microliter) of the peripheral blood CD3⁺CD4⁺ T cell subset (A) and CD3⁺CD8⁺ T cell subset (B). The indicated differentiation stages in both subsets were determined in three ICF2 patients and eight healthy age-matched controls (age range 0.8–4.3 yr) by flow cytometry. Phenotypical definitions: naive T cells, CD45RA⁺CCR7⁺; central memory T cells, CD45RA⁺CCR7⁺; antigen experienced CD4⁺ T cells, CD45RA⁻CCR7⁻: early CD28⁺CD27⁺; intermediate CD28⁺CD27⁺; late CD28⁻CD27⁻; antigen experienced CD8⁺ T cells, CD45RA⁻CCR7⁻: early CD28⁺CD27⁺; intermediate CD28⁻CD27⁻; late CD28⁻CD27⁻. Statistical significance was calculated using Student's *t* test (*, *P* < 0.05; **, *P* < 0.01; ns, not significant). **(C and D)** Heatmaps showing the relative frequency of the combinatorial diversity of VH and JH genes (C) or DH and JH genes (D) of unique productive IgH rearrangements (defined by the unique combination of VH, DH, JH, and nucleotide sequences of CDR3) amplified from control (*n* = 4,789) and ICF2 patients ICF2-49 (*n* = 757), ICF2-55 (*n* = 3,723), and ICF2-67 (*n* = 1,663). **(E)** The ICF2 patients (*n* = 3) display normal numbers of deletions and N-nucleotides compared with control (*n* = 12), in contrast to XRCC4- and LIG4-deficient patients (*n* = 7; Murray et al., 2015), who display increased numbers of deletions and decreased numbers of N-nucleotides in unique unproductive IGH rearrangements. Statistical significance was calculated using the Mann-Whitney *U* test (*, *P* < 0.05; ****, *P* < 0.0001; ns, not significant).

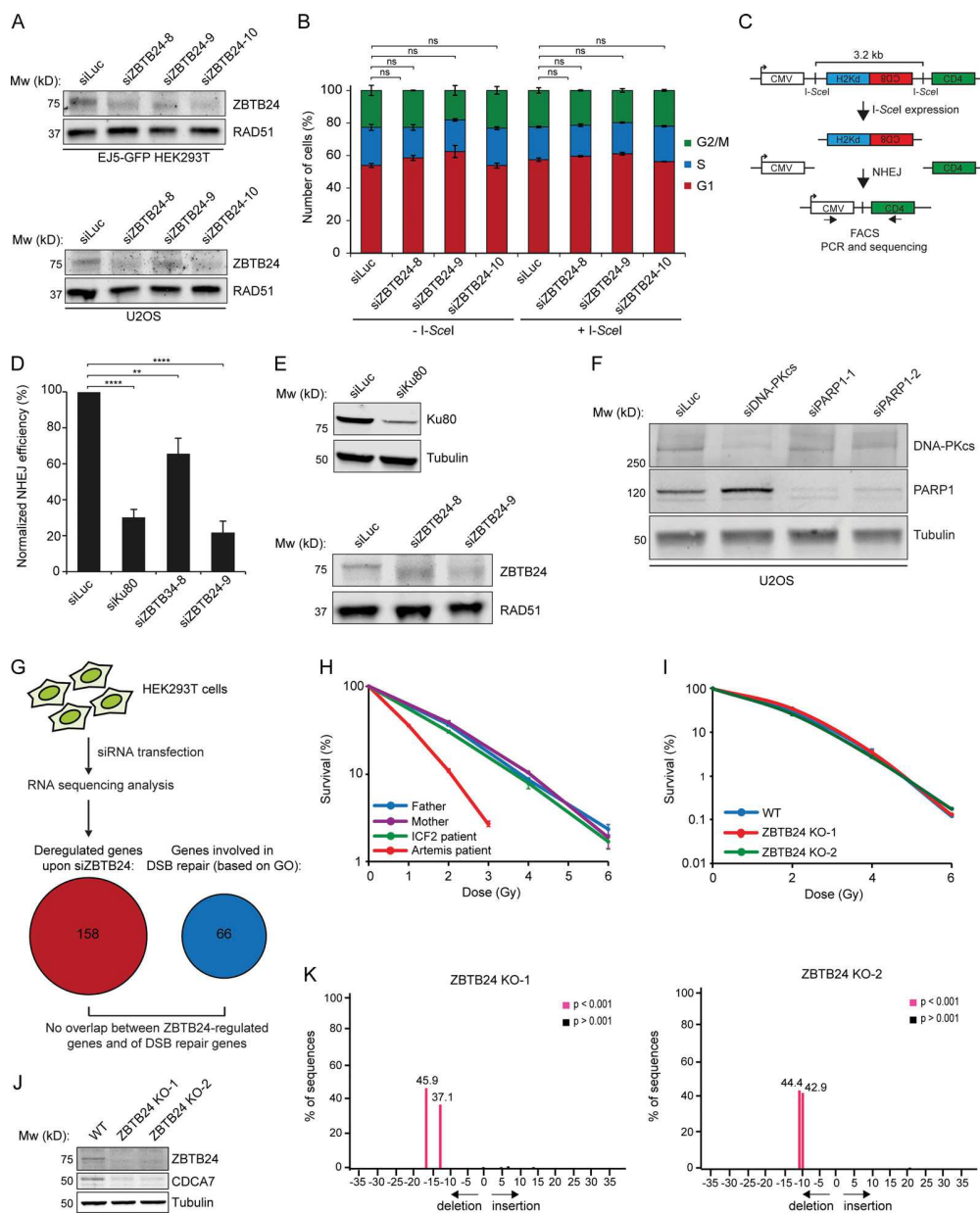


Figure S2. Knockdown of ZBTB24 does not affect cell cycle progression and the expression of genes involved in DSB repair. ICF2 patient-derived fibroblasts and ZBTB24 KO U2OS cells are not sensitive to IR (related to Fig. 2). **(A)** U2OS cells and HEK293T cells containing the EJ5-GFP reporter were treated with the indicated siRNAs. WCEs were prepared 48 h later and subjected to Western blot analysis for ZBTB24. RAD51 is a loading control. **(B)** HEK293T cells containing the EJ5-GFP reporter were transfected with the indicated siRNAs. 48 h later, cells were transfected with a control vector or the I-SceI expression vector (pCBASce). After an additional 24 h, cells were subjected to propidium iodide staining followed by flow cytometry analysis. The percentage of cells in G1 (red bar), S (blue bar), and G2/M (green bar) phase is presented. The mean \pm SEM from two independent experiments is shown. Statistical significance was calculated using Student's *t* test (ns, not significant). **(C)** Schematic of the GC92 reporter for NHEJ. **(D)** Fibroblasts containing the GC92 reporter were treated with the indicated siRNAs and, 48 h later, cotransfected with I-SceI (pCBASce) and mCherry expression vectors. The ratio of CD4-FITC/mCherry-expressing cells was counted by flow cytometry 48 h later. The mean \pm SEM from three to four independent experiments is shown. Statistical significance was calculated using Student's *t* test (**, $P < 0.01$; ****, $P < 0.0001$). **(E)** Cells from D were subjected to Western blot analysis of KU80 and ZBTB24 expression. Tubulin and RAD51 are loading controls. **(F)** U2OS cells were treated with the indicated siRNAs. WCEs were prepared 48 h later and subjected to Western blot analysis for DNA-PKcs and PARP1. Tubulin is a loading control. **(G)** HEK293T cells were treated with control siRNAs against luciferase or three different siRNAs against ZBTB24. 4 d later, RNA was isolated and subjected to RNA sequencing analysis. The number of genes found to be commonly misregulated after ZBTB24-depletion with each of the siRNAs is presented (false discovery rate < 0.05). Importantly, gene ontology term analysis (0006302; DSB repair) did not reveal the presence of DSB repair genes among the misregulated genes. **(H)** ICF2 patient-derived fibroblasts were exposed to different doses of IR and scored for clonogenic survival. The mean \pm SEM from two independent experiments is shown. **(I)** ZBTB24 KO U2OS cells were exposed to different doses of IR and scored for clonogenic survival. The mean \pm SEM from two independent experiments is shown. **(J)** Western blot analysis of ZBTB24 and CDCA7 expression in ZBTB24 KO U2OS clones from I. Tubulin is a loading control. **(K)** TIDE analysis of ZBTB24 KO clones from I, showing 17- and 13-bp out-of-frame deletions in ZBTB24 KO-1 and 11 and 10-bp out-of-frame deletions in ZBTB24 KO-2.

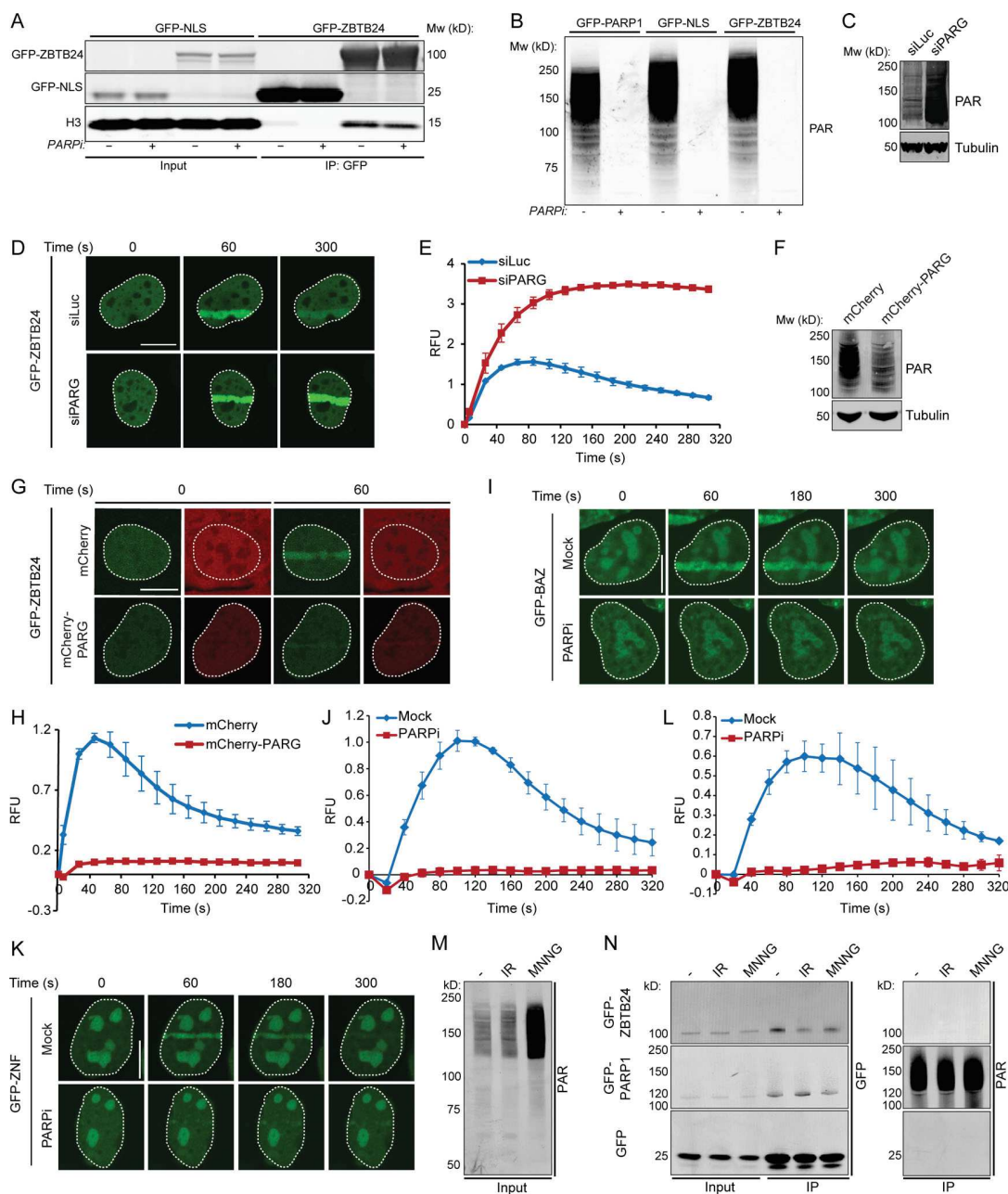


Figure S3. PARG-dependent turnover of PAR chains modulates the accumulation of ZBTB24 at sites of DNA damage. The ZNF domain of ZBTB24 accumulates at sites of DNA damage in a PARP-dependent manner. ZBTB24 is not PARYlated after DNA damage induction (related to Figs. 3 and 4). **(A)** U2OS cells transiently expressing GFP-ZBTB24 or GFP-NLS were treated with either DMSO (Mock) or PARPi. WCEs were subjected to GFP IP followed by Western blot analysis of the indicated proteins. **(B)** U2OS cells transfected with the indicated GFP-tagged proteins were treated with either DMSO (Mock) or PARPi. WCEs were prepared and subjected to Western blot analysis to assess total PAR levels. **(C)** Western blot analysis showing total PAR levels in U2OS cells transfected with the indicated siRNAs and transiently expressing GFP-ZBTB24. Tubulin is loading control. **(D)** GFP-ZBTB24 accumulation as monitored at the indicated time points after laser microirradiation in cells from C. Scale bar, 10 μ m. **(E)** Quantification of the results from D. The mean \pm SEM from two independent experiments is shown. **(F)** As in C, except that cells were cotransfected with a GFP-ZBTB24 and either an mCherry or mCherry-PARG expression vector were used. **(G)** As in D, except that cells from F were used. Scale bar, 10 μ m. **(H)** Quantification of the results from G. The mean \pm SEM from three independent experiments is shown. **(I)** U2OS cells transiently expressing GFP-tagged BAZ domains of ZBTB24 were treated with DMSO (Mock) or PARPi and subjected to laser microirradiation to follow GFP-BAZ accumulation at sites of DNA damage at the indicated time points after irradiation. Scale bar, 10 μ m. **(J)** Quantification of I. The mean \pm SEM from two independent experiments is shown. **(K)** As in I, except for the GFP-tagged ZNF domain of ZBTB24 (GFP-ZNF). Scale bar, 10 μ m. **(L)** Quantification of K. The mean \pm SEM from two independent experiments is shown. **(M)** U2OS cells expressing GFP were left untreated or treated with IR or MNNG. WCE were prepared and subjected to Western blot analysis for global PAR levels. **(N)** WCE extracts from M and from cells expressing GFP-ZBTB24 or GFP-PARP1 were subjected to GFP IP. Washes were performed under high-salt conditions to remove interacting proteins. Western blot analysis was done for the indicated proteins and PAR. The experiment was performed two times for PARP1 and four times for ZBTB24. Blots from a representative experiment are shown.

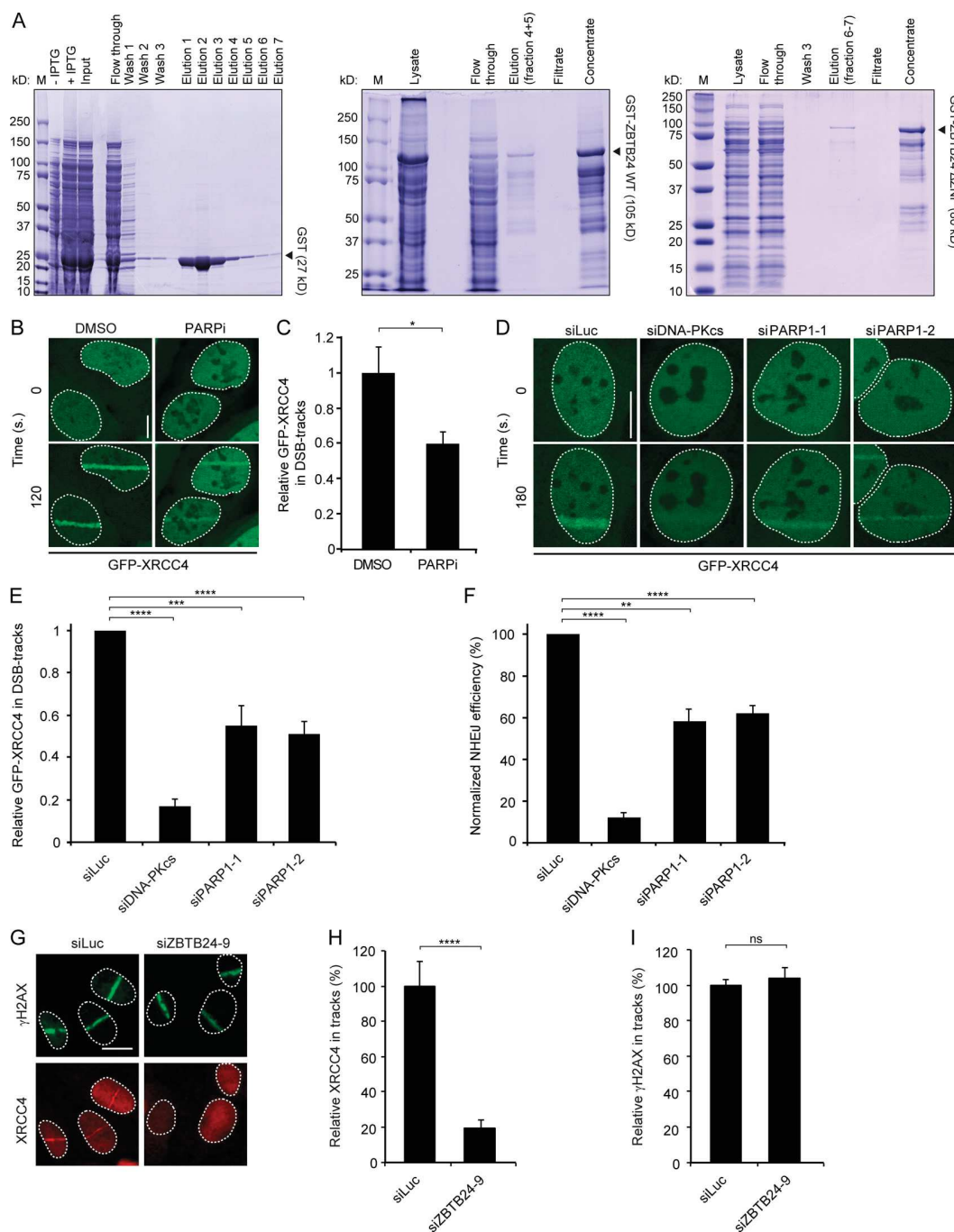


Figure S4. **Purification of recombinant ZBTB24 and PARP1 promotes XRCC4/LIG4 assembly and NHEJ at DNA damage sites.** Related to Figs. 5 and 6. **(A)** Coomassie-stained gel of recombinant GST, GST-tagged ZBTB24, and GST-tagged ZBTB24 Δ ZNF, which were purified after expression in *E. coli*. The indicated samples from the purification procedure were loaded and run on a 4–12% polyacrylamide gel. **(B)** U2OS stably expressing GFP-XRCC4 were treated with DMSO (Mock) or PARPi and subjected to laser microirradiation. Representative images of unirradiated and irradiated cells (taken at the indicated time point after irradiation) are shown. Scale bar, 10 μ m. **(C)** Quantification of B. The mean \pm SEM from two independent experiments is shown. Statistical significance was calculated using Student's *t* test (*, $P < 0.05$). **(D)** As in B, except that cells were transfected with the indicated siRNAs. Scale bar, 10 μ m. **(E)** Quantification of D. The mean \pm SEM from three to five independent experiments is shown. Statistical significance was calculated using Student's *t* test (***, $P < 0.001$; ****, $P < 0.0001$). **(F)** Plasmid integration assays in U2OS cells transfected with indicated siRNAs. The mean \pm SEM from two to four independent experiments is shown. Statistical significance was calculated using Student's *t* test (**, $P < 0.01$; ****, $P < 0.0001$). **(G)** Accumulation of γ H2AX and endogenous XRCC4 at sites of laser-inflicted DNA damage. U2OS cells were treated with the indicated siRNAs, subjected to laser microirradiation, and 10 min later, fixed and immunostained for γ H2AX and endogenous XRCC4. Scale bar, 10 μ m. **(H)** Quantification of endogenous XRCC4 levels in laser tracks from G. The mean \pm SEM from two independent experiments is shown. Statistical significance was calculated using Student's *t* test (****, $P < 0.0001$). **(I)** As in H, except for γ H2AX. The mean \pm SEM from two independent experiments is shown. Statistical significance was calculated using Student's *t* test (ns, not significant).

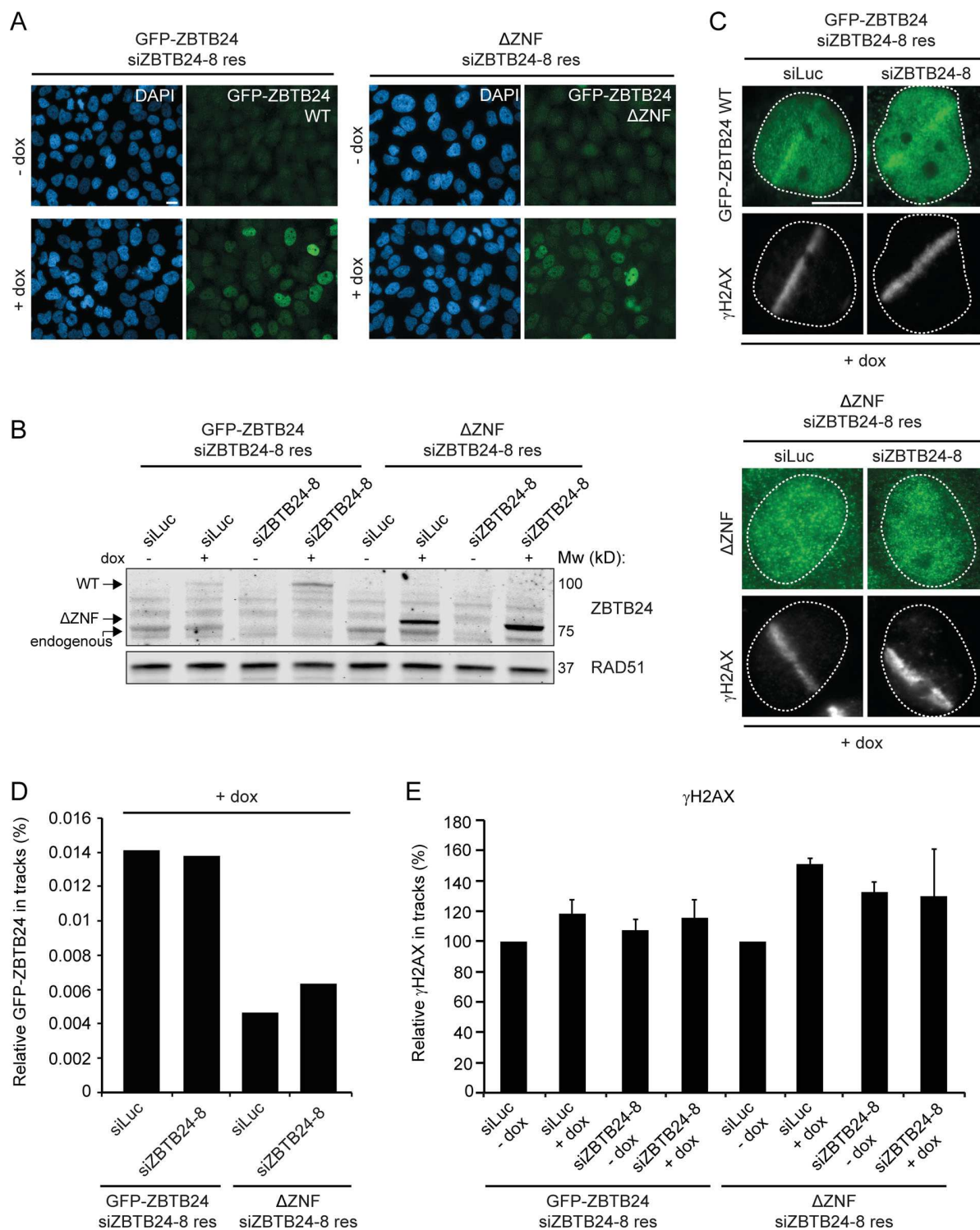


Figure S5. **Analysis of HeLa Flp-In/T-REx cells expressing GFP-ZBTB24 or GFP-ZBTB24 ΔZNF.** Related to Fig. 6. **(A)** HeLa Flp-In/T-REx cells carrying stably integrated inducible and siZBTB24-resistant GFP-ZBTB24 or GFP-ZBTB24 ΔZNF expression vectors express the GFP-tagged proteins upon doxycycline (dox) treatment. Representative microscope images showing dox-inducible expression. Scale bar, 10 μm. **(B)** Western blot analysis of ZBTB24 expression in cells from A. RAD51 is a loading control. **(C)** GFP-ZBTB24 and GFP-ZBTB24 ΔZNF recruitment at sites of laser-inflicted DNA damage in HeLa Flp-In/T-REx cells induced with dox. Cells were transfected with the indicated siRNAs, subjected to laser microirradiation, and 10 min later, fixed and immunostained. γH2AX is a DNA damage marker. Scale bar, 10 μm. **(D)** Quantification of the recruitment in cells from C. The mean from one experiment is shown. **(E)** Quantification of γH2AX levels in laser tracks from cells in Fig. 6 G. The mean ± SEM of two to three independent experiments is shown.

Tables S1–S5 are provided online as separate files. Table S1 shows Ig isotype concentrations at first analysis. Table S2 lists ZBTB24-regulated genes identified by RNA sequencing. Table S3 lists proteins identified as ZBTB24 interactors by SILAC MS, ordered by H/L. Table S4 lists proteins identified as PARP1 interactors by SILAC MS, ordered by H/L. Table S5 lists primers and siRNAs. Data S1, S2, and S3 are available online as PDFs. Data S1 contains the S μ -S α junctions from ICF2 patients. Data S2 contains the S μ -S α junctions from healthy children. Data S3 contains the S μ -S γ junctions from ICF2 patients.

II- Discussion

L'ensemble de ces résultats montre que Zbtb24 est nécessaire pour une commutation efficace et qu'il promeut la voie de réparation du C-NHEJ avec la coopération de PARP1 en régulant l'assemblage du complexe XRCC4/LIG4.

La réduction sévère de la production d'Ig et de la capacité de diversification observée dans les cellules B des patients atteints d'ICF2 est cohérente avec les précédentes affirmations et confirme qu'une altération du mécanisme de CI peut en être la cause. L'usage de longues micro-homologies dans les jonctions des cellules B de patients atteints d'ICF2 montre une réorientation vers la voie A-EJ dans celles-ci et donc que ZBTB24 pourrait être un nouveau facteur du NHEJ impliqué durant la CI. En effet, nous avons ensuite révélé que ZBTB24 est recruté aux sites de dommages de l'ADN d'une manière dépendante de PARP1 en s'associant avec les chaînes PAR générées par PARP1 via son domaine ZNF. Zbtb24 stimule la synthèse des chaînes PAR médiée par PARP1 et les protège de la dégradation dépendante de PARG (poly(ADP-ribose) glycohydrolase), renforçant ainsi le recrutement du complexe LIG4/XRCC4 dépendant de PARP1 pour favoriser une réparation efficace des DSBs par le C-NHEJ. Nous avons ensuite montré que la perte de Zbtb24 dans la lignée cellulaire CH12 entraîne un défaut de CI. L'ensemble de ces résultats montre que ZBTB24 est impliqué dans la voie du C-NHEJ durant la CI, fournissant ainsi le mécanisme moléculaire menant à l'immunodéficiência observée chez les patients atteints d'ICF2. De plus, il a été très récemment montré que les facteurs CDCA7 et HELLS, impliqués dans les ICF3 et 4 respectivement, participent aussi à la voie de réparation C-NHEJ (He et al., 2020; Unoki et al., 2019).

Cette étude permet de rappeler que Parp1 qui est à l'origine du recrutement de Mre11 et Nbs1 aux sites de DSB durant l'A-EJ, peut également stimuler une réparation efficace des DSBs par le biais du C-NHEJ. Il est important de noter que ZBTB24 pourrait également promouvoir l'A-EJ étant donné son interaction avec PARP1.

En conclusion, nos travaux ont permis de mettre en lumière le mécanisme provoquant l'immunodéficiência dans le syndrome ICF en présence de mutations

dans le gène ZBTB24 (ICF2). Ainsi, nous avons révélé que Zbtb24 est impliqué dans la commutation isotypique en favorisant la réparation des cassures double brin par la voie C-NHEJ dépendante de Parp1.

CHAPITRE 4 :

**Caractérisation fonctionnelle de
l'enhancer γ 1E lors de la CI.**

I- Une nouvelle région régulatrice contrôle la transcription du locus IgH et la CI à un sous-ensemble d'isotypes

La CI est un évènement de recombinaison de longue distance qui se produit au sein du locus IgH et implique la cohésion de deux régions S qui peuvent être séparées par plusieurs centaines de kb. Par conséquent, pour que la CI soit réalisée, les régions S donneuse et acceptrice doivent être rapprochées, et cela se produit par des changements de conformation tri-dimensionnelle qui impliquent la formation de boucles d'ADN couplées à la transcription (Kenter et al., 2012). Dans les cellules B quiescentes, les enhancers E_{μ} et 3'RR interagissent grâce à la formation d'une boucle (Kenter et al., 2012; Wuerffel et al., 2007). Puis, lors de l'activation des cellules B, la conformation de la boucle change et mène au recrutement de la région S acceptrice dans la boucle principale, se terminant par une interaction tripartite entre les enhancers E_{μ} , 3'RR et la région S acceptrice (Kenter et al., 2012; Wuerffel et al., 2007). Ces changements de conformation sont notamment dépendants du complexe Mediator (Thomas-Claudepierre et al., 2016). Cependant le mécanisme précis et les facteurs additionnels requis pour ces changements de conformation sont encore inconnus.

Par le biais d'expériences 4C, notre équipe a trouvé qu'une région située en aval du gène $C\gamma 1$, appelée $\gamma 1E$ interagit dynamiquement avec les enhancers E_{μ} et 3'RR durant la CI. De plus, cette région se lie aux sous unités Med1 et Med12 du complexe Mediator et est transcrite dans les cellules B activées (Thomas-Claudepierre et al., 2016). Celle-ci a également été décrite dans de précédents travaux comme un enhancer putatif dans des cellules pro-B (Medvedovic et al., 2013; Predeus et al., 2014). En effet, ces travaux ont suggéré que cette région contiendrait un enhancer putatif, lors du développement des cellules B, car elle présente des caractéristiques d'enhancer telles que la liaison de facteurs de transcription (IRF4, IRF8, YY1, E2A, PAX5), la présence de sites hypersensibles à la DNase I et des marques épigénétiques de la chromatine active (Medvedovic et al., 2013; Predeus et al., 2014).

Basés sur ces observations, nous avons émis l'hypothèse que la région située en aval du gène $C\gamma 1$ comporte un enhancer putatif $\gamma 1E$ et que celui-ci pourrait avoir un rôle important dans l'activation transcriptionnelle du locus IgH et dans l'établissement d'interactions longue distance durant la recombinaison V(D)J, la CI et l'hypermutation somatique.

J'ai contribué à la caractérisation de l'enhancer putatif $\gamma 1E$ dans le modèle cellulaire CH12 durant la CI ainsi que dans le modèle souris $\gamma 1E^{-/-}$ durant le développement des cellules B.

Nous avons tout d'abord montré que l'absence de de la région $\gamma 1E$ dans les cellules CH12 réduit de 50% l'efficacité de CI vers IgA et de manière surprenante n'affecte pas les niveaux de transcrits $S\alpha$. Afin d'étudier le rôle de $\gamma 1E$ vers les autres isotypes, nous avons généré un modèle souris $\gamma 1E^{-/-}$ qui nous a permis de révéler que son absence n'affecte pas le développement des cellules B mais réduit l'efficacité de CI vers IgG3, IgG2b et IgG2a qui corrèle avec un défaut du niveau des GLTs correspondants. Ceci montre que l'enhancer $\gamma 1E$ régule la transcription du locus IgH et la CI d'une manière spécifique à l'isotype.



CORRESPONDENCE

A novel regulatory region controls IgH locus transcription and switch recombination to a subset of isotypes

Rocío Amoretti-Villa^{1,2,3,4}, Mélanie Rogier^{1,2,3,4}, Isabelle Robert^{1,2,3,4}, Vincent Heyer^{1,2,3,4} and Bernardo Reina-San-Martin^{1,2,3,4}

Cellular & Molecular Immunology _____; <https://doi.org/10.1038/s41423-019-0267-4>

Class switch recombination (CSR) occurs at the IgH locus and replaces the immunoglobulin (Ig) isotype expressed from IgM to IgG, IgE or IgA, endowing the B cell receptor with novel effector functions. CSR is triggered by activation-induced cytidine deaminase (AID),¹ an enzyme that deaminates cytosines to uracils in single-stranded DNA exposed by transcription. The distinct antibody isotypes are encoded in the IgH locus in individual transcription units composed of a cytokine-inducible promoter, an intronic exon, and a switch region (Sx), followed by the exons encoding the constant region (Cx) (Fig. S1a). During CSR, the choice of recombination to a particular isotype is determined by the stimulation-dependent activation of specific promoters, triggering the generation of noncoding germline transcripts (GLTs).² Thus far, the transcriptional regulation of the IgH locus is known to be controlled by the E μ enhancer, located downstream of the variable region and upstream of the donor switch region (S μ), and the 3' regulatory region (3'RR) super-enhancer located downstream of C α .³

Because the donor and acceptor switch (S) regions may be separated by up to 200 kb, CSR requires long-range interactions to occur. In mature resting B cells, both E μ and the 3'RR super-enhancer are located in close proximity.^{4,5} Upon B cell stimulation, the locus undergoes 3D conformational changes that bring the two S regions that will recombine near the E μ and the 3'RR.⁵ These conformational changes are in part dependent on the Mediator complex,⁵ but the precise mechanism, additional regulatory regions, and factors involved are poorly understood. Through 4C-Seq experiments, we found that a region located downstream of the IgG1 gene (termed here γ 1E) dynamically interacts with the E μ and 3'RR enhancers during CSR.⁵ This region is not only bound by the Med1 and Med12 subunits of Mediator⁵ but it also bears chromatin marks and features characteristic of enhancers in pro-B cells.^{6,7} Moreover, this region is transcribed in activated B cells (Fig. 1a), a feature shared by active enhancers.⁸ Based on these observations, we hypothesized that the γ 1E region could have a transcriptional regulatory function during CSR. To test this hypothesis, we used a CRISPR/Cas9 knockout strategy in CH12 cells, a murine B cell line that can be efficiently stimulated to perform CSR to IgA.⁹ Cells were transfected with a plasmid coexpressing two gRNAs flanking the γ 1E region and Cas9 fused to EGFP. One day after transfection, cells were sorted for EGFP expression and cultured under limiting dilutions to generate individual clones that were genotyped by PCR and sequencing (Fig. S1a). Notably, in CH12 cells, only one IgH allele is functional.¹⁰ The other allele has a D-J rearrangement,¹¹ is pre-recombined

(between S μ and S α) and thus lacks the γ 1E region. Therefore, wild-type CH12 cells are denoted γ 1E^{+/+}. Four γ 1E^{-/-} clones and one γ 1E^{+/-} clone with comparable levels of AID expression (Fig. S1b, c) were selected for further functional analysis. To determine whether deletion of γ 1E has an impact on the efficiency of CSR, these five clones together with the parental cell line (pCH12) were induced to undergo CSR. Three days later, the percentage of IgA⁺ cells was determined by flow cytometry (Fig. 1b, c). The efficiency of CSR was reduced by ~50% in all four γ 1E^{-/-} clones when compared with the γ 1E^{+/-} control clone or the pCH12 cells (Fig. 1b, c), showing that the deletion of the γ 1E region results in defective CSR to IgA in CH12 cells, independent of differences in AID expression.

To determine whether deletion of the γ 1E region has an impact on S region transcription, we measured the levels of I μ -C μ and I α -C α GLTs by RT-qPCR (Fig. 1d). Surprisingly, deletion of the γ 1E region had no negative effect on S region transcription, and both transcripts appeared to accumulate in the γ 1E^{-/-} clones (Fig. 1d). Hence, the reduced CSR efficiency observed in γ 1E^{-/-} clones cannot be explained by defects in transcription. A possible explanation is that the γ 1E region participates in the 3D conformational changes occurring at the IgH locus during CSR.^{4,5} Alternatively, the S α region may be insensitive to a putative transcriptional effect of the γ 1E region due to its close proximity to the 3'RR super-enhancer (2 kb).

To determine whether the γ 1E region has a role in CSR to other isotypes, we generated a γ 1E knockout mouse model using the CRISPR/Cas9 system and the same gRNAs used for the knockout in CH12 cells. Despite homozygous deletion of the γ 1E region, all the different B cell subsets were found to be represented in the bone marrow and in the spleen at normal numbers and proportions (Fig. S2), indicating that γ 1E deletion has no obvious effect on B cell development and that it does not lead to a block during differentiation.

To assess the ability of γ 1E-deficient B cells to undergo CSR, CFSE-labeled splenic B cells isolated from wild-type and γ 1E^{-/-} mice were cultured under conditions that induce CSR to different isotypes. After 72 h, we determined the surface expression of the different isotypes by flow cytometry (Fig. 1e, f). While the efficiency of CSR to IgG1 and IgA appeared similar between genotypes, a significant reduction in CSR to IgG3 (-32%), IgG2b (-86.5%), and IgG2a (-24.5%) was observed in γ 1E^{-/-} B cells compared with control B cells (Fig. 1e, f), independent of AID expression (Fig. S1d). We concluded that the γ 1E region has an isotype-specific role in the regulation of CSR.

¹Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Illkirch, France; ²Institut National de la Santé et de la Recherche Médicale (INSERM), U1258, Illkirch, France; ³Centre National de la Recherche Scientifique (CNRS), UMR7104, Illkirch, France and ⁴Université de Strasbourg, Illkirch, France
Correspondence: Bernardo Reina-San-Martin (reinab@igbmc.fr)

Received: 5 July 2019 Accepted: 10 July 2019
Published online: 05 August 2019

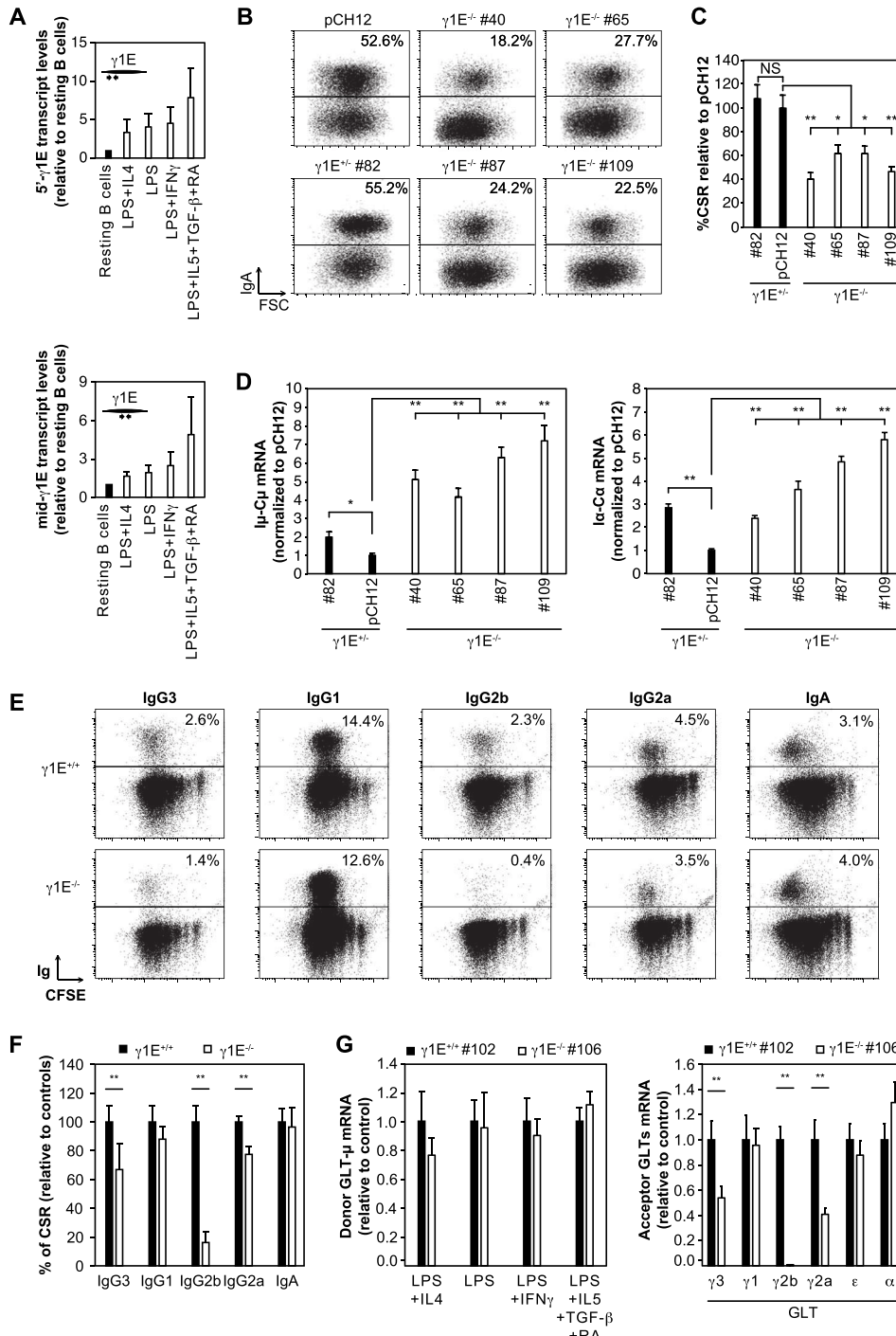


Fig. 1 Role of the $\gamma 1E$ region in CSR and IgH locus transcription in CH12 and primary B cells. **a** RT-qPCR of $\gamma 1E$ transcript levels using primers located in the 5' (top) or middle (bottom) region of the $\gamma 1E$. **b** Surface expression of IgA analyzed by flow cytometry in $\gamma 1E^{+/-}$ and $\gamma 1E^{-/-}$ clones and pCH12 cells after 3 days in culture with TGF- β , IL-4, and anti-CD40 antibody. The percentage of switched cells is indicated. Representative dot plots of six experiments are shown. **c** Percentage of CSR in $\gamma 1E^{+/-}$ and $\gamma 1E^{-/-}$ clones relative to pCH12 cells. Data are pooled from six independent experiments. Statistical significance was determined by two-tailed Student's *t*-test. **d** RT-qPCR for μ -C μ (left panel) and α -C α (right panel) transcripts. Triplicates were normalized to the abundance of Ig β , set as 1. Statistical significance was determined by a two-tailed Student's *t*-test. (**p*-value < 0.05; ***p*-value < 0.005). Data are representative of six experiments. **e** Flow cytometry analysis of surface Ig expression in CFSE-labeled primary B cells purified from $\gamma 1E^{+/-}$ or $\gamma 1E^{-/-}$ mice and cultured in vitro for 3 days with LPS (CSR to IgG3 and IgG2b), LPS + IL4 (CSR to IgG1 and IgE), or LPS + IFN γ (CSR to IgG2a) or for 4 days with LPS + IL5 + TGF β + retinoic acid (RA) (CSR to IgA). The percentage of switched cells is indicated. Representative dot plots from six experiments are shown. **f** CSR efficiency in primary B cells obtained from $\gamma 1E^{-/-}$ (*n* = 13) or $\gamma 1E^{+/-}$ (*n* = 12) mice, shown as the percentage of CSR. The mean + SD from six experiments is presented relative to $\gamma 1E^{+/+}$ control B cells. Statistical significance was determined by two-tailed Student's *t*-test (**p*-value < 0.05; ***p*-value < 0.005). **g** RT-qPCR for donor (μ -C μ ; left panel) and acceptor ($\gamma 3$ -C $\gamma 3$, $\gamma 1$ -C $\gamma 1$, $\gamma 2b$ -C $\gamma 2b$, $\gamma 2a$ -C $\gamma 2a$, ϵ -C ϵ , and α -C α ; right panel) GLT transcripts in primary B cells purified from $\gamma 1E^{+/+}$ or $\gamma 1E^{-/-}$ mice and cultured as in **e**. Mean + SD of triplicate values was normalized to the abundance of Ig β and is shown relative to $\gamma 1E^{+/+}$ B cells, set as 1. Statistical significance was determined by a two-tailed Student's *t*-test. (**p*-value < 0.05; ***p*-value < 0.005). For biological replicates, refer to Fig. S1e

Consistent with the reduction in CSR observed for IgG3, IgG2b, and IgG2a, we found that the level of the corresponding GLTs (Iy3-Cy3, Iy2b-Cy2b, and Iy2a-Cy2a) was significantly decreased in $\gamma 1E^{-/-}$ B cells compared with controls (Figs. 1g and S1e), while no differences were found for the Iy1-Cy1, Ie-Ce, and Ia-Ca transcript levels (Figs. 1g and S1e). We concluded that $\gamma 1E$ deletion impairs sterile transcription at the $\gamma 3$, $\gamma 2b$ and $\gamma 2a$ genes, perfectly correlating with the isotype-specific deficiencies observed in CSR.

Given the defect in CSR to IgA revealed in CH12 cells, it was surprising to find that CSR to IgA was not affected in primary $\gamma 1E^{-/-}$ B cells. Notably, CH12 cells present a B1 lineage-like phenotype, and B1 B cells are known to be biased towards CSR to IgA.¹² Interestingly, the 3'RR super-enhancer did not regulate CSR to IgA in B1 cells.¹³ In this regard, a lineage-specific regulatory role of the 3'RR and the $\gamma 1E$ region could be a plausible explanation for the disparities observed in the transcriptional regulation of $\gamma 1E^{-/-}$ CH12 and primary B cells.

The role of insulator regions in the transcriptional regulation of the IgH locus is also notable. In this regard, a recent study showed that the 5'hs1RI region acts as an insulator, restraining the 3'RR super-enhancer effect, and that it contributes to the regulation of CSR to IgG3, IgG2b, and IgG2a by blocking the premature activation of their corresponding promoters.¹⁴ Therefore, upon activation, the $\gamma 1E$ region may counteract the insulator function of the 5'hs1RI region. Hence, deletion of the $\gamma 1E$ region could result in unchecked 5'hs1RI insulator activity, resulting in defective transcription and reduced CSR to these isotypes.

Although we found an effect of $\gamma 1E$ in the transcriptional regulation of the switch region promoters, which could explain the defect in IgG3, IgG2b, and IgG2a CSR, we cannot exclude the possibility that the $\gamma 1E$ region is involved in the dynamic conformational changes of the IgH locus that occur during CSR. Indeed, this region is brought to the E μ enhancer and the 3'RR super-enhancer during CSR after B cell stimulation.⁵ Therefore, deletion of this region might affect IgH locus looping and consequently transcription.

Altogether, our results are consistent with a model in which the $\gamma 1E$ region regulates IgH locus transcription and CSR in an isotype-specific manner.

ACKNOWLEDGEMENTS

We thank members of the Reina-San-Martin laboratory for discussions, A. Khamlichi for comments on the manuscript, C. Ebel for assistance with cell sorting, M. Gendron for animal care, and Marie-Christine Birling for help with the generation of the $\gamma 1E^{-/-}$ mouse model. R.A.-V. was supported by the IGBMC's International PhD Program LABEX fellowship and by the Fondation Recherche Médicale. This study was supported 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 labeled ANR-10-IDEX-0002-02.

AUTHOR CONTRIBUTIONS

Conceptualization, B.R.-S.-M., R.A.-V., M.R., and I.R. Methodology, B.R.-S.-M., R.A.-V., and I.R. Investigation, R.A.-V., M.R., I.R., and V.H. Writing—original draft, R.A.-V. Writing—review and editing, B.R.-S.-M. and I.R. Supervision, B.R.-S.-M. and I.R. Funding Acquisition, B.R.-S.-M.

ADDITIONAL INFORMATION

The online version of this article (<https://doi.org/10.1038/s41423-019-0267-4>) contains supplementary material.

Competing interests: The authors declare no competing interests.

REFERENCES

- Muramatsu, M. et al. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* **102**, 553–563 (2000).
- Yewdell, W. T. & Chaudhuri, J. A transcriptional serenAID: the role of noncoding RNAs in class switch recombination. *Int. Immunol.* **29**, 183–196 (2017).
- Chaudhuri, J. et al. Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv. Immunol.* **94**, 157–214 (2007).
- Wuerffel, R. et al. S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity* **27**, 711–722 (2007).
- Thomas-Claudepierre, A. S. et al. Mediator facilitates transcriptional activation and dynamic long-range contacts at the IgH locus during class switch recombination. *J. Exp. Med.* **213**, 303–312 (2016).
- Medvedovic, J. et al. Flexible long-range loops in the VH gene region of the Igh locus facilitate the generation of a diverse antibody repertoire. *Immunity* **39**, 229–244 (2013).
- Predeus, A. V. et al. Targeted chromatin profiling reveals novel enhancers in Ig H and Ig L chain Loci. *J. Immunol.* **192**, 1064–1070 (2014).
- Li, W., Notani, D. & Rosenfeld, M. G. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat. Rev. Genet.* **17**, 207–223 (2016).
- Nakamura, M. et al. High frequency class switching of an IgM+ B lymphoma clone CH12F3 to IgA+ cells. *Int. Immunol.* **8**, 193–201 (1996).
- Arnold, L. W., Grdina, T. A., Whitmore, A. C. & Houghton, G. Ig isotype switching in B lymphocytes. Isolation and characterization of clonal variants of the murine Ly-1+ B cell lymphoma, CH12, expressing isotypes other than IgM. *J. Immunol.* **140**, 4355–4363 (1988).
- Ono, S. J. et al. Identification of a stimulus-dependent DNase I hypersensitive site between the Ialpha and Calpha exons during immunoglobulin heavy chain class switch recombination. *FEBS Lett.* **467**, 268–272 (2000).
- Kaminski, D. A. & Stavnezer, J. Enhanced IgA class switching in marginal zone and B1 B cells relative to follicular/B2 B cells. *J. Immunol.* **177**, 6025–6029 (2006).
- Issaoui, H. et al. The IgH 3' regulatory region super-enhancer does not control IgA class switch recombination in the B1 lineage. *Cell. Mol. Immunol.* **15**, 289–291 (2018).
- Braikia, F. Z. et al. Inducible CTCF insulator delays the IgH 3' regulatory region-mediated activation of germline promoters and alters class switching. *Proc. Natl Acad. Sci. USA* **114**, 6092–6097 (2017).

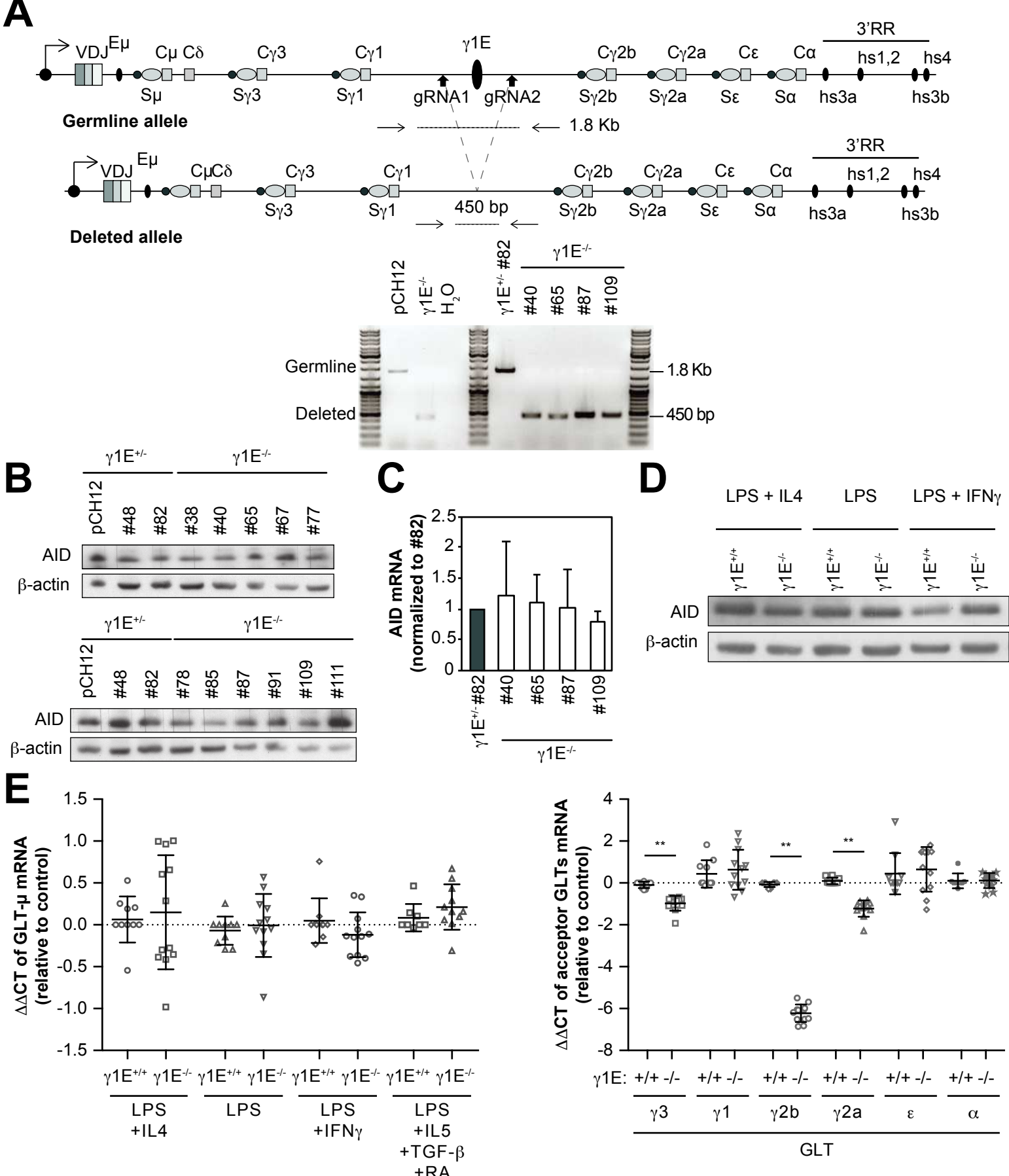
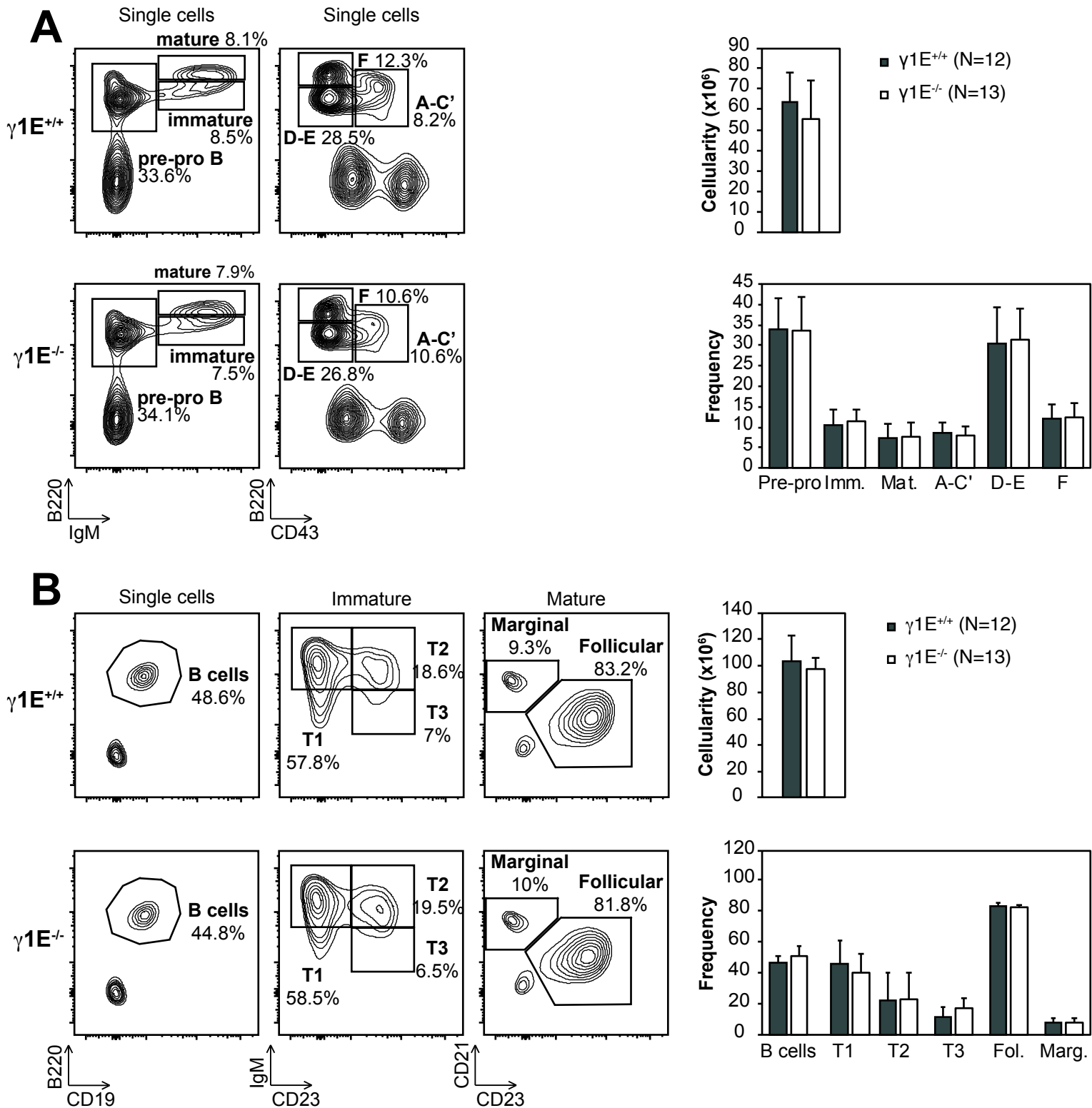


Figure S1. Generation and characterization of $\gamma 1E^{-/-}$ CH12 cells and $\gamma 1E^{-/-}$ mouse model. A (top). Schematic representation of the IgH locus before and after CRISPR/Cas9-mediated deletion of the $\gamma 1E$ region. gRNAs and PCR primers for genotyping are indicated. A (bottom) Representative agarose gel of genotyping by PCR. The bands corresponding to the germline or the deleted alleles are indicated. B. Western blot analysis for AID and β -Actin of CH12 cell clones of different genotypes after 72h in culture with TFG- β , IL-4 and anti-CD40 antibody. Theoretical molecular weights (kDa) are indicated. C. RT-qPCR for AID expression in selected CH12 clones after 72h in culture with TFG- β , IL-4 and anti-CD40 antibody. Mean+SD of AID values normalized to Ig β and relative to $\gamma 1E^{+/-}$ clone #82 from 4 independent experiments is shown. D. Western blot analysis for AID and β -Actin of primary B cells from $\gamma 1E^{+/+}$ and $\gamma 1E^{-/-}$ mice cultured for 72h with LPS+IL4, LPS, or LPS+IFN γ . E. RT-qPCR of the donor (left) and acceptor (right) GLTs of $\gamma 1E^{+/+}$ and $\gamma 1E^{-/-}$ B cells cultured as in Fig. 1E. Mean \pm SD of biological replicates from 5 independent experiments were normalized to the abundance of Ig β . Two $\gamma 1E^{+/+}$ mice were used for each experiment, and one of them was randomly set as control. Statistical significance was determined by a one-way ANOVA (**p-value < 0.005).

Amoretti-Villa et al.



Materials and Methods

Mice

The $\gamma 1E^{-/-}$ mutant mouse line was established at the MCI/ICS (Mouse Clinical Institute - Institut Clinique de la Souris-, Illkirch, France; <http://www.ics-mci.fr/en/>). Briefly, *in vitro* transcribed Cas9 mRNA (10 ng/ μ l) and 2 gRNAs (10 ng/ μ l) were co-injected into 803 C57BL/6N fertilized oocytes. Seventy pups were born. Two founders with the expected deletion were identified, one of which gave germline transmission. The $\gamma 1E^{-/-}$ line was maintained and bred under specific pathogen-free conditions. All animals used in this study were 8 to 12 weeks old. All animal work was performed under protocols approved by an ethics committee.

Cell culture and flow cytometry

CH12 cells were induced to undergo CSR with IL-4 (5 ng/ml; Sigma), TGF- β (3 ng/ml; R&D Systems), and anti-CD40 antibody (100 ng/ml; eBioscience) for 72h. Resting B cells were isolated from the spleen using CD43 Microbeads (Miltenyi Biotec), labeled with 5 μ g/ml CFSE (Invitrogen) for 10 minutes at 37°C, and cultured with 50 μ g/ml LPS (Sigma-Aldrich), 5ng/ml IL-4 (PeproTech), 100 ng/ml IFN- γ (PeproTech), 5ng/ml IL-5 (BD), 3 ng/ml TGF- β (R&D Systems), or 0.3 ng/ml retinoic acid (Sigma-Aldrich). Cells were stained with biotin-anti-IgG1 (BD), biotin-anti-IgG3 (BD), biotin-anti-IgG2b (BioLegend), biotin-anti-IgG2a (BD), PE-anti-IgA (SouthernBiotech) antibodies and PE-Streptavidin (Beckman Coulter). ToPro-3 (Invitrogen) was used to exclude dead cells in the analysis. For B cell development, bone marrow and splenic lymphocytes were stained with anti B220-PECy7 (e-biosciences), anti IgM-Cy5 (Southern Biotech), anti CD43-BB515 (BD Pharmingen), anti CD93-Biotin (e-biosciences), Streptavidine-BUV395 (BD Pharmingen), anti CD21-FITC (BD Biosciences), anti CD23-PE (BD Biosciences), and anti IgM-Cy5 (Southern Biotech) antibodies. All flow cytometry data were acquired in a Fortessa (BD Biosciences) flow cytometer and analyzed using the FlowJo software.

Generation of $\gamma 1E^{-/-}$ CH12 cells.

5×10^6 CH12 cells were electroporated with 0.5 μ g of a plasmid co-expressing 2 gRNAs and Cas9 fused to EGFP using a Neon Transfection System (Thermo Fisher Scientific). 24h later, EGFP+ cells were sorted with a BD FACSAria (BD Biosciences) and plated under limiting dilution conditions to obtain single-cell clones, which were cultured for approximately 2 weeks. Genomic DNA was extracted and individual clones were genotyped by PCR and sequencing. See table S1 for gRNAs and primer sequences.

Western blot

Proteins were fractionated on 4-12% gradient SDS-PAGE gels (Invitrogen), transferred to Immobilon PVDF membranes (Millipore) and stained with anti-AID (Strasbg9, AID-2E11; IGBMC) or β -actin (Sigma) antibodies.

RT-qPCR

RNA and cDNA from CH12 or primary B cells were obtained following standard protocols. qPCR was performed in triplicates using the Universal Probe Library system (Roche) or SyberGreen (QIAGEN) and a LightCycler 480 (Roche). Transcript quantities were calculated relative to standard curves and normalized to Ig β mRNA. PCR primers and probes used are shown in Table S1.

Materials and Methods

Table S1. Primers, probes and gRNAs			
RT-qPCR			
Primer	Sequence (5'-3')	Probe and Reference	
Igβ-F	TGGTGCTGTCTTCCATGC	UPL Probe 18 ¹	
Igβ-R	TTGCTGGTACCGGCTCAC	UPL Probe 18 ¹	
Iμ-Cμ-F	ACCTGGGAATGTATGGTTGTGGCTT	2	
Iμ-Cμ-R	TCTGAACCTTCAAGGATGCTCTTG	2	
Iμ-Cμ2-F	CCCAGACCTGGGAATGTATG	UPL Probe 29 ³	
Iμ-Cμ2-R	GGAAGACATTTGGGAAGGACT	UPL Probe 29 ³	
Iγ3-Cγ3-F	GCAGAAATCTGCAGGACTAACA	UPL Probe 71 ³	
Iγ3-Cγ3-R	ACCGAGGATCCAGATGTGTC	UPL Probe 71 ³	
Iγ2b-Cγ2b-F	TGGGCCTTTCCAGACCTAAT	UPL Probe 88 ³	
Iγ2b-Cγ2b-R	GGGCTGATCTGTCAACTCCT	UPL Probe 88 ³	
Iγ2a-Cγ2a-F	CAGCCTGGGATCAAGCAG	UPL Probe 109 ³	
Iγ2a-Cγ2a-R	TGGGGCTGTTGTTTTGGT	UPL Probe 109 ³	
Iγ1-Cγ1-F	GGCCCTCCAGATCTTTGAG	4	
Iγ1-Cγ1-R	ATGGAGTTAGTTTGGGCAGCA	4	
Iμ-Cε-F	CCCAGACCTGGGAATGTATG	UPL Probe 29 ⁵	
Iμ-Cε-R	GGGTAGAGCTGAGGGTTCCT	UPL Probe 29 ⁵	
Iμ-Cα-F	GGAGACTCCCAGGCTAGACA	UPL Probe 27 ¹	
Iμ-Cα-R	CGGAAGGGAAGTAATCGTGA	UPL Probe 27 ¹	
Genotyping primers (CH12 cells and γ1E^{-/-} mice)			
Primer	Sequence (5'-3')		
g1E-F	GCTGAGCAAAACACCACCTG		
g1E-R	GACCTCTCCAGTTTCGGCTG		
gRNAs			
Primer	Sequence (5'-3')	PAM	Strand
gRNA1	AACATTGGCCTCCCAACA	TGG	-
gRNA2	CCAATGGCATTGGTAACC	AGG	+

1. Thomas-Claudepierre, A.S. *et al.* The cohesin complex regulates immunoglobulin class switch recombination. *The Journal of experimental medicine* **210**, 2495-502 (2013).

2. Jeevan-Raj, B.P. *et al.* Epigenetic tethering of AID to the donor switch region during immunoglobulin class switch recombination. *The Journal of experimental medicine* **208**, 1649-60 (2011).

3. Robert, I. *et al.* Parp3 negatively regulates immunoglobulin class switch recombination. *PLoS genetics* **11**, e1005240 (2015).

4. Park, S.R. *et al.* HoxC4 binds to the promoter of the cytidine deaminase AID gene to induce AID expression, class-switch DNA recombination and somatic hypermutation. *Nature immunology* **10**, 540-50 (2009).

5. Thomas-Claudepierre, A.S. *et al.* Mediator facilitates transcriptional activation and dynamic long-range contacts at the IgH locus during class switch recombination. *J Exp Med* **213**, 303-12 (2016).

Amoretti-Villa *et al.*

II- Discussion

Afin d'étudier le rôle de l'enhancer putatif $\gamma 1E$ dans la régulation transcriptionnelle du locus IgH durant la CI, nous avons généré un modèle CH12 $\gamma 1E^{-/-}$. L'absence de $\gamma 1E$ dans la lignée cellulaire CH12 a révélé un défaut de CI et de manière surprenante, des niveaux normaux de transcrits $S\alpha$. Pour expliquer ce manque de corrélation, au vu de la proximité relative entre $S\alpha$ et le super-enhancer 3'RR, nous pensions que celui-ci pouvait jouer un rôle dans la régulation de $S\alpha$. Cependant, il a été montré que le 3'RR ne régule pas la CI vers IgA dans les cellules B de type B1 et que son absence ne réduit que modestement la CI vers IgA dans les CH12 (Issaoui et al., 2018; Kim et al., 2016). La lignée cellulaire CH12 présentant un phénotype semblable à la lignée B1 (Kaminski and Stavnezer, 2006), il n'est donc pas possible que la régulation des transcrits $S\alpha$ dans les cellules CH12 $\gamma 1E^{-/-}$ provienne du super-enhancer 3'RR. La cause de ce manque de corrélation n'est donc pas claire et met en avant des caractéristiques particulières de la lignée CH12. En effet, il est important de noter qu'il existe déjà un niveau basal de GLT $S\alpha$ sans induction de la CI. Le fait que la CI vers IgA puisse se produire sans le 3'RR dans la lignée cellulaire CH12 indique que celui-ci ne dispose pas de motif de ciblage d'AID en soi pour la CI. Il est donc probable que le locus IgH dans les cellules CH12 soit déjà en conformation de boucle, contournant ainsi l'exigence du 3'RR. Etant donné que la régulation transcriptionnelle du locus IgH dans la lignée cellulaire CH12 montre de grandes différences avec celle des cellules B primaires (Issaoui et al., 2018; Kim et al., 2016), le modèle cellulaire CH12 n'est certainement pas le meilleur modèle pour étudier le rôle d'un enhancer dans la régulation transcriptionnelle du locus IgH.

Afin d'étudier le rôle de $\gamma 1E$ vers les autres isotypes, nous avons donc généré un modèle souris $\gamma 1E^{-/-}$ en utilisant le système CRISPR/Cas9. Avec les précédentes études identifiant un enhancer putatif $\gamma 1E$ dans la région en aval de $C\gamma 1$ dans les cellules Pro-B (Medvedovic et al., 2013; Predeus et al., 2014), nous avons émis l'hypothèse que cet enhancer pouvait jouer un rôle dans la recombinaison V(D)J. Afin d'évaluer celui-ci, nous nous sommes donc intéressés au développement des

cellules B provenant de la moelle osseuse et de la rate. L'absence de défaut dans les différentes populations analysées affiche que le développement des cellules B n'est pas affecté par l'inactivation de $\gamma 1E$, suggérant que $\gamma 1E$ ne contrôle pas la recombinaison V(D)J. L'analyse de l'efficacité de CI dans les cellules B primaires $\gamma 1E^{-/-}$ vers les autres isotypes a révélé un défaut de CI vers les isotypes IgG3, IgG2b et IgG2a mais pas vers IgA. La divergence des résultats de CI vers IgA dans les cellules CH12 $\gamma 1E$ et les cellules B primaires $\gamma 1E$ peut s'expliquer par le fait que la lignée cellulaire CH12 est de type B1 tandis que les cellules B primaires spléniques sont de type B2 (Kaminski and Stavnezer, 2006). Plusieurs études ont reporté des différences majeures entre les cellules B1 et B2 en regardant leur capacité à induire la CI vers IgA (Issaoui et al., 2018; Kaminski and Stavnezer, 2006; Kim et al., 2016; Li et al., 2018). L'observation d'une réduction des niveaux des GLTs $S\gamma 3$, $S\gamma 2b$ et $S\gamma 2a$ corrèle avec le défaut de CI observé aux isotypes correspondants. Cette corrélation permet de conclure que l'enhancer $\gamma 1E$ régule la transcription du locus IgH de manière spécifique à l'isotype. Ceci n'est pas une première dans le locus IgH car le super-enhancer 3'RR montre aussi une très faible régulation de l'isotype IgG1 comparé aux autres isotypes (Pinaud et al., 2001; Vincent-Fabert et al., 2010).

Néanmoins, étant donné que notre laboratoire a montré qu'après l'activation des cellules B la région de $\gamma 1E$ est recrutée dans la boucle principale à proximité des enhancers $E\mu$ et du 3'RR durant la CI, il est possible que l'enhancer putatif $\gamma 1E$ joue aussi un rôle dans l'architecture du locus IgH et plus précisément, dans le maintien de la boucle d'ADN entre enhancer/promoteur durant la CI. Afin de dresser une carte génomique regroupant tous les éléments régulateurs du locus IgH, ainsi que leur fonction, cette hypothèse nécessite de plus amples investigations.

Conclusion générale

Conclusion générale

La compréhension de l'immunité adaptative médiée par les cellules B a évolué depuis la découverte d'AID, il y a environ vingt ans. Les recherches menées ces dernières années ont révolutionné notre compréhension sur les mécanismes de diversification des anticorps, l'hypermutation somatique et la commutation isotypique, initiés par AID, ainsi que sur les mécanismes protégeant l'intégrité du génome des cellules B. L'induction de dommages à l'ADN induits par AID est restreinte aux régions S et V aux loci des Ig durant la CI et l'HMS respectivement, mais AID peut également cibler des gènes non-Ig, menaçant ainsi la stabilité génomique des lymphocytes B. Cependant, les cellules B ont développé des mécanismes qui favorisent la réparation de l'ADN mutagène et recombino-gène dépendant d'AID au sein des loci Ig tout en réparant fidèlement les dommages collatéraux des loci non-Ig en utilisant les voies de réparation de l'ADN canoniques. Des mécanismes de ciblage d'AID aux régions S et V ainsi que de nombreux facteurs contribuant à la CI et à l'HMS ont déjà été mis en lumière mais ne sont pas suffisants pour répondre à toutes les questions encore non élucidées.

Avec le but de clarifier la régulation des mécanismes de CI, nous nous sommes concentrés sur l'identification de mécanismes additionnels encore inconnus requis pour une CI efficace. Les différents projets menés durant ma thèse ont conduit à l'identification de nouveaux acteurs moléculaires ou séquence régulatrice requis pour les différentes étapes de la réaction de recombinaison. Avec une étude complète du rôle de Fam72a, nous espérons pouvoir apporter un élément de réponse au mécanisme précis menant à la réparation mutagène des résidus uraciles au sein des loci Ig. La compréhension des mécanismes fondamentaux impliqués durant la diversification des anticorps a des implications directes dans l'amélioration de l'efficacité des vaccinations ainsi que dans la lutte de l'apparition des lymphomes B.

Matériels et Méthodes

Clonage

Les vecteurs ont été générés selon les procédures classiques de clonage ou assemblés par une réaction unique de digestion/ligation à l'aide d'enzyme de type IIS, à partir de modules provenant de l'UES (Système d'Expression Universel-Labo Reina). Les vecteurs d'expression utilisés pour les projets principaux sont répertoriés dans le tableau ci-dessous.

Vecteur	Construction	Nom utilisé
pMXPIE	Kozak-Cas9-mCherry-P2A-HygroR	pMX-Cas9
pMXPIE	Lacz-CMV-EGFP-SV40-PuroR	pMX-librairies
pMXPIE	Librairie mGeCKO-v2-CMV-EGFP-SV40-PuroR	pMX-mGeCKO-v2
pMXPIE	Librairie mBrie-CMV-EGFP-SV40-PuroR	pMX-mBrie
pMXPIE	U6-LacZ-oScaffold-CMV-EGFP-SV40-PuroR	pMX-25
pMXPIE	mFam72a	pMX-mFam72a
pMXPIE	mFam72aW125R	pMX-mFam72a ^{W125R}
pMXPIE	hFam72a	pMX-hFam72a
pMXPIE	hFam72b	pMX-hFam72b
pMXPIE	hFam72c	pMX-hFam72c
pMXPIE	hFam72d	pMX-hFam72d
pLentiCRISPR-V2	U6-LacZ-oScaffold-EFS-Cas9-EGFP-P2A-PuroR	pLentiCRISPR-v6/pL6
pX-883	pGH-Fam72a.1-pGH-Fam72a.2-CMV-Kozak-Cas9-HF-SV40-PuroR	pX-Fam72a
pX-903	pGH-Ung1.1-pGH-Ung1.2-CMV-Kozak-Cas9-HF-SV40-PuroR	pX-Ung1
pX-904	pGH-Ung2.1-pGH-Ung2.2-CMV-Kozak-Cas9-HF-SV40-PuroR	pX-Ung2
pX-905	pGH-Ung.1-pGH-Ung.2-CMV-Kozak-Cas9-HF-SV40-PuroR	pX-Ung

CMV : promoteur du CytoMégaloVirus

EGFP: enhanced green fluorescent protein

U6: promoteur ARN polymérase II

EFS: promoteur EF1-Alpha short

oScaffold: Scaffold optimisé selon (Chen et al., 2013)

pGH:ARNg pour la Cas9-HF

Clonage librairie contrôle (Preuve de principe du crible)

Les amorces contenant la séquence de l'ARNg ont tout d'abord été hybridées puis clonées dans le vecteur pMX-25 en une réaction de digestion (SapI)/ligation. La séquence des différents ARNg utilisés dans la librairie contrôle sont répertoriés dans le tableau ci-dessous.

Gène	Séquence ARNg	Gène	Séquence ARNg
NT-1	GCGAGGTATTCCGGCTCCGCG	Kap1-1	GAGGAACCAACGTAAACTCT
NT-2	GCTTTCACGGAGGTTTCGACG	Kap1-2	GGACCTGCTAAGACTCGAGA
NT-3	GACTCCGGGTACTAAATGTC	Kap1-3	GCTCTCACAGAACAGCACGA
NT-4	GCTTCTACTCGCAACGTATT	Parp1-1	GATGCGCCTGTCCAAGAAGA
NT-5	GTCAAGCCGAACGCTGCCGG	Parp1-2	GGACAAGGATAGTAGTAAGC
NT-6	GCGGTTACCGCGAAAACCAT	Parp1-3	GCGCTGGGATGATCAGCAGA
AID-1	GTAGGAACAACAATTCCACG	Parp9-1	GAATCGCTATCACCGGTGCG
AID-2	GTGACGCGGTAACACCGGCC	Parp9-2	GATTCATCTGGTGAGCAATG
AID-3	GCCGGTGTTACCGCGTCACC	Parp9-3	GCCGATGCGAAGAGTCATGG
UNG-1	GAGCTGTCTACAGACATCGA	CTNNBL1-1	GCTTACTTCTCCGGATTGTC
UNG-2	GCCGGTCCCAGGGAGCGGAG	CTNNBL1-2	GCTGCTGCAGATTATTGACA
UNG-3	GGAGTGCGGCGGAGATCGG	CTNNBL1-3	GCGCCTGGATGAGTCTGTGA
53BP1-1	GCTGAGAATCTTCTATTATC	Parp3-1	GGGTACAGTCATAGTCCTCG
53BP1-2	GCACAAAGAGAACCCCGTGT	Parp3-2	GATGAACCACTTCACCTGCC
53BP1-3	GGGGTGGGAGCAAACGCTCC	Parp3-3	GGAGCCCTCAGTCTGCACAG
Med1-1	GCCAGCAATGACTGATCGTT	Smc1-1	GGACCGCTATCAGCGCCTGA
Med1-2	GATGTAACACTCAGTGCCAC	Smc1-2	GCCATTGGGTCCAATGATGG
Med1-3	GTGTTGGTTATCTCACCCCG	Smc1-3	GGAGATTGAGAAGCTCAACA
H2AX-1	GGCGCCGCGGTCGGCAAGA	Smc3-1	GTATTTCTAGATAAGAAGA
H2AX-2	GAGGCCTTCTTGCCGACCGC	Smc3-1	GCTGGCCAGTATCAGAAGT
H2AX-3	GCCGGCCTTCAGTTCCTGT	Smc3-1	GGAGTACACCATCTACAACC
ATM-1	GGATCACGGAGTACATCCAG	SPT5-1	GGAAGAGGAAGAATACGATG
ATM-2	GTGAAGTATCTCAGCAACAG	SPT5-2	GCAGACGACGAGCCCCAGAG
ATM-3	GCAAGGTACTGTAATCAG	SPT5-3	GATGACCGATGTAATCAAGG

Infection lentivirale

Les particules lentivirales sont produites à partir de cellules HEK 293T par co-transfection en utilisant le réactif Fugene6 (Roche) supplémenté avec 1µg du vecteur lentiviral d'intérêt (LentiCRISPR-v6) et 1µ du vecteur d'emballage (combinaison de vecteurs auxiliaires testés répertoriés dans la **Figure 16**). 48h après transfection, l'efficacité de transfection est analysée par cytométrie en flux. Les surnageants lentiviraux sont ensuite récupérés, filtrés, supplémentés avec du polybrène (10µg/mL) et de l'Hepes (20mM) et ajoutés aux cellules CH12. Les cellules sont ensuite centrifugées à 2300 rpm pendant 90 minutes puis incubées à 37°C. 24h après infection, l'efficacité d'infection est analysée par cytométrie en flux.

Test de l'activité d'Ung

Le test de l'activité d'Ung a été réalisé comme décrit précédemment (Cortizas et al., 2016). Les cellules ont été lavées dans du PBS 1X, resuspendues dans du tampon HEN (25mM Hepes, 5mM NaCl, 5mM EDTA, 1M DTT et 10% de glycérol) avec un cocktail d'inhibiteurs de protéase (PIC) (Roche), et lysées par sonication (Q700). Après centrifugation à 14 000 rcf pendant 20 min à 4°C, le surnageant a été congelé en aliquots dans l'azote liquide et stocké à -80°C. Les tests UNG ont été réalisés dans un tampon HEN en mélangeant 10 µg d'extrait cellulaire avec 1 pmol de substrat oligonucléotidique marqué à la fluorescéine dans un volume final de 10 µl pendant 2 h à 37°C. L'oligonucléotide double brin avec un seul mésappariement dU/dG a été obtenu par hybridation 5' Biotine-ATTATTATTATTCCGUGGATTTATTTATTTATT TATTTATTT-Fluorescéine à l'oligonucléotide complémentaire, 5' AAATAAATAAA TAAATAAATAAATCCGCGGAATAATAATAAT 3'. La réaction a été arrêtée par l'ajout de 10 µl de formamide, et les produits ont été résolus sur un gel de polyacrylamide à 15 % de Tris/borate/EDTA (TBE)-urée. Le signal de fluorescéine des produits de la réaction a été visualisé à l'aide d'un Typhoon PhosphorImager (GE Healthcare).

Bibliographie

- 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.
- Akbari, M. 2004. Repair of U/G and U/A in DNA by UNG2-associated repair complexes takes place predominantly by short-patch repair both in proliferating and growth-arrested cells. *Nucleic Acids Research* 32:5486-5498.
- 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. Ilicic, 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, R., R. Armitage, M. Conley, H. Rosenblatt, N. Jenkins, N. Copeland, M. Bedell, S. Edelhoff, C. Disteche, D. Simoneaux, and A. Et. 1993. CD40 ligand gene defects responsible for X-linked hyper-IgM syndrome. *Science* 259:990-993.
- Alt, F.W., G.D. Yancopoulos, T.K. Blackwell, C. Wood, E. Thomas, M. Boss, R. Coffman, N. Rosenberg, S. Tonegawa, and D. Baltimore. 1984. Ordered rearrangement of immunoglobulin heavy chain variable region segments. *The EMBO Journal* 3:1209-1219.
- 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.
- Amoretti-Villa, R., M. Rogier, I. Robert, V. Heyer, and B. Reina-San-Martin. 2019. A novel regulatory region controls IgH locus transcription and switch recombination to a subset of isotypes. *Cell Mol Immunol* 16:887-889.
- Anderson, L., C. Henderson, and Y. Adachi. 2001. Phosphorylation and Rapid Relocalization of 53BP1 to Nuclear Foci upon DNA Damage. *21:1719-1729.*
- 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. Faili, 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. 2002. Requirement of the Activation-Induced Deaminase (AID) Gene for Immunoglobulin Gene Conversion. *Science* 295:1301-1306.
- 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.

- Aruffo, A., M. Farrington, D. Hollenbaugh, X. Li, A. Milatovich, S. Nonoyama, J. Bajorath, L.S. Grosmaire, R. Stenkamp, M. Neubauer, R.L. Roberts, R.J. Noelle, Jeffrey A. Ledbetter, U. Francke, and a.H.D. Ochs. 1993. The CD40 Ligand, gp39, Is Defective in Activated T Cells from Patients with X-Linked Hyper-IgM Syndrome *Cell* 72:291-300.
- 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 β prevents AID-instigated mutagenic non-canonical mismatch DNA repair. In Cold Spring Harbor Laboratory.
- Bannard, O., and J.G. Cyster. 2017. Germinal centers: programmed for affinity maturation and antibody diversification. *Current Opinion in Immunology* 45:21-30.
- 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. 2003. C-Terminal Deletion of AID Uncouples Class Switch Recombination from Somatic Hypermutation and Gene Conversion. *Molecular 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.
- Bassing, C.H., W. Swat, and F.W. Alt. 2002. The Mechanism and Regulation of Chromosomal V(D)J Recombination. *Cell* 109:S45-S55.
- Basso, K., and R. Dalla-Favera. 2015. Germinal centres and B cell lymphomagenesis. *Nature Reviews Immunology* 15:172-184.
- 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. *The Journal of Experimental Medicine* 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, D.R. Wesemann, K. Januszzyk, 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.
- 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., K. Kinoshita, N. Kakazu, M. Muramatsu, H. Nagaoka, R. Shinkura, D. Biniszkiwicz, 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. *Journal of Biological Chemistry* 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.
- Bergqvist, P., A. Stensson, N.Y. Lycke, and M. Bemark. 2010. T Cell-Independent IgA Class Switch Recombination Is Restricted to the GALT and Occurs Prior to Manifest Germinal Center Formation. 184:3545-3553.

- 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.
- Bishop, G.A., and G. Haughton. 1986. Induced differentiation of a transformed clone of Ly-1+ B cells by clonal T cells and antigen. *Proceedings of the National Academy of Sciences* 83:7410-7414.
- 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.
- Blanco-Betancourt, C.E. 2004. Defective B-cell-negative selection and terminal differentiation in the ICF syndrome. *103*:2683-2690.
- Boboila, C., F.W. Alt, and B. Schwer. 2012. 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., 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. 2003. 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. *Journal of Biological Chemistry* 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.
- Bunting, S.F., E. Callén, N. Wong, H.-T. Chen, F. Polato, A. Gunn, A. Bothmer, N. Feldhahn, O. Fernandez-Capetillo, L. Cao, X. Xu, C.-X. Deng, T. Finkel, M. Nussenzweig, J.M. Stark, and A. Nussenzweig. 2010. 53BP1 Inhibits Homologous Recombination in Brca1-Deficient Cells by Blocking Resection of DNA Breaks. *Cell* 141:243-254.

- Burma, S., B.P. Chen, M. Murphy, A. Kurimasa, and D.J. Chen. 2001. ATM Phosphorylates Histone H2AX in Response to DNA Double-strand Breaks. *Journal of Biological Chemistry* 276:42462-42467.
- 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. McDougle, 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.
- C. Berek, A.B., and M. Ape1. 1991. Maturation of the Immune Response in Germinal Centers. *Cell* 67:1121-1129.
- 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. 2016. Mutations, kataegis and translocations in B cells: understanding AID promiscuous activity. *Nature Reviews Immunology* 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 CTNBL1 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., 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.
- Chapman, J.R., P. Barral, J.-B. Vannier, V. Borel, M. Steger, A. Tomas-Loba, Alessandro, Ian, Facundo, and Simon. 2013. RIF1 Is Essential for 53BP1-Dependent Nonhomologous End Joining and Suppression of DNA Double-Strand Break Resection. *Molecular Cell* 49:858-871.
- 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, B., L.A. Gilbert, B.A. Cimini, J. Schnitzbauer, W. Zhang, G.W. Li, J. Park, E.H. Blackburn, J.S. Weissman, L.S. Qi, and B. Huang. 2013. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell* 155:1479-1491.
- Chen, Z., and J.H. Wang. 2019. Chapter Four - Signaling control of antibody isotype switching. In *Advances in Immunology*. F. Alt, editor Academic Press, 105-164.
- 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.

- Cheng, X., P. Deng, H. Cui, and A. Wang. 2015. Visualizing double-stranded RNA distribution and dynamics in living cells by dsRNA binding-dependent fluorescence complementation. *Virology* 485:439-451.
- 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. *J Immunol* 204:13-22.
- Chouery, E., J. Abou-Ghoch, S. Corbani, N. El Ali, R. Korban, N. Salem, C. Castro, S. Klayme, M. Azoury-Abou Rjeily, R. Khoury-Matar, G. Debo, M. Germanos-Haddad, V. Delague, G. Lefranc, and A. Megarbane. 2012. A novel deletion in ZBTB24 in a Lebanese family with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *Clin Genet* 82:489-493.
- Chowdhury, M., O. Forouhi, S. Dayal, N. McCloskey, H.J. Gould, G. Felsenfeld, and D.J. Fear. 2008. Analysis of intergenic transcription and histone modification across the human immunoglobulin heavy-chain locus. 105:15872-15877.
- Chu, V.T., R. Graf, T. Wirtz, T. Weber, J. Favret, X. Li, K. Petsch, N.T. Tran, M.H. Sieweke, C. Berek, R. Kuhn, and K. Rajewsky. 2016. Efficient CRISPR-mediated mutagenesis in primary immune cells using CrispRGold and a C57BL/6 Cas9 transgenic mouse line. *Proc Natl Acad Sci U S A* 113:12514-12519.
- Cobb, R.M., K.J. Oestreich, O.A. Osipovich, and E.M. Oltz. 2006. Accessibility control of V(D)J recombination. *Adv Immunol* 91:45-109.
- Cong, L., F.A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P.D. Hsu, X. Wu, W. Jiang, L.A. Marraffini, and F. Zhang. 2013. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* 339:819-823.
- 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.
- 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.D. Noia, and R.E. Verdun. 2016. UNG protects B cells from AID-induced telomere loss. *Journal of Experimental Medicine* 213:2459-2472.
- Cristina Rada, M.R.E., † Michael S. Neuberger,* and Cesar Milstein. 1998. Hot Spot Focusing of Somatic Hypermutation in MSH2-Deficient Mice Suggests Two Stages of Mutational Targeting. *Immunity* 9:135-141.
- 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.
- 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.
- 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 (New York, N.Y.)* 329:917-923.
- de Greef, J.C., J. Wang, J. Balog, J.T. den Dunnen, R.R. Frants, K.R. Straasheijm, C. Aytekin, M. van der Burg, L. Duprez, A. Ferster, A.R. Gennery, G. Gimelli, I. Reisli, C. Schuetz, A. Schulz, D. Smeets, Y. Sznajder, C. Wijmenga, M.C. van Eggermond, M.M. van Ostaijen-Ten Dam, A.C. Lankester,

- M.J.D. van Tol, P.J. van den Elsen, C.M. Weemaes, and S.M. van der Maarel. 2011. Mutations in ZBTB24 are associated with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *Am J Hum Genet* 88:796-804.
- de Miranda, N.F., A. Bjorkman, and Q. Pan-Hammarstrom. 2011. DNA repair: the link between primary immunodeficiency and cancer. *Ann N Y Acad Sci* 1246:50-63.
- 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 η is the sole contributor of A/T modifications during immunoglobulin gene hypermutation in the mouse. *Journal of Experimental Medicine* 204:17-23.
- Delgado-Benito, V., D.B. Rosen, Q. Wang, A. Gazumyan, J.A. Pai, T.Y. Oliveira, D. Sundaravinayagam, W. Zhang, M. Andreani, L. Keller, K.-R. Kieffer-Kwon, A. Pękowska, S. Jung, M. Driesner, R.I. Subbotin, R. Casellas, B.T. Chait, M.C. Nussenzweig, and M. Di Virgilio. 2018. The Chromatin Reader ZMYND8 Regulates Igh Enhancers to Promote Immunoglobulin Class Switch Recombination. *Molecular Cell* 72:636-649.e638.
- 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 Yebenes, and A.R. Ramiro. 2020. Interplay between UNG and AID governs intratumoral heterogeneity in mature B cell lymphoma. In Cold Spring Harbor Laboratory.
- Dhanjal, J.K., N. Radhakrishnan, and D. Sundar. 2019. CRISPCut: A novel tool for designing optimal sgRNAs for CRISPR/Cas9 based experiments in human cells. *Genomics* 111:560-566.
- 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., and M.S. Neuberger. 2007. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem* 76:1-22.
- 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.
- Di Virgilio, M., E. Callen, A. Yamane, W. Zhang, M. Jankovic, A.D. Gitlin, N. Feldhahn, W. Resch, T.Y. Oliveira, B.T. Chait, A. Nussenzweig, R. Casellas, D.F. Robbiani, and M.C. Nussenzweig. 2013. Rif1 Prevents Resection of DNA Breaks and Promotes Immunoglobulin Class Switching. *Science* 339:711-715.
- Diaz, A.A., H. Qin, M. Ramalho-Santos, and J.S. Song. 2015. HiTSelect: a comprehensive tool for high-complexity-pooled screen analysis. *Nucleic Acids Res* 43:e16.
- Dickerson, S.K., E. Market, E. Besmer, and F.N. Papavasiliou. 2003. 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. *Nature Structural & Molecular Biology* 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. *Nat Struct Mol Biol* 16:808-813.
- Disanto, J.P., J.Y. Bonnefoy, J.F. Gauchatt, A. Fischer, and G.D. Saint Basile. 1993. CD40 ligand mutations in X-linked immunodeficiency with hyper-IgM. *Nature* 361:541-543.
- 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.
- Doench, J.G., N. Fusi, M. Sullender, M. Hegde, E.W. Vaimberg, K.F. Donovan, I. Smith, Z. Tothova, C. Wilen, R. Orchard, H.W. Virgin, J. Listgarten, and D.E. Root. 2016. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 34:184-191.
- Doench, J.G., E. Hartenian, D.B. Graham, Z. Tothova, M. Hegde, I. Smith, M. Sullender, B.L. Ebert, R.J. Xavier, and D.E. Root. 2014. Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nature Biotechnology* 32:1262-1267.
- Döffinger, R., A. Smahi, C. Bessia, F. Geissmann, J. Feinberg, A. Durandy, C. Bodemer, S. Kenwrick, S. Dupuis-Girod, S. Blanche, P. Wood, S.H. Rabia, D.J. Headon, P.A. Overbeek, F. Le Deist, S.M. Holland, K. Belani, D.S. Kumararatne, A. Fischer, R. Shapiro, M.E. Conley, E. Reimund, H. Kalhoff, M. Abinun, A. Munnich, A. Israël, G. Courtois, and J.-L. Casanova. 2001. X-linked anhidrotic ectodermal dysplasia with immunodeficiency is caused by impaired NF- κ B signaling. *Nature Genetics* 27:277-285.
- 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.
- Dorsett, Y., K.M. McBride, M. Jankovic, A. Gazumyan, T.H. Thai, D.F. Robbiani, 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* 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.
- Du, L., R. Peng, A. Björkman, N. Filipe De Miranda, C. Rosner, A. Kotnis, M. Berglund, C. Liu, R. Rosenquist, G. Enblad, C. Sundström, M. Hojjat-Farsangi, H. Rabbani, M.R. Teixeira, P. Revy, A. Durandy, Y. Zeng, A.R. Gennery, J.-P. De Villartay, and Q. Pan-Hammarström. 2012. Cernunnos influences human immunoglobulin class switch recombination and may be associated with B cell lymphomagenesis. 209:291-305.
- 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., G.Z. Hertz, L. Scappino, and C. Gritzmacher. 1993. DNA sequences at immunoglobulin switch region recombination sites. *Nucleic Acids Research* 21:365-372.
- Duquette, M.L. 2004. Intracellular transcription of G-rich DNAs induces formation of G-loops, novel structures containing G4 DNA. *Genes & Development* 18:1618-1629.
- Durandy, A., S. Peron, N. Taubenheim, and A. Fischer. 2006. Activation-induced cytidine deaminase: structure-function relationship as based on the study of mutants. 27:1185-1191.

- Durandy, A., P. Revy, K. Imai, and A. Fischer. 2005. Hyper-immunoglobulin M syndromes caused by intrinsic B-lymphocyte defects. *203:67-79*.
- Durandy, A., C. Schiff, J.Y. Bonnefoy, M. Forveille, F. Rousset, G. Mazzei, M. Milili, and A. Fischer. 1993. Induction by anti-CD40 antibody or soluble CD40 ligand and cytokines of IgG, IgA and IgE production by B cells from patients with X-linked hyper IgM syndrome. *23:2294-2299*.
- 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. *J Immunol* 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, K. Kinoshita, T. Morisawa, T. Sakurai, I.M. Okazaki, K. Watashi, K. Shimotohno, T. Honjo, and T. Chiba. 2007. Expression of activation-induced cytidine deaminase in human hepatocytes via NF- κ B signaling. *Oncogene* 26:5587-5595.
- 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.
- Escribano-Diaz, C., A. Orthwein, A. Fradet-Turcotte, M. Xing, J.T. Young, J. Tkac, M.A. Cook, A.P. Rosebrock, M. Munro, M.D. Canny, D. Xu, and D. Durocher. 2013. A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. *Mol Cell* 49:872-883.
- Etzioni, A., A. Ben-Barak, S. Peron, and a.A. Durandy. 2007. Ataxia-Telangiectasia in Twins Presenting as Autosomal Recessive Hyper-Immunoglobulin M Syndrome. *IMAJ* 9:406-407.
- 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.
- Feldman, S., R. Wuerffel, I. Achour, L. Wang, P.B. Carpenter, and A.L. Kenter. 2017. 53BP1 Contributes to Igh Locus Chromatin Topology during Class Switch Recombination. *The Journal of Immunology* 198:2434-2444.
- Feng, L., K.-W. Fong, J. Wang, W. Wang, and J. Chen. 2013. RIF1 Counteracts BRCA1-mediated End Resection during DNA Repair. *Journal of Biological Chemistry* 288:11135-11143.
- Feng, Y., N. Seija, J. Di Noia, and A. Martin. 2020. AID in Antibody Diversification: There and Back again. *Trends in Immunology*
- Ferrari, S., S. Giliani, A. Insalaco, A. Al-Ghoniaim, A.R. Soresina, M. Loubser, M.A. Avanzini, M. Marconi, R. Badolato, A.G. Ugazio, Y. Levy, N. Catalan, A. Durandy, A. Tbakhi, L.D. Notarangelo, and A. Plebani. 2001. Mutations of CD40 gene cause an autosomal recessive form of immunodeficiency with hyper IgM. *Proceedings of the National Academy of Sciences* 98:12614-12619.
- 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.
- Fuleihan, R., N. Ramesh, R. Loh, H. Jabara, R.S. Rosen, T. Chatila, S.M. Fu, I. Stamenkovic, and R.S. Geha. 1993. Defective expression of the CD40 ligand in X chromosome-linked immunoglobulin deficiency with normal or elevated IgM. *Proceedings of the National Academy of Sciences* 90:2170-2173.
- 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.
- 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. *Molecular and Cellular Biology* 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. *Proceedings of the National Academy of Sciences* 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.
- 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*.

- Graf, R., X. Li, V.T. Chu, and K. Rajewsky. 2019. sgRNA Sequence Motifs Blocking Efficient CRISPR/Cas9-Mediated Gene Editing. *Cell Rep* 26:1098-1103 e1093.
- 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.
- 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. 2008. Cell cycle-specific UNG2 phosphorylations regulate protein turnover, activity and association with RPA. *EMBO J* 27:51-61.
- Halabelian, L., M. Ravichandran, Y. Li, H. Zeng, A. Rao, L. Aravind, and C.H. Arrowsmith. 2019. Structural basis of HMCES interactions with abasic DNA and multivalent substrate recognition. *Nature Structural & Molecular Biology* 26:607-612.
- Halverson, R., R.M. Torres, and R. Pelanda. 2004. Receptor editing is the main mechanism of B cell tolerance toward membrane antigens. *Nature Immunology* 5:645-650.
- Han, L., S. Masani, and K. Yu. 2010. Cutting Edge: CTNBL1 Is Dispensable for Ig Class Switch Recombination. 185:1379-1381.
- Hansen, R.S., C. Wijmenga, P. Luo, A.M. Stanek, T.K. Canfield, C.M.R. Weemaes, and S.M. Gartler. 1999. The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. 96:14412-14417.
- Hanson, E.P., L. Monaco-Shawver, L.A. Solt, L.A. Madge, P.P. Banerjee, M.J. May, and J.S. Orange. 2008. Hypomorphic nuclear factor- κ B essential modulator mutation database and reconstitution system identifies phenotypic and immunologic diversity. 122:1169-1177.e1116.
- 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.
- 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, B., W. Xu, P.A. Santini, A.D. Polydorides, A. Chiu, J. Estrella, M. Shan, A. Chadburn, V. Villanacci, A. Plebani, D.M. Knowles, M. Rescigno, and A. Cerutti. 2007. Intestinal Bacteria Trigger T Cell-Independent Immunoglobulin A2 Class Switching by Inducing Epithelial-Cell Secretion of the Cytokine APRIL. *Immunity* 26:812-826.
- He, Y., J. Ren, X. Xu, K. Ni, A. Schwader, R. Finney, C. Wang, L. Sun, K. Klarmann, J. Keller, A. Tubbs, A. Nussenzweig, and K. Muegge. 2020. Lsh/HELLS is required for B lymphocyte development and immunoglobulin class switch recombination. *Proc Natl Acad Sci U S A* 117:20100-20108.
- Heese, K. 2013. The protein p17 signaling pathways in cancer. *Tumour Biol* 34:4081-4087.
- 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.

- Hogerkor, C.M., and C.A.K. Borrebaeck. 2006. The Human CD77- B Cell Population Represents a Heterogeneous Subset of Cells Comprising Centroblasts, Centrocytes, and Plasmablasts, Prompting Phenotypical Revision. *177:4341-4349*.
- 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.
- Hsiao, K.Y., and C.A. Mizzen. 2013. Histone H4 deacetylation facilitates 53BP1 DNA damage signaling and double-strand break repair. *Journal of Molecular Cell Biology* 5:157-165.
- 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. *Proceedings of the National Academy of Sciences* 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.
- Hwang, J.K., F.W. Alt, and L.-S. Yeap. 2015. Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination. In *American Society of Microbiology*, 325-348.
- 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., N. Catalan, A. Plebani, L. Maródi, Ö. Sanal, S. Kumaki, V. Nagendran, P. Wood, C. Glastre, F. Sarrot-Reynauld, O. Hermine, M. Forveille, P. Revy, A. Fischer, and A. Durandy. 2003a. Hyper-IgM syndrome type 4 with a B lymphocyte–intrinsic selective deficiency in Ig class-switch recombination. *Journal of Clinical Investigation* 112:136-142.
- 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. *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. 2003c. Human uracil-DNA glycosylase deficiency associated with profoundly impaired immunoglobulin class-switch recombination. *Nat Immunol* 4:1023-1028.
- Imai, K., Y. Zhu, P. Revy, T. Morio, S. Mizutani, A. Fischer, S. Nonoyama, and A. Durandy. 2005. Analysis of class switch recombination and somatic hypermutation in patients affected with autosomal dominant hyper-IgM syndrome type 2. *115:277-285*.
- 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.
- Issaoui, H., N. Ghazzaui, A. Saintamand, C. Carrion, C. Oblet, and Y. Denizot. 2018. The IgH 3' regulatory region super-enhancer does not control IgA class switch recombination in the B1 lineage. *Cell Mol Immunol* 15:289-291.
- Ito, S., H. Nagaoka, R. Shinkura, N. Begum, M. Muramatsu, M. Nakata, and T. Honjo. 2004. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proceedings of the National Academy of Sciences* 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.
- Jain, A., C.A. Ma, E. Lopez-Granados, G. Means, W. Brady, J.S. Orange, S. Liu, S. Holland, and J.M.J. Derry. 2004. Specific NEMO mutations impair CD40-mediated c-Rel activation and B cell terminal differentiation. *Journal of Clinical Investigation* 114:1593-1602.
- 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. *The Journal of Experimental Medicine* 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.
- Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, and E. Charpentier. 2012. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337:816-821.
- Joung, J., S. Konermann, J.S. Gootenberg, O.O. Abudayyeh, R.J. Platt, M.D. Brigham, N.E. Sanjana, and F. Zhang. 2016. Protocol: Genome-Scale CRISPR-Cas9 Knockout and Transcriptional Activation Screening.
- Joung, J., S. Konermann, J.S. Gootenberg, O.O. Abudayyeh, R.J. Platt, M.D. Brigham, N.E. Sanjana, and F. Zhang. 2017. Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat Protoc* 12:828-863.
- Jung, S., K. Rajewsky, and A. Radbruch. 1993. Shutdown of class switch recombination by deletion of a switch region control element. 259:984-987.
- Kaminski, D.A., and J. Stavnezer. 2006. Enhanced IgA Class Switching in Marginal Zone and B1 B Cells Relative to Follicular/B2 B Cells. *The Journal of Immunology* 177:6025-6029.
- 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.
- Kenter, A.L., S. Feldman, R. Wuerffel, I. Achour, L. Wang, and S. Kumar. 2012. Three-dimensional architecture of the IgH locus facilitates class switch recombination. 1267:86-94.

- Khamlichi, A.A., F. Glaudet, 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.
- Kim, A., L. Han, G.E. Santiago, R.E. Verdun, and K. Yu. 2016. Class-Switch Recombination in the Absence of the IgH 3' Regulatory Region. *J Immunol* 197:2930-2935.
- 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.
- Kleinstiver, B.P., V. Pattanayak, M.S. Prew, S.Q. Tsai, N.T. Nguyen, Z. Zheng, and J.K. Joung. 2016. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529:490-495.
- 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.
- 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 γ -H2AX Foci through Interaction with the FHA/BRCT Domain. *Current Biology* 12:1846-1851.
- 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.
- 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.
- Korthäuer, U., D. Graf, H.W. Mages, F. Brière, M. Padayachee, S. Malcolm, A.G. Ugazio, L.D. Notarangelo, R.J. Levinsky, and R.A. Kroczeck. 1993. Defective expression of T-cell CD40 ligand causes X-linked immunodeficiency with hyper-IgM. 361:539-541.
- 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. Schwartzenruber, 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.
- 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.
- Kuang, F.L., Z. Luo, and M.D. Scharff. 2009. H3 trimethyl K9 and H3 acetyl K9 chromatin modifications are associated with class switch recombination. 106:5288-5293.
- 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.
- Küppers, R., U. Klein, M.-L. Hansmann, and K. Rajewsky. 1999. Cellular Origin of Human B-Cell Lymphomas. *The New England Journal of Medicine* 341:1520-1529.
- Kuscu, C., S. Arslan, R. Singh, J. Thorpe, and M. Adli. 2014. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat Biotechnol* 32:677-683.
- 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.
- 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.
- 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.
- Lander, E.S. 2016. The Heroes of CRISPR. *Cell* 164:18-28.
- Landree, M.A., J.A. Wibbenmeyer, and D.B. Roth. 1999. Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. *Genes & Development* 13:3059-3069.
- 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.
- 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.
- 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.
- Lee, C.M., T.H. Davis, and G. Bao. 2018. Examination of CRISPR/Cas9 design tools and the effect of target site accessibility on Cas9 activity. *Exp Physiol* 103:456-460.
- Lee, S.U., and T. Maeda. 2012. POK/ZBTB proteins: an emerging family of proteins that regulate lymphoid development and function. *Immunol Rev* 247:107-119.
- 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.
- Li, C., T. Irrazabal, C.C. So, M. Berru, L. Du, E. Lam, A.K. Ling, J.L. Gommerman, Q. Pan-Hammarstrom, and A. Martin. 2018. The H2B deubiquitinase Usp22 promotes antibody class switch recombination by facilitating non-homologous end joining. *Nat Commun* 9:1006.
- Li, F., Y. Yan, J. Pieretti, D.A. Feldman, and L.A. Eckhardt. 2010. Comparison of Identical and Functional Igh Alleles Reveals a Nonessential Role for E μ in Somatic Hypermutation and Class-Switch Recombination. *The Journal of Immunology* 185:6049-6057.
- Li, G., F.W. Alt, H.-L. Cheng, J.W. Brush, P.H. Goff, M.M. Murphy, S. Franco, Y. Zhang, and S. Zha. 2008. Lymphocyte-Specific Compensation for XLF/Cernunnos End-Joining Functions in V(D)J Recombination. 31:631-640.

- Li, G., Clayton, T. Lam, Egest, Daniel, Ken, H. Zan, Z. Xu, and P. Casali. 2013. Combinatorial H3K9acS10ph Histone Modification in IgH Locus S Regions Targets 14-3-3 Adaptors and AID to Specify Antibody Class-Switch DNA Recombination. 5:702-714.
- Li, W., H. Xu, T. Xiao, L. Cong, M.I. Love, F. Zhang, R.A. Irizarry, J.S. Liu, M. Brown, and X.S. Liu. 2014. MAGECK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* 15:554.
- 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, 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, X., T. Liu, Y. Shang, P. Dai, W. Zhang, B.J. Lee, M. Huang, D. Yang, Q. Wu, L.D. Liu, X. Zheng, B.O. Zhou, J. Dong, L.S. Yeap, J. Hu, T. Xiao, S. Zha, R. Casellas, X.S. Liu, and F.L. Meng. 2020. ERCC6L2 promotes DNA orientation-specific recombination in mammalian cells. *Cell Res* 30:732-744.
- Liu, Y.J., D.E. Joshua, G.T. Williams, C.A. Smith, J. Gordon, and I.C.M. MacLennan. 1989. Mechanism of antigen-driven selection in germinal centres. 342:929-931.
- 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.
- Loeffler, M., M. Kreuz, A. Haake, D. Hasenclever, H. Trautmann, C. Arnold, K. Winter, K. Koch, W. Klapper, R. Scholtysik, M. Rosolowski, S. Hoffmann, O. Ammerpohl, M. Szczepanowski, D. Herrmann, R. Küppers, C. Pott, and R. Siebert. 2015. Genomic and epigenomic co-evolution in follicular lymphomas. *Leukemia* 29:456-463.
- 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.
- Luby, T.M., C.E. Schrader, J. Stavnezer, and E. Selsing. 2001. The μ Switch Region Tandem Repeats Are Important, but Not Required, for Antibody Class Switch Recombination. 193:159-168.
- 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. Wayne. 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.
- Macara, I.G. 2001. Transport into and out of the nucleus. *Microbiol Mol Biol Rev* 65:570-594, table of contents.
- Macpherson, A.J., K.D. McCoy, F.E. Johansen, and P. Brandtzaeg. 2008. The immune geography of IgA induction and function. *Mucosal Immunology* 1:11-22.

- 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*.
- Mali, P., L. Yang, K.M. Esvelt, J. Aach, M. Guell, J.E. Dicarlo, J.E. Norville, and G.M. Church. 2013. RNA-Guided Human Genome Engineering via Cas9. *Science* 339:823-826.
- Manis, J.P., Y. Gu, R. Lansford, E. Sonoda, R. Ferrini, L. Davidson, K. Rajewsky, and F.W. Alt. 1998a. 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. 2004. 53BP1 links DNA damage-response pathways to immunoglobulin heavy chain class-switch recombination. *5:481-487*.
- Manis, J.P., N. Van Der Stoep, M. Tian, R. Ferrini, L. Davidson, A. Bottaro, and F.W. Alt. 1998b. Class Switching in B Cells Lacking 3' Immunoglobulin Heavy Chain Enhancers. *188:1421-1431*.
- Marquet, M., A. Garot, S. Bender, C. Carrion, P. Rouaud, S. Lecardeur, Y. Denizot, M. Cogné, and E. Pinaud. 2014. The E μ Enhancer Region Influences H Chain Expression and B Cell Fate without Impacting IgVH Repertoire and Immune Response In Vivo. *The Journal of Immunology* 193:1171-1183.
- 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*.
- Marusawa, H. 2008. Aberrant AID expression and human cancer development. *The International Journal of Biochemistry & Cell Biology* 40:1399-1402.
- 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, T., T. Shimizu, N. Nishijima, A. Ikeda, Y. Eso, Y. Matsumoto, T. Chiba, and H. Marusawa. 2015. Hepatic inflammation facilitates transcription-associated mutagenesis via AID activity and enhances liver tumorigenesis. *36:904-913*.
- 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. Zheng, L.J. DiMenna, and J. Chaudhuri. 2014. Regulation of immunoglobulin class-switch recombination: choreography of noncoding transcription, targeted DNA deamination, and long-range DNA repair. *Adv Immunol* 122:1-57.
- 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.
- McBlane, J.F., D.C. van Gent, D.A. Ramsden, C. Romeo, C.A. Cuomo, M. Gellert, and M.A. Oettinger. 1995. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell* 83:387-395.
- 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.
- Medvedovic, J., A. Ebert, H. Tagoh, I.M. Tamir, T.A. Schwickert, M. Novatchkova, Q. Sun, P.J. Huis In 't Veld, C. Guo, H.S. Yoon, Y. Denizot, S.J. Holwerda, W. de Laat, M. Cogne, Y. Shi, F.W. Alt, and M. Busslinger. 2013. Flexible long-range loops in the VH gene region of the Igh locus facilitate the generation of a diverse antibody repertoire. *Immunity* 39:229-244.
- Meffre, E., and H. Wardemann. 2008. B-cell tolerance checkpoints in health and autoimmunity. *Current Opinion in Immunology* 20:632-638.
- Meier, J.A., F. Zhang, and N.E. Sanjana. 2017. GUIDES: sgRNA design for loss-of-function screens. *Nat Methods* 14:831-832.
- 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.
- Methot, S.P., and J.M. Di Noia. 2017. Molecular Mechanisms of Somatic Hypermutation and Class Switch Recombination. In Elsevier, 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. Buschiazzo, 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.
- Michlits, G., J. Jude, M. Hinterdorfer, M. de Almeida, G. Vainorius, M. Hubmann, T. Neumann, A. Schleiffer, T.R. Burkard, M. Fellner, M. Gijbsbertsen, A. Traunbauer, J. Zuber, and U. Elling. 2020. Multilayered VBC score predicts sgRNAs that efficiently generate loss-of-function alleles. *Nat Methods* 17:708-716.
- 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.
- Mombaerts, P., J. Iacomini, R.S. Johnson, K. Herrup, S. Tonegawa, and V.E. Papaioannou. 1992. RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* 68:869-877.
- 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.
- Mu, Y., M.A. Zelazowska, and K.M. McBride. 2017. Phosphorylation promotes activation-induced cytidine deaminase activity at the Myc oncogene. *J Exp Med* 214:3543-3552.
- 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. 1999. 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.
- Nakagawa, H., J.C. Lockman, W.L. Frankel, H. Hampel, K. Steenblock, L.J. Burgart, S.N. Thibodeau, and a.A.d.I. Chappelle. 2004. Mismatch Repair Gene PMS2: Disease-Causing Germline Mutations Are Frequent in Patients Whose Tumors Stain Negative for PMS2 Protein, but Paralogous Genes Obscure Mutation Detection and Interpretation. 64:4721-4727.
- 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. *International Immunology* 8:193-201.
- Nambu, Y., M. Sugai, H. Gonda, C.-G. Lee, T. Kataikai, Y. Agata, Y. Yokota, and A. Shimizu. 2003. Transcription-Coupled Events Associating with Immunoglobulin Switch Region Chromatin. *Science* 302:2137-2140.
- 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.
- Niessen, R.C., M.J.W. Berends, Y. Wu, R.H. Sijmons, H. Hollema, M.J.L. Ligtenberg, H.E.K. De Walle, E.G.E. De Vries, A. Karrenbeld, C.H.C.M. Buys, A.G.J. Van Der Zee, R.M.W. Hofstra, and J.H. Kleibeuker. 2006. Identification of mismatch repair gene mutations in young patients with colorectal cancer and in patients with multiple tumours associated with hereditary non-polyposis colorectal cancer. 55:1781-1788.
- Nieuwenhuis, P., and D. Opstelten. 1984. Functional anatomy of germinal centers. *THE AMERICAN JOURNAL OF ANATOMY* 170:421-435.
- 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., 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.
- Nimmerjahn, F., and J.V. Ravetch. 2008. Fc γ receptors as regulators of immune responses. *Nature Reviews Immunology* 8:34-47.
- 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.
- Nitta, H., M. Unoki, K. Ichiyonagi, T. Kosho, T. Shigemura, H. Takahashi, G. Velasco, C. Francastel, C. Picard, T. Kubota, and H. Sasaki. 2013. Three novel ZBTB24 mutations identified in Japanese and Cape Verdean type 2 ICF syndrome patients. *J Hum Genet* 58:455-460.

- Nonoyama, S., D. Hollenbaugh, A. Aruffo, J.A. Ledbetter, and H.D. Ochs. 1993. B cell activation via CD40 is required for specific antibody production by antigen-stimulated human B cells. *Journal of Experimental Medicine* 178:1097-1102.
- Noordermeer, S.M., S. Adam, D. Setiapatra, 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*
- Noordzij, J.G., N.M. Wulffraat, A. Haraldsson, I. Meyts, L.J. van't Veer, F.B. Hogervorst, A. Warris, and C.M. Weemaes. 2009. Ataxia-telangiectasia patients presenting with hyper-IgM syndrome. *Arch Dis Child* 94:448-449.
- 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.
- Obata, M., T. Kataoka, S. Nakai, H. Yamagishi, N. Takahashi, Y. Yamawaki-Kataoka, T. Nikaido, A. Shimizu, and T. Honjo. 1981. Structure of a rearranged gamma 1 chain gene and its implication to immunoglobulin class-switch mechanism. 78:2437-2441.
- Ochs, H.D., D. Hollenbaugh, and A. Aruffo. 1994. The role of CD40L (gp39)/CD40 in T/B cell interaction and primary immunodeficiency. *Research in Immunology* 145:210-215.
- Oettgen, H.C. 2000. Regulation of the IgE isotype switch: new insights on cytokine signals and the functions of ϵ germline transcripts. *Curr opin in Immunology* 12:618-623.
- 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. 2003. 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.
- 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. 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., E.-K. Lee, B.-C. Kim, and P.-H. Kim. 2003. p300 cooperates with Smad3/4 and Runx3 in TGF β 1-induced IgA isotype expression. *European Journal of Immunology* 33:3386-3392.
- 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. *Nature Immunology* 10:540-550.
- 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., H. Khiabani, M. Fangazio, M. Vasishtha, M. Messina, Antony, P. Ouillette, V. Trifonov, D. Rossi, F. Tabbò, M. Ponzoni, A. Chadburn, Vundavalli, G. Bhagat, G. Gaidano, G. Inghirami, Sami, R. Rabadan, and R. Dalla-Favera. 2014. Genetics of Follicular Lymphoma Transformation. *Cell Reports* 6:130-140.
- 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. 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. Kuppers, 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. 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., 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.

- Pefanis, E., J. Wang, G. Rothschild, J. Lim, D. Kazadi, J. Sun, A. Federation, J. Chao, O. Elliott, Z.P. Liu, A.N. Economides, J.E. Bradner, R. Rabadan, and U. Basu. 2015. RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell* 161:774-789.
- Pei, H., X. Wu, T. Liu, K. Yu, D.F. Jelinek, and Z. Lou. 2013. The Histone Methyltransferase MMSET Regulates Class Switch Recombination. *The Journal of Immunology* 190:756-763.
- 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.
- Peng, J., Y. Zhou, S. Zhu, and W. Wei. 2015. High-throughput screens in mammalian cells using the CRISPR-Cas9 system. *FEBS J* 282:2089-2096.
- 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.
- Perlot, T., F.W. Alt, C.H. Bassing, H. Suh, and E. Pinaud. 2005. Elucidation of IgH intronic enhancer functions via germ-line deletion. 102:14362-14367.
- 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.
- PéRon, S., Q. Pan-Hammarström, K. Imai, L. Du, N. Taubenheim, O. Sanal, L. Marodi, A. Bergelin-Besançon, M. Benkerrou, J.-P. De Villartay, A. Fischer, P. Revy, and A. Durandy. 2007. A primary immunodeficiency characterized by defective immunoglobulin class switch recombination and impaired DNA repair. *Journal of Experimental Medicine* 204:1207-1216.
- Petersen-Mahrt, S.K., R.S. Harris, and M.S. Neuberger. 2002. 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.
- Pettersen, H.S., O. Sundheim, K.M. Gilljam, G. Slupphaug, H.E. Krokan, and B. Kavli. 2007. Uracil-DNA glycosylases SMUG1 and UNG2 coordinate the initial steps of base excision repair by distinct mechanisms. *Nucleic Acids Research* 35:3879-3892.
- 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., A.A. Khamlichi, C. Le Morvan, M. Drouet, V. Nalesso, M. Le Bert, and M. Cogné. 2001. Localization of the 3' IgH Locus Elements that Effect Long-Distance Regulation of Class Switch Recombination. 15:187-199.
- 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- κ B pathway. *Nature Communications* 3:1-12.
- 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.
- Predeus, A.V., S. Gopalakrishnan, Y. Huang, J. Tang, A.J. Feeney, E.M. Oltz, and M.N. Artyomov. 2014. Targeted Chromatin Profiling Reveals Novel Enhancers in Ig H and Ig L Chain Loci. 192:1064-1070.

- Pucella, J.N., and J. Chaudhuri. 2017. AID Invited to the G4 Summit. *Molecular Cell* 67:355-357.
- Qiao, Q., L. Wang, F.-L. Meng, J.K. Hwang, F.W. Alt, and H. Wu. 2017. AID Recognizes Structured DNA for Class Switch Recombination. *Molecular 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. 2004. 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. *Proceedings of the National Academy of Sciences* 99:7003-7008.
- Rada, C., and C. Milstein. 2001. The intrinsic hypermutability of antibody heavy and light chain genes decays exponentially. *The EMBO Journal* 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.
- Rajagopal, D., R.W. Maul, A. Ghosh, T. Chakraborty, A.A. Khamlichi, R. Sen, and P.J. Gearhart. 2009. Immunoglobulin switch μ sequence causes RNA polymerase II accumulation and reduces dA hypermutation. *The Journal of Experimental Medicine* 206:1237-1244.
- Rajewsky, K. 1996. Clonal selection and learning in the antibody system. *Nature* 381:751.
- 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.
- Ramezani-Rad, P., and R.C. Rickert. 2017. Murine models of germinal center derived-lymphomas. *Current Opinion in Immunology* 45:31-36.
- 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. *Nature Immunology* 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.
- Rappold, I., K. Iwabuchi, T. Date, and J. Chen. 2001. Tumor Suppressor P53 Binding Protein 1 (53bp1) Is Involved in DNA Damage-Signaling Pathways. 153:613-620.
- 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.
- 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*.
- Revy, P., T. Muto, Y. Levy, F. Geissmann, A. Plebani, O. Sanal, N. Catalan, M. Forveille, R. Dufourcq-Lagelouse, 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.-A., V. Anquez, A. Dahan, and J.-C. Weill. 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.
- Reynaud, C.-A., A. Dahan, V. Anquez, and J.-C. Weill. 1989. Somatic hyperconversion diversifies the single VH gene of the chicken with a high incidence in the D region. *Cell* 59:171-183.
- Ribeiro De Almeida, C., S. Dhir, A. Dhir, A.E. Moghaddam, Q. Sattentau, A. Meinhart, and N.J. Proudfoot. 2018. RNA Helicase DDX1 Converts RNA G-Quadruplex Structures into R-Loops to Promote IgH Class Switch Recombination. *Molecular Cell* 70:650-662.e658.
- Rivera-Munoz, P., P. Soulas-Sprauel, G. Le Guyader, V. Abramowski, S. Bruneau, A. Fischer, F. Paques, and J.P. De Villartay. 2009. Reduced immunoglobulin class switch recombination in the absence of Artemis. *114:3601-3609*.
- 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. *Molecular 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*.
- 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. Victora, 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.
- Rogakou, E.P., D.R. Pilch, A.H. Orr, V.S. Ivanova, and W.M. Bonner. 1998. DNA Double-stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139. *Journal of Biological Chemistry* 273:5858-5868.
- Rogozin, I.B., and N.A. Kolchanov. 1992. 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 η 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.
- Rommel, P.C., T.Y. Oliveira, M.C. Nussenzweig, and D.F. Robbiani. 2017. RAG1/2 induces genomic insertions by mobilizing DNA into RAG1/2-independent breaks. *Journal of Experimental Medicine* 214:815-831.
- Roth, D.B. 2002. Amplifying Mechanisms of Lymphomagenesis. *Molecular Cell* 10:1-2.
- 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.
- 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., C. Vincent-Fabert, M. Marquet, N. Ghazzaui, V. Magnone, E. Pinaud, M. Cogné, and Y. Denizot. 2017. E μ and 3'RR IgH enhancers show hierarchic unilateral dependence in mature B-cells. *Scientific Reports* 7:
- Sakofsky, C.J., S.A. Roberts, E. Malc, P.A. Mieczkowski, M.A. Resnick, D.A. Gordenin, and A. Malkova. 2014. Break-induced replication is a source of mutation clusters underlying kataegis. *Cell Rep* 7:1640-1648.
- Sanjana, N.E., O. Shalem, and F. Zhang. 2014. Improved vectors and genome-wide libraries for CRISPR screening. *Nat Methods* 11:783-784.
- 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. 2019. Uracil–DNA glycosylase UNG1 isoform variant supports class switch recombination and repairs nuclear genomic uracil. *Nucleic Acids Research* 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., and P.C. Swanson. 2011. V(D)J Recombination: Mechanisms of Initiation. *Annual Review of Genetics* 45:167-202.
- Scheeren, F.A., M. Nagasawa, K. Weijer, T. Cupedo, J.R. Kirberg, N. Legrand, and H. Spits. 2008. T cell-independent development and induction of somatic hypermutation in human IgM+IgD+CD27+ B cells. *Journal of Experimental Medicine* 205:2033-2042.
- 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. 2005. Inducible DNA breaks in Ig S regions are dependent on AID and UNG. *Journal of Experimental Medicine* 202:561-568.
- Schwab, K.R., S.R. Patel, and G.R. Dressler. 2011. Role of PTIP in Class Switch Recombination and Long-Range Chromatin Interactions at the Immunoglobulin Heavy Chain Locus. *Molecular and Cellular Biology* 31:1503-1511.
- Seifert, M., R. Scholtysik, and R. Küppers. 2013. Origin and Pathogenesis of B Cell Lymphomas. In Humana Press, 1-25.
- Sekiguchi, J., D.O. Ferguson, H.T. Chen, E.M. Yang, J. Earle, K. Frank, S. Whitlow, Y. Gu, Y. Xu, A. Nussenzweig, and F.W. Alt. 2001. Genetic interactions between ATM and the nonhomologous end-joining factors in genomic stability and development. 98:3243-3248.
- 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.
- Setiaputra, D., and D. Durocher. 2019. Shieldin – the protector of DNA ends. *EMBO reports* 20:e47560.
- Shalem, O., N.E. Sanjana, E. Hartenian, X. Shi, D.A. Scott, T. Mikkelsen, D. Heckl, B.L. Ebert, D.E. Root, J.G. Doench, and F. Zhang. 2014. Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science* 343:84-87.
- Shalem, O., N.E. Sanjana, and F. Zhang. 2015. High-throughput functional genomics using CRISPR-Cas9. *Nat Rev Genet* 16:299-311.
- 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.
- Shinkai, Y., G. Rathbun, K.P. Lam, E.M. Oltz, V. Stewart, M. Mendelsohn, J. Charron, M. Datta, F. Young, A.M. Stall, and e. al. 1992. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* 68:855-867.

- 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.
- Simsek, D., and M. Jasin. 2010. Alternative end-joining is suppressed by the canonical NHEJ component Xrcc4–ligase IV during chromosomal translocation formation. *Nature Structural & Molecular Biology* 17:410-416.
- 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.
- Smith, I., P.G. Greenside, T. Natoli, D.L. Lahr, D. Wadden, I. Tirosh, R. Narayan, D.E. Root, T.R. Golub, A. Subramanian, and J.G. Doench. 2017. Evaluation of RNAi and CRISPR technologies by large-scale gene expression profiling in the Connectivity Map. *PLoS Biol* 15:e2003213.
- Sohail, A.K., Joanna; Samaranayake, Mala; Ullah, Asad ; S. Bhagwat, Ashok. 2003. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. *Nucleic Acids Research* 31:2990-2994.
- 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. 107:22190-22195.
- Stavnezer, J., and C.E. Schrader. 2014. Ig heavy chain class switch recombination: mechanism and regulation. *Journal of immunology (Baltimore, Md. : 1950)* 193:5370-5378.
- 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.
- 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. 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.
- 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. Müschen. 2015. Mechanisms of clonal evolution in childhood acute lymphoblastic leukemia. *Nature Immunology* 16:766-774.
- 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. 2003. AID mutant analyses indicate requirement for class-switch-specific cofactors. *Immunity* 18:843-848.
- 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. Wening, 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.
- Thijssen, P.E., Y. Ito, G. Grillo, J. Wang, G. Velasco, H. Nitta, M. Unoki, M. Yoshihara, M. Suyama, Y. Sun, R.J. Lemmers, J.C. de Greef, A. Gennery, P. Picco, B. Kloeckener-Gruissem, T. Gungor, I. Reisli, C. Picard, K. Kebaili, B. Roquelaure, T. Iwai, I. Kondo, T. Kubota, M.M. van Ostaïjen-Ten Dam, M.J. van Tol, C. Weemaes, C. Francastel, S.M. van der Maarel, and H. Sasaki. 2015. Mutations in CDCA7 and HELLS cause immunodeficiency-centromeric instability-facial anomalies syndrome. *Nat Commun* 6:7870.
- 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. *Journal of Experimental Medicine* 210:2495-2502.
- Tian, M., and F.W. Alt. 2000. Transcription-induced Cleavage of Immunoglobulin Switch Regions by Nucleotide Excision Repair Nucleases in Vitro. *Journal of Biological Chemistry* 275:24163-24172.
- 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.
- 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.
- 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.

- Unoki, M., H. Funabiki, G. Velasco, C. Francastel, and H. Sasaki. 2019. CDCA7 and HELLS mutations undermine nonhomologous end joining in centromeric instability syndrome. *J Clin Invest* 129:78-92.
- Vaidyanathan, B., W.-F. Yen, J.N. Pucella, and J. Chaudhuri. 2014. AIDing Chromatin and Transcription-Coupled Orchestration of Immunoglobulin Class-Switch Recombination. 5:
- 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.
- Van Zelm, M.C., S.J.W. Bartol, G.J. Driessen, F. Mascart, I. Reisli, J.L. Franco, B. Wolska-Kusnierz, H. Kanegane, L. Boon, J.J.M. Van Dongen, and M. Van Der Burg. 2014. Human CD19 and CD40L deficiencies impair antibody selection and differentially affect somatic hypermutation. 134:135-144.e137.
- Veeneman, B., Y. Gao, J. Grant, D. Fruhling, J. Ahn, B. Bosbach, J. Bienkowska, M. Follettie, K. Arndt, J. Myers, and W. Zhong. 2020. PINCER: improved CRISPR/Cas9 screening by efficient cleavage at conserved residues. *Nucleic Acids Res* 48:9462-9477.
- Victora, G.D., and M.C. Nussenzweig. 2012. Germinal Centers. *Annual Review of Immunology* 30:429-457.
- Victora, 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. Cogné, M. Cogné, 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.
- Volpi, S.A., J. Verma-Gaur, R. Hassan, Z. Ju, S. Roa, S. Chatterjee, U. Werling, H. Hou, B. Will, U. Steidl, M. Scharff, W. Edelman, A.J. Feeney, and B.K. Birshstein. 2012. Germline Deletion of Igh 3' Regulatory Region Elements hs 5, 6, 7 (hs5-7) Affects B Cell-Specific Regulation, Rearrangement, and Insulation of the Igh Locus. 188:2556-2566.
- 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. *Nat Immunol* 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. *Nat Immunol* 10:420-426.
- Wang, J.H., F.W. Alt, M. Gostissa, A. Datta, M. Murphy, M.B. Alimzhanov, K.M. Coakley, K. Rajewsky, J.P. Manis, and C.T. Yan. 2008. Oncogenic transformation in the absence of Xrcc4 targets peripheral B cells that have undergone editing and switching. 205:3079-3090.
- Wang, L., N. Whang, R. Wuerffel, and A.L. Kenter. 2006a. AID-dependent histone acetylation is detected in immunoglobulin S regions. 203:215-226.
- 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. 206:1817-1830.
- Wang, M., W. Wu, W. Wu, B. Rosidi, L. Zhang, H. Wang, and G. Iliakis. 2006b. 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, T., J.J. Wei, D.M. Sabatini, and E.S. Lander. 2014a. Genetic screens in human cells using the CRISPR-Cas9 system. *Science* 343:80-84.

- Wang, X., M. Fan, S. Kalis, L. Wei, and M.D. Scharff. 2014b. A source of the single-stranded DNA substrate for activation-induced deaminase during somatic hypermutation. *Nat Commun* 5:4137.
- Ward, I.M., K. Minn, K.G. Jorda, and J. Chen. 2003. Accumulation of Checkpoint Protein 53BP1 at DNA Breaks Involves Its Binding to Phosphorylated Histone H2AX. *Journal of Biological Chemistry* 278:19579-19582.
- Ward, I.M., B. Reina-San-Martin, A. Oлару, 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.
- 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.
- White, D.E., D. Negorev, H. Peng, A.V. Ivanov, G.G. Maul, and F.J. Rauscher. 2006. KAP1, a Novel Substrate for PIKK Family Members, Colocalizes with Numerous Damage Response Factors at DNA Lesions. *Cancer Research* 66:11594-11599.
- 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. 209:2099-2111.
- Woof, J.M., and M.A. Kerr. 2006. The function of immunoglobulin A in immunity. *The Journal of Pathology* 208:270-282.
- Wu, L.C., and A.A. Zarrin. 2014. The production and regulation of IgE by the immune system. *Nature Reviews Immunology* 14:247-259.
- Wu, X., D.A. Scott, A.J. Kriz, A.C. Chiu, P.D. Hsu, D.B. Dadon, A.W. Cheng, A.E. Trevino, S. Konermann, S. Chen, R. Jaenisch, F. Zhang, and P.A. Sharp. 2014. Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat Biotechnol* 32:670-676.
- Wuerffel, R., L. Wang, F. Grigera, J. Manis, E. Selsing, T. Perlot, F.W. Alt, M. Cogne, E. Pinaud, and A.L. Kenter. 2007. S-S Synapsis during Class Switch Recombination Is Promoted by Distantly Located Transcriptional Elements and Activation-Induced Deaminase. 27:711-722.
- 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, G.L., T.H. Bestor, D. Bourc'his, C.L. Hsieh, N. Tommerup, M. Bugge, M. Hulten, X. Qu, J.J. Russo, and E. Viegas-Pequignot. 1999. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* 402:187-191.
- 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. *Nature Structural & Molecular Biology* 17:1124-1135.
- 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*^{-/-} *ung*^{-/-} mice. *J Exp Med* 203:2085-2094.
- Yadav, A., A. Oлару, 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.
- Yamane, A., W. Resch, N. Kuo, S. Kuchen, Z. Li, H.-W. Sun, D.F. Robbiani, K. McBride, M.C. Nussenzweig, and R. Casellas. 2011. Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. 12:62-69.
- 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. 449:478-482.

- Yeap, L.S., J.K. Hwang, Z. Du, R.M. Meyers, F.L. Meng, A. Jakubauskaite, M. Liu, V. Mani, D. Neuberg, 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. 111:E1016-E1024.
- 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. *Nature Immunology* 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. *Journal of Biological Chemistry* 279:6496-6500.
- Yu, K., D. Roy, M. Bayramyan, I.S. Haworth, and M.R. Lieber. 2005. Fine-Structure Analysis of Activation-Induced Deaminase Accessibility to Class Switch Region R-Loops. 25:1730-1736.
- Yun, M.H., and K. Hiom. 2009. CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. *Nature* 459:460-463.
- 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.
- Zarrin, A.A., F.W. Alt, J. Chaudhuri, N. Stokes, D. Kaushal, L. Du Pasquier, and M. Tian. 2004. An evolutionarily conserved target motif for immunoglobulin class-switch recombination. *Nature Immunology* 5:1275-1281.
- Zeng, X., D.B. Winter, C. Kasmer, K.H. Kraemer, A.R. Lehmann, and P.J. Gearhart. 2001. DNA polymerase η 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. 1993. AselectivedefectingG2bswitchingasaresultoftargetedmutationofthely2bpromoterandexon. *TheEMBOJournal* 12:
- Zhang, Z.Z., N.R. Pannunzio, L. Han, C.L. Hsieh, K. Yu, and M.R. Lieber. 2014. The strength of an Ig switch region is determined by its ability to drive R loop formation and its number of WGCW sites. *Cell Rep* 8:557-569.
- Zheng, S., Bao, B. Vaidyanathan, J.-Y. Lin, F.-T. Huang, and J. Chaudhuri. 2015. 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.
- Zimmermann, M., F. Lotterberger, S.B. Buonomo, A. Sfeir, and T. De Lange. 2013. 53BP1 Regulates DSB Repair Using Rif1 to Control 5' End Resection. *Science* 339:700-704.

- Ziv, Y., D. Bielopolski, Y. Galanty, C. Lukas, Y. Taya, D.C. Schultz, J. Lukas, S. Bekker-Jensen, J. Bartek, and Y. Shiloh. 2006. Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. *Nature Cell Biology* 8:870-876.
- 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 η 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.