



**HAL**  
open science

# Rôle des éléments cis-régulateurs du locus des chaînes lourdes d'immunoglobuline sur l'organisation du noyau et l'intégrité du génome du lymphocyte B

Charlotte Bruzeau

► **To cite this version:**

Charlotte Bruzeau. Rôle des éléments cis-régulateurs du locus des chaînes lourdes d'immunoglobuline sur l'organisation du noyau et l'intégrité du génome du lymphocyte B. Médecine humaine et pathologie. Université de Limoges, 2023. Français. NNT : 2023LIMO0039 . tel-04275671

**HAL Id: tel-04275671**

**<https://theses.hal.science/tel-04275671v1>**

Submitted on 8 Nov 2023

**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 Limoges

### ED 652 - Biologie, Chimie, Santé (BCS)

Laboratoire CRIBL : Contrôle de la Réponse Immune B et Lymphoproliférations UMR CNRS 7276  
INSERM U1262 – B NATION : B cell Nuclear Architecture, Ig genes and Oncogenes

Thèse pour obtenir le grade de

Docteur de l'Université de Limoges

Spécialité : Immunologie, Oncologie, Inflammation et Infectiologie

Présentée et soutenue par

**Charlotte Bruzeau**

Le 19 septembre 2023

## **Rôle des éléments *cis*-régulateurs du locus des chaînes lourdes d'immunoglobuline dans l'organisation du noyau et l'intégrité du génome des cellules B**

Thèse dirigée par le Dr. Sandrine Le Noir et le Dr. Eric Pinaud

JURY :

Président du jury

Dr. Salvatore Spicuglia, DR INSERM, TAGC INSERM U1090, Marseille

Rapporteurs

Dr. Ludovic Deriano, DR Institut Pasteur, INSERM 1223, Paris

Dr. Salvatore Spicuglia, DR INSERM, TAGC INSERM U1090, Marseille

Examineurs

Dr. Stéphanie Durand, MCU, CAPTuR, Université de Limoges

Dr. Philippe Ruminy, IR-HDR Centre Henri Becquerel, Rouen

Invités

Dr. Sandrine Le Noir, CR CNRS, CRIBL, Université de Limoges

Dr. Eric Pinaud, DR CNRS, CRIBL, Université de Limoges





Je remercie la fondation ARC pour la recherche sur le cancer pour le financement de ma quatrième année de thèse.





*À mes parents,*

*À Antoine,*

*Et à ces personnes qui se sont envolées trop tôt...*

*« La vie ce n'est pas attendre que l'orage passe, c'est apprendre à danser sous la pluie. »*  
**Sénèque**

*« Je ne perds jamais. Soit je gagne, soit j'apprends. »*  
**Nelson Mandela**

*« Il y a des jours, des mois, des années interminables où il ne se passe presque rien. Il y a des minutes et des secondes qui contiennent tout un monde. »*  
**Jean d'Ormesson**

## Remerciements

---

Ça y est, l'écriture des remerciements... Synonyme de la fin, c'est un mélange curieux de sentiments : reconnaissance, soulagement, satisfaction mais aussi nostalgie... Se remémorer ces années et les liens créés au fil du temps, se souvenir aussi de la personne qu'on était avant de commencer et constater la différence avec la personne qu'on est devenue à la fin de la thèse, sacré chemin parcouru à tous les niveaux ! Cette petite partie du manuscrit peut sembler insignifiante et pourtant tellement compliquée à écrire... Sachez que les prochains mots n'expriment pas la moitié de la reconnaissance que j'ai envers vous.

Je vais bien sûr commencer par mes directeurs de thèse, **Sandrine** et **Éric**, merci de m'avoir accordé votre confiance pour faire cette thèse et de m'avoir accompagné dans ce parcours pendant quatre ans. Merci de m'avoir grandement initié à la science. **Éric**, merci pour tes phrases philosophiques du style « Choisir, c'est renoncer » ou encore « Il est urgent d'attendre ». Merci pour nos discussions scientifiques et tes conseils. Promis, je vais essayer de toujours me mettre dans mon tube 😊. **Sandrine**, j'admire ta capacité à te balader si facilement sur l'ADN, merci pour toutes ces fois où tu m'as aidé à me repérer sur ce fameux locus IgH et pour ces galères d'analyses bio-info, même quand tu étais à la maternité : « Allô, j'ai un problème avec l'analyse de 3C, les coordonnées ne correspondent pas au primer ! ». Je n'oublierai pas non plus tout ce temps passé à chercher des jonctions freqcut, uncut, mapqual et j'en passe dans les tlx filtrés et les tlx non filtrés... bonne continuation à la team **B-NATION**, en espérant que ces futures années soient riches en publications 🙏 !

Merci aux Drs. **Ludovic Deriano** & **Salvatore Spicuglia** d'avoir accepté de prendre de leur temps afin d'évaluer mon travail de thèse ainsi qu'aux Drs. **Stéphanie Durand** & **Philippe Rummy** d'être membres de ce jury.

Bien évidemment un énorme merci à l'ensemble de **CRIJBL**, un labo où je me suis épanouie pendant ces années de thèse. Je souhaite spécialement remercier **Christophe**, car sans ta recommandation auprès de Sandrine pour mon stage de M2 je n'aurais certainement jamais atterri au labo 😊.

Passons à ces piliers du labo, sans qui **CRIJBL** aurait une toute autre tête... **Tata Carrion**, merci d'avoir ri à toutes (ou presque ?!) ces blagues (plus ou moins drôles c'est vrai), et plus spécialement, merci d'avoir volé à mon secours lorsque le trieur était capricieux, et désolée pour tous ces accidents de travail, mais t'as vu je me suis calmée 😊 Et surtout, puisse la ♥bio mol♥ gagner ton cœur ! **Tatie Steph**, merci pour tes ADIIIIIIII ou tes multiples « BONJOUR ! » (à lire avec l'intonation caractéristique qui va bien) au quotidien. Merci d'avoir allumé la lumière après 19h et d'avoir marché jusqu'au fond du couloir, tu es une des seules à l'avoir fait de si nombreuses fois (et ce, même en étant éclopée ! Je t'avais dit que je l'écrirais non ?! 😊), car de toute façon à partir d'une certaine heure, comme tu sais si bien le dire : **il ne reste que les meilleurs** ! Merci aussi pour cette paire de collants (qui m'a bien sauvé quand le mien a fondu 😊) et d'en avoir toujours une de secours pour moi dans ton bureau 😊 ! Une pensée aussi pour ma petite **Cendrine**, ma formatrice **ELISA** préférée, merci pour ces discussions et les apéros qui vont avec. **Christelle O**, la N°1 de la cellule comm', un merci tout particulier pour cette fois où tu m'as aidé à ramasser mes tubes alors que je venais de renverser mon bac à glace et que je craquais complet et merci aussi pour tes talents d'illustratrice 😊 !

**Tiffany**, merci d'avoir supporté nos chansons, nos (mini ?) coups de gueule et nos conversations pendant un petit moment dans ton bureau. Merci aussi à **Amélie** pour sa bonne humeur pendant les pauses repas et puis bon courage à **Ophélie M** pour la suite du projet DLBCL (tu l'auras compris, il me tient ~~un~~ **peu** beaucoup à cœur). Merci à **Justine** pour les analyses bio-info, que j'ai demandé de refaire (sûrement un peu trop) souvent, et à **Cathy** pour l'allumage du trieur ces (trop ?) nombreuses fois. Une pensée pour **Laurent** et **Jean-Claude** qui se sont bien souvent arrêtés dans notre bureau pour discuter science et un petit peu vacances, notamment randonnée et Pyrénées. Merci aussi à **Patrick**, pour m'avoir proposé « l'écran géant » pour la période de rédac, ses conseils avertis sur le choix du post-doc et pour les « Salut Charlotte » matinaux au quotidien. **Eden**, tu as été une collègue de bureau super sympa et surtout TRÈS PATIENTE, bonne continuation à toi et merci pour ta grande gentillesse et ta touche d'humour toujours à point. Merci **Jeanne**, pour tes corrections en anglais évidemment mais surtout pour ta douceur et ta gentillesse ainsi que ton bonjour matinal chaque jour, qui m'ont bien manqué cette dernière année, profite de ta retraite comme il se doit. Merci également à **Sophie**, **Christelle VF**, **Danielle**, **Chantal**, **Nathalie** et **Jean**. Bon courage à tous les thésards, tout spécialement à ceux de la **B-NATION** team, **Pierre & Alex**, ainsi qu'à toi **Kenza**, mon petit poulain de M1 🥺, et bonne continuation aux docteurs de ma promo (pensées spéciales pour **Léa**) et aux suivants 😊.

Merci également à ces « personnes de l'ombre », **Julien** et **Christophe**, pour m'avoir permis de toujours travailler dans un environnement propre et agréable. J'adresse mes remerciements aux personnels de l'animalerie, tout particulièrement **Magali** pour sa gentillesse et sa sollicitude ainsi qu'**Ophélie**, **Elodie** et **Agnès**, qui ont grandement prit soin de mes petites souris pendant ces quatre années.

J'ai aussi à cœur de remercier des collègues un peu spéciaux, devenus des amis chers, que la thèse a mis sur mon chemin, chacun d'entre eux ayant contribué à rendre ce parcours de thèse un peu plus fun et jovial.

En premier, **Seb**, merci pour toutes ces discussions, qu'elles fussent scientifiques ou non, elles ont toujours été enrichissantes. Merci d'être passé dans le bureau, pour prendre des nouvelles, demander où j'en étais et surtout pour parler Science et planifier des manip souris en période de rédac, ça m'a permis de retrouver mes chères pipettes 😊 ! Je te souhaite d'avoir toujours des supers thésards à tes côtés à qui transmettre tes savoirs et partager ton expérience comme tu as pu le faire avec moi et tant d'autres avant 😊. J'espère que tu vas continuer le fractionné pour tout déchirer aux GVT 😊, en espérant qu'on s'y revoit un de ces jours !

**Audrey**, aka **Dédé**, la fille la plus rayonnante que je connaisse ! Tu as toujours le sourire et cette facilité à voir le positif dans chaque situation, même la plus pourrie ! Tu es partie du labo il y a maintenant un bon moment mais malgré ça, je suis reconnaissante d'avoir partagé ce petit bout de route « doctoresque » avec toi et d'avoir eu ta présence pour les moments suivants. Je garde en mémoire tes petites impros chansons, parfaites illustrations de ta joie de vivre. Merci pour ces midis sushis, ces petits verres en terrasse et merci d'être la personne que tu es, je ne te souhaite que ce que tu mérites : le bonheur. Au plaisir de te recevoir avec Virgile dans ma prochaine demeure 🍷.

Ma petite **Gemma**, mon **Krapo** préféré, la plus Bretonne des Espagnoles ou la plus Espagnole des Bretonnes ? En fait je ne saurais dire quelle est ta nationalité avec tes traductions un peu bancales, ta prononciation « approximative » (poupitre ! 😊) et en plus, telle une vraie Bretonne, tu cites le Mont Saint Michel comme merveille Armoricaine 😊 ! Blague à part, merci pour ta grande générosité, toutes ces bières partagées et les gossips du soir ainsi que les quelques

séances de sport (je dois reconnaître que tu m'as initié à la salle 😊). Bientôt ton tour viendra aussi et tu pourras te vanter d'être le plus **mignon des mini-docteurs de CRJBL** (mini par ta taille mais pas par tes connaissances, et écris en rose pour te combler de bonheur !). Bon courage pour cette dernière ligne droite et pour la suite, on aura bien l'occasion d'en reparler, toujours au détour d'une bière 🍷. Et surtout, n'oublie pas tu peux compter sur ma présence, je serais toujours au bout de mes messages et puis tu es bien évidemment la bienvenue à Köln, si l'envie de boire une délicieuse (ou pas, désolée d'avance ce n'est pas moi qui la brasse !) kölsch te prends 🍷. Pour finir avec toi sur une note espagnole : « Un amigo es el que siendo ciego te puede mirar y siendo sordo te puede escuchar ».

Passons au must des lab mates : **Morgane** ou plutôt **Morguy**, ces années de thèse ensemble, elles sont inoubliables et elles resteront à jamais gravées dans ma mémoire. Ma partenaire « de crime » (sans intention de faire un mauvais jeu de mot), toutes ces manips, ces repas et les quelques pauses café, parfois ~~très~~ très tard, devant le trieur et aussi une pensée à ces bouteilles sabrées (on taira à quel endroit 😊). Merci pour ton aide sur les manips souris, parce-que, toi-même tu sais : « tu peux sacrifier 10 souris, tu ne peux pas en sacrifier 30 [...] Ah bah si tu peux 😊 ! » On s'est payé de sacrées rigolades, ces longues manips n'auraient pas été les mêmes sans toi. Tellement de souvenirs ensemble, et sans aucun doute, ce sont les meilleurs de ma thèse (on n'oubliera pas que les meilleurs stim se lancent à minuit !) ! Merci d'avoir été ce pilier dans de nombreuses situations. Merci d'avoir été cette fille aussi timbrée que moi, j'ai la chance d'avoir partagé un grand morceau de ce chemin avec toi. Merci aussi pour ces blagues, je repense notamment à la fois où tu as caché ma souris (d'ordi) et je pensais perdre la tête à cause de la fatigue 😊. Et merci d'avoir toujours été au bout des messages pendant cette dernière année... Discuter avec toi a été un appui précieux, j'espère que tu quitteras de temps en temps ton sud Montpelliérain pour venir voir un peu la grisaille allemande de Cologne (pas pire que celle de Limoges je suis sûre) 😊. Je te souhaite le meilleur dans ta vie pro et perso, car c'est tout simplement la seule chose que tu mérites. Et surtout ; Salsamente !

Et puis, un merci tout particulier à quelques personnes extérieures au labo/à la science, dont le soutien, l'amitié et l'amour sont une force inébranlable au cours de ces années.

Je remercie donc **Audrey P.**, ma fidèle amie depuis plus de dix ans maintenant (un petit coup de vieux dit comme ça quand même !!) ! On ne se voit pas beaucoup mais à chaque soirée partagée ensemble c'est comme si on s'était vu la veille. Merci d'avoir toujours été présente, même au bout du téléphone, et d'avoir cru en moi, sans jamais douter de mes choix et de ma réussite. Je te souhaite le meilleur dans ta vie, tu le mérites tellement.

**Manon**, que pourrais-je te dire ? ! Ma cousine « coup de cœur » par alliance ! Tu es tellement rayonnante, merci pour tous ces moments de joie, ces fous rires incontrôlables et mention spéciale à la mission Ninja (seuls toi et Antoine pouvez comprendre). Tu fais une merveilleuse maman pour Nino et merci d'avoir été à mes côtés et de continuer à l'être.

Merci à **Coco**, **Sève** et **Ma belette**, pour les week-ends et moments en famille partagés pendant ces années de thèse. Merci aussi de m'avoir accueilli auparavant (en M1, ça remonte maintenant !) chez vous pour me permettre de faire mon stage où je le souhaitais ! Je vous souhaite plein de bonheurs. Maintenant vous aurez une nouvelle destination de vacances, un peu plus « fun » que Limoges (mais sans Bijou) !

Je remercie également ma « belle famille », et tout particulièrement **Nathalie**, pour s'être toujours réjouie pour moi ainsi que **Francis** pour nous avoir reçu pendant nos virées parisiennes et pour leur accueil chaleureux quand j'avais besoin d'un pied à terre pour aller en Allemagne et merci Francis parce que j'ai bien failli louper mon train cette fois-ci... 😊

Je souhaite exprimer toute ma gratitude à mes parents, ma petite **Mamita** et mon **Papou**. Depuis mon plus jeune âge vous me soutenez et m'encouragez dans tous ces défis, et vous ne m'avez pas lâché dans ce défi qu'est la thèse. Vous avez souvent pensé que j'étais malade (pour les manips qui durent 19h !) et vous aviez peut-être raison, mais vous m'avez toujours montré l'exemple du travail et donné les moyens de réussir, j'espère ne jamais vous décevoir. Je n'aurais jamais assez de mots et assez d'espace ici pour vous remercier comme il se doit. Pour toutes ces visios, toutes ces fois où vous êtes venus nous voir, à chaque instant vous avez été présents, MERCI pour tout, infiniment.

Enfin la dernière personne à remercier, et pas des moindres, **Antoine**. Tout ce chemin parcouru n'aurait pas été le même si tu n'avais pas été à mes côtés. Merci pour ta patience, ton soutien et ta joie de vivre, le tout sans faille ! Simplement, merci d'être la personne que tu es. Merci de m'avoir supportée (tout le monde sera d'accord pour dire que ce n'est pas si simple !), pendant ces quatre années de thèse mais aussi durant les sept précédentes 😊 et en avance pour toutes les suivantes... Merci aussi de m'accompagner à chaque étape, départ imminent pour le post doc, même pas dans une nouvelle ville mais carrément un autre pays, j'ai conscience de la chance que j'ai de t'avoir avec moi. Je n'en dirais pas plus ici mais tu sais à quel point tu es essentiel à mon équilibre, Merci, mille fois. ❤️ A notre future vie (Auf unser zukünftiges Leben in Köln 😊), nos futurs défis et bien sûr nos futures randos (GR20 et Tour du Mont Blanc n'attendent que nous !).

Et puis évidemment « miaourci » à mon chat, **Lucky**, pour tous ses câlins si réconfortants et sa compagnie pendant ce temps de rédaction. 🐾

Loin de moi l'envie de plomber ces remerciements, mais cela me tient particulièrement à cœur de finir en adressant mes pensées les plus tendres, à toi mon amie d'enfance partie beaucoup trop tôt, **Laure**. Ceci est un petit peu ma façon de te rendre hommage, à toi, la 18<sup>ème</sup> victime d'un horrible féminicide de cette année 2023... et aussi de témoigner tout mon soutien à tes parents, qu'ils obtiennent justice pour toi et qu'ils puissent soutenir ton fils comme ils le souhaitent. Ton départ prématuré m'a percuté en plein cœur et m'a fait prendre pleinement conscience de la chance que j'ai de vivre... Alors, chaque fois que j'hésiterais à prendre une décision par peur, je regarderais le ciel et repenserais encore plus à toi pour toujours me souvenir que je possède cette chance inouïe de pouvoir vivre et prendre des risques. J'aime me dire que tu as rejoint ton oncle, cet homme que je considérais presque comme un « deuxième père » et qu'à ses côtés, tu peux reposer en paix.

Et puis, pour quand même clôturer ces remerciements sur une note plus légère, merci **la thèse**, tu m'as permis de développer des qualités inespérées pour moi, comme la patience, tu m'as également permis de passer au-dessus de ma timidité... Tu m'as bien évidemment construit un socle solide pour continuer dans la recherche et aller expérimenter ailleurs... Et puis, surtout tu as mis sur mon chemin des personnes incroyables que je n'aurais jamais rencontré sans toi. Et maintenant, grâce à toi, en route vers de nouveaux horizons et de nouvelles aventures car **la vie n'a pas le temps d'attendre pour être vécue**.

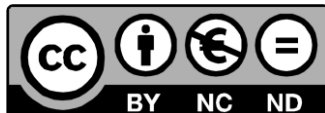
## Droits d'auteurs

---

Cette création est mise à disposition selon le Contrat :

« **Attribution-Pas d'Utilisation Commerciale-Pas de modification 3.0 France** »

disponible en ligne : <http://creativecommons.org/licenses/by-nc-nd/3.0/fr/>



## Table des matières

---

Avant-Propos .....	2
Partie I. Immunoglobulines : de la protéine aux gènes .....	5
I.1. Les immunoglobulines .....	5
I.2. Les gènes d'Ig .....	6
I.2.1. Loci des chaînes légères (IgL).....	6
I.2.1.1. Le locus <i>Igλ</i> .....	6
I.2.1.2. Le locus <i>Igκ</i> .....	6
I.2.2. Locus des chaînes lourdes ( <i>IgH</i> ) .....	7
Partie II. Ontogenèse B normale et Lymphomagenèse.....	9
II.1. Ontogenèse B normale .....	9
II.1.1. Phase précoce de l'ontogenèse B : indépendance à l'antigène .....	9
II.1.1.1. Cellules souches et précurseurs non B .....	9
II.1.1.2. Stade pré-pro B : porte d'entrée du lignage B .....	10
II.1.1.3. Stade pro-B : acquisition d'une chaîne lourde réarrangée .....	10
II.1.1.4. Stade pré-B : acquisition du BCR.....	11
II.1.1.5. Cellule B immature : acquisition de la tolérance centrale.....	11
II.1.2. Phase tardive du développement : dépendance à l'antigène .....	12
II.1.2.1. Cellules B transitionnels : phénotype intermédiaire .....	12
II.1.2.2. Cellules B matures .....	13
II.1.2.2.1. Cellules B1 .....	13
II.1.2.2.2. Cellules B2 .....	13
II.1.2.3. Activation des cellules B : formation du centre germinatif.....	13
II.1.2.4. Différenciation plasmocytaire .....	15
II.1.2.5. Cellules B mémoires .....	16
II.1.3. Remaniements géniques associés au développement B.....	16
II.1.3.1. La recombinaison V(D)J, évènement précoce .....	16
II.1.3.1.1. Mécanisme .....	16
II.1.3.1.2. Régulation .....	17
II.1.3.2. Hypermutation somatique et commutation de classe : évènements caractéristiques du développement tardif.....	19
II.1.3.2.1. AID, enzyme essentielle .....	19
II.1.3.2.2. Hypermutation somatique .....	19
II.1.3.2.2.1. Mécanisme.....	20
II.1.3.2.2.2. Dommages collatéraux.....	21
II.1.3.2.2.3. Régulation.....	22
II.1.3.2.3. Commutation de classe .....	23
II.1.3.2.3.1. Mécanisme.....	23
II.1.3.2.3.2. Régulation.....	24
II.2. Lymphomagenèse B .....	25
II.2.1. Généralités.....	25
II.2.2. Altérations liées à la maturation des gènes d'Ig.....	26
II.2.2.1. Evènements de translocation pendant la recombinaison V(D)J.....	26
II.2.2.2. Altérations géniques lors du développement tardif .....	27
II.2.2.2.1. Associées à la CSR.....	27
II.2.2.2.2. Associées à la SHM.....	27



Partie III. Éléments <i>cis</i> -régulateurs du locus <i>IgH</i> .....	28
III.1. Régions isolatrices et promoteurs .....	28
III.1.1. Régions isolatrices .....	28
III.1.2. Promoteurs .....	28
III.1.2.1. Promoteurs $pV_H$ .....	28
III.1.2.2. Promoteur pDQ52 .....	29
III.1.2.3. Promoteurs I .....	29
III.2. La région $E\mu$ MARs .....	29
III.2.1. L'enhancer $cE\mu$ .....	29
III.2.2. Les régions d'attachement à la matrice ( <i>MAR</i> ) .....	31
III.3. La région régulatrice en 3' ( $3'RR$ ) .....	32
Partie IV. Organisation du noyau et architecture de la chromatine .....	53
IV.1. Territoires chromosomiques et compartiments A et B .....	53
IV.2. TAD: "Topologically Associated Domains" .....	54
IV.3. Mécanisme d'extrusion de boucle .....	55
IV.3.1. Le modèle de l'extrusion de boucle .....	55
IV.3.2. L'extrusion de boucle chez le lymphocyte B .....	56
IV.3.2.1. L'extrusion de boucle pendant la recombinaison V(D)J .....	57
IV.3.2.2. L'extrusion de boucle pendant la CSR .....	57
IV.4. Organisation du noyau chez le lymphocyte B et rôle des enhancers .....	59
Objectifs .....	73
Résultats .....	76
Discussion et Perspectives .....	89
Références bibliographiques .....	101
Annexes .....	124

## Liste des abréviations

---

**3C** : Chromosome  
Conformation Capture

**3'RR** : 3' Regulatory Region

### - A -

**A** : Adénine

**ABC** : Activated B Cell

**ADN** : Acide Désoxyribo  
Nucléique

**A-EJ** : Alternative End  
Joining

**AID** : Activation Induced  
cytidine Deaminase

**APE** : Apurinique/  
Apyrimidique  
endonucleases

**ATM** : Ataxia aelangiectasia  
mutated

### - B -

**BCL** : B Cell Lymphoma

**BCR** : B Cell Receptor

**BER** : Base Excision Repair

### - C -

**C** : Cytosine

**CBE** : CTCF Binding Element

**CD** : Cluster de  
différenciation

**CDB** : Cassures Double Brin

**CDF** : Cellule Dendritique  
Folliculaire

**CDR** : Complementary  
Determining Region

**ChIP** : Immunoprécipitation  
de Chromatine

**CLP** : Common Lymphoid  
Precursors

**CMH** : Complexe Majeur  
d'Histocompatibilité

**CPA** : Cellule Présentatrice  
d'Antigène

**CSR** : Classe Switch  
Recombination

**CSRC** : Centre de CSR

**CTCF** : CCCTC Binding Factor

**CtIP** : Ct-interacting Protein

### - D -

**DICE** : Downstream Ig  
Control Element

**DIVAC** : DIVersification  
ACTivaotr

**DLBCL** : Lymphome Diffus à  
Grande Cellules B

**DNA-PK** : Protéine Kinase  
dépendante de l'ADN

### - E -

**EBF** : Early B cell Factor

### - F -

**FISH** : Hybridation *in situ*  
fluorescente

**FO** : Folliculaire

**FR** : Framework

### - G -

**G** : Guanine

**GC** : Centre Germinatif

### - H -

**HMGB** : High Mobility  
Group Box

**Hs** : Hypersensibility

**HTGTS** : High Throughput  
Genome wide Translocation  
Sequencing

### - I -

**I** : Intervening

**Ig** : Immunoglobuline

**IgH** : locus des chaines  
lourdes d'Ig

**IgL** : locus des chaines  
légères d'Ig

**Igκ** : locus de la chaine  
légère κ

**Igλ** : locus de la chaine  
légère λ

**IL** : Interleukine

### - K -

**kb** : Kilobases

**KO** : Knock Out

### - L -

**LAM** : Linear Amplification  
Mediated

**LAD** : Lamina Associated  
Domains

**LNH** : Lymphome Non  
Hodgkinien

- M -

**MAR** : Matrix Attachment Regions

**MLH** : MutL Homologie

**MSH** : MutS Homolog

**MZ** : Zone Marginale

- N -

**NHEJ** : Non Homologous End Joining

**NK** : Natural Killer

- O -

**OMS** : Organisation Mondiale de la Santé

- P -

**PAIR** : Pax-Activated Intergenic Repeat

**PARP1** : Poly-ADP Ribose Polymerase 1

**PAX5** : Paired Box 5

**pb** : paire de bases

**PCNA** : Proliferating Cell Nuclear Antigen

**PMS** : PostMeiotic Segregation

- R -

**RAG** : Recombination Activating Gene

**RSS** : Séquence Signal de Recombinaison

- S -

**S** : Switch

**SHM** : Hypermutation Somatique

**SMC** : Structural Maintenance of Chromosome

**SPT5** : Suppressor or Ty5 homolog

- T -

**T** : Thymine

**TAD** : Topologically Associated Domain

**TC** : Territoire Chromosomique

**TCR** : T Cell Receptor

**TdT** : Terminal Desoxynucleotidyl Transferase

**T<sub>H</sub>** : Lymphocyte T Helper

**TLR** : Toll Like Receptor

- U -

**U** : Uracile

**UNG** : Uracile N Glycosylase

- W -

**wt** : wild type

- Z -

**ZC** : Zone Claire

**ZS** : Zone Sombre

## Table des illustrations

---

### Synthèse Bibliographique

**Figure 1** : Structure d'une immunoglobuline

**Figure 2** : Représentation schématique des loci des immunoglobulines

**Figure 3** : Phase précoce de l'ontogénie B, indépendante de l'antigène

**Figure 4** : Phase tardive de l'ontogénie B

**Figure 5** : Structure schématique d'un centre germinatif

**Figure 6** : Mécanisme de la recombinaison V(D)J

**Figure 7** : Représentation schématique de la structure d'AID

**Figure 8** : Localisation de l'hypermutation somatique

**Figure 9** : Illustration des différentes prises en charge du mésappariement dU:dG initié par AID

**Figure 10** : Classement des gènes AID "off-target" en trois groupes

**Figure 11** : Régulation du ciblage par AID lors de la SHM

**Figure 12** : Mécanisme de la recombinaison de classe

**Figure 13** : Voies de réparation recrutées au niveau des cassures doubles brins de l'ADN pendant la CSR

**Figure 14** : Régulation du recrutement de AID aux régions switch pendant la CSR

**Figure 15** : Représentation des contreparties physiologiques des principaux lymphomes non hodgkiniens et du myélome multiple

**Figure 16** : Translocation pendant la recombinaison VDJ, exemple de la translocation t(14;18) dans le lymphome folliculaire

**Figure 17** : Translocation pendant la CSR, exemple de la translocation t(8;14) dans le lymphome de Burkitt

**Figure 18** : Mutation d'oncogènes pendant le processus de SHM, exemple de Bcl6 dans le DLBCL

**Figure 19** : Eléments cis-régulateurs du locus IgH

**Figure 20** : Représentation de l'organisation du noyau en territoires chromosomiques et compartimentation A/B

**Figure 21** : Exemple de carte de contacts révélant la présence de TAD

**Figure 22** : Schématisation des fonctions des TAD

**Figure 23** : Représentation schématique de la cohésine et du mécanisme d'extrusion de boucle

**Figure 24** : Mécanisme d'extrusion de boucle pendant la recombinaison VDJ

**Figure 25** : Mécanisme d'extrusion de boucle pendant la CSR



## Résultats

**Figure 26** : Involvement of *EμMARs* region in *IgH* chromatin loops

**Figure 27** : Scheme of mouse models and gating strategy to sort GC B cells

**Figure 28** : Study of *Bcl6* mutation frequency

**Figure 29** : Quantification of *Bcl6* primary transcripts

**Figure 30** : *Bcl6* position in nucleus is modified

**Figure 31** : Study of *cMyc* mutations

**Figure 32** : Quantification of *cMyc* primary transcription

**Figure 33** : Nuclear localization of *cMyc*

**Figure 34** : Mutation Analysis in intron 1 of *Pim1* and *Cd83*

**Figure 35** : Quantification of primary transcription

## Discussion et perspectives

**Figure 36** : Localisation de *Bcl6* en contexte *wt* et *3'RR*-mutant

**Figure 37** : Localisation de *cMyc* en contexte *wt* et *3'RR*-mutant

**Figure 38** : Taux de mutation de *Bcl6* dans des lignées cellulaires de DLBCL

**Figure 39** : Répartition des évènements de recombinaison dans *IgH* vs. le reste du génome

**Figure 40** : Fréquence de mutation des gènes « AID off target » dans les mutants *EμMARs*

# *Avant-Propos*

---

## Avant-Propos

---

Le système immunitaire est constitué d'un ensemble d'organes, de tissus et de cellules spécialisés qui coopèrent afin d'assurer la défense d'un organisme vivant. Chacun de ces éléments possède une fonction précise, et ensemble, ils participent au maintien de l'homéostasie cellulaire et permettent la dégradation et l'élimination d'agents pathogènes.

En cas d'intrusion d'un pathogène dans l'organisme, la réponse immune se déploie en deux temps. Une première réponse, très rapide et non spécifique, fait intervenir les acteurs de l'immunité innée. Ces acteurs sont cellulaires avec l'intervention des cellules NK (pour *Natural Killer*), des macrophages ou encore des mastocytes. Les barrières physiques telles que la peau ou les muqueuses constituent également des acteurs importants de l'immunité innée. Ensuite, la seconde ligne de défense, appelée immunité adaptative, intervient. Bien qu'elle prenne plus de temps à être mise en place, elle est plus précise et efficace puisqu'elle s'adapte au pathogène rencontré. Les lymphocytes T, responsables de la réponse à médiation cellulaire, et les lymphocytes B, acteurs de la réponse humorale *via* la production d'anticorps, sont les deux types cellulaires qui interviennent dans cette phase adaptative. Contrairement à la réponse innée, la réponse adaptative est à l'origine de l'acquisition d'une mémoire immunitaire qui permet une réaction plus rapide et plus efficace lors de rencontres ultérieures avec le même pathogène.

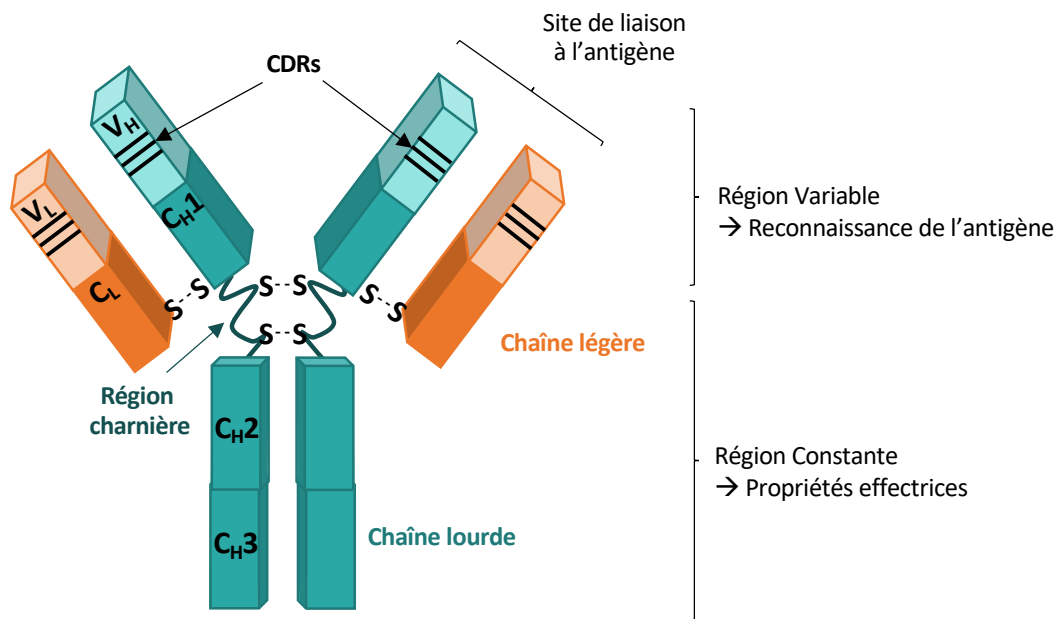
Pour fournir une réponse spécifique, les lymphocytes exercent des remaniements géniques leur permettant d'exprimer à leur surface un récepteur spécifique de l'antigène : le TCR (pour *T Cell Receptor*) pour les lymphocytes T et le BCR (pour *B Cell Receptor*) pour les lymphocytes B. Chez la cellule B, le BCR correspond à une immunoglobuline (Ig) de surface, de même spécificité que sa forme sécrétée (anticorps). Le lymphocyte B est capable de produire une grande variété d'Ig à partir de seulement 3 loci grâce à des remaniements géniques qui se produisent à différents moments de son développement. Bien qu'il constitue une caractéristique unique de la cellule B et qu'il soit impératif à la mise en place d'une réponse adaptée, ce dernier point présente un risque pour l'intégrité du génome de la cellule B et peut entraîner l'émergence de pathologies telles que les leucémies ou les lymphomes. Il est ainsi crucial que les remaniements géniques des loci d'Ig soient précisément régulés afin d'éviter les anomalies génétiques potentielles. Les régions *cis*-régulatrices des loci d'Ig, telles que la région *E $\mu$ MARs* et la région régulatrice en 3' (*3'RR*), sont responsables de la coordination ordonnée de ces événements. Dans le cadre de ma thèse, j'ai exploré le rôle de ces régions dans l'organisation du génome de la cellule B à différentes échelles et leur implication dans le maintien de l'intégrité du génome. À l'aide de plusieurs modèles murins présentant des délétions totales ou partielles de chacune de ces deux régions, nous avons exploré par un ensemble de techniques innovantes l'adressage du locus *IgH* au sein du noyau, son accessibilité et la formation de boucles nécessaires au

bon déroulement des remaniements géniques. Nous nous sommes également intéressés au rôle des régions régulatrices et de l'organisation du noyau dans le maintien de l'intégrité du génome de la cellule B.



# *Synthèse Bibliographique*

---



**Figure 1 : Structure d'une immunoglobuline**

Exemple d'une IgG constituée de 2 chaînes lourdes (bleu), chacune constituée d'un domaine variable (V<sub>H</sub>) et de trois domaines constants (C<sub>H</sub>1, C<sub>H</sub>2 et C<sub>H</sub>3) et de 2 chaînes légères (orange), comprenant chacune un domaine variable (V<sub>L</sub>) et un seul domaine constant (C<sub>L</sub>). Les domaines C<sub>H</sub>1 et C<sub>H</sub>2 sont séparés par la région charnière permettant l'association des 2 chaînes lourdes par formation de ponts disulfures et conférant la flexibilité à l'Ig. Les régions V (H+L) contiennent trois régions hypervariables nommées CDR 1 à 3 (« Complementary determining region »). La chaîne lourde et la chaîne légère s'associent grâce à la formation de pont disulfure.

## Partie I. Immunoglobulines : de la protéine aux gènes

---

### I.1. Les immunoglobulines

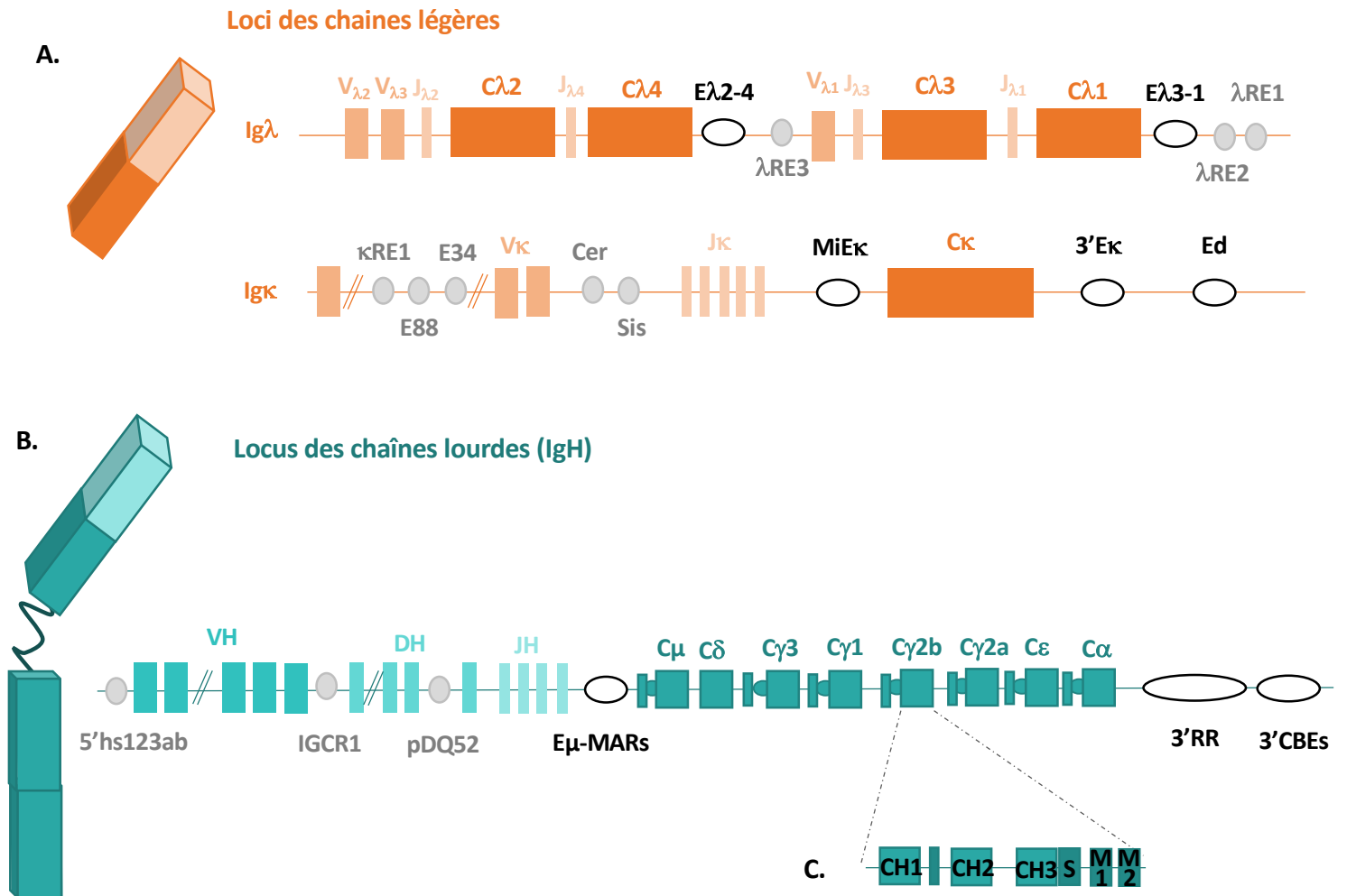
Une immunoglobuline (Ig) est une glycoprotéine de taille importante ( $\approx 150$  KDa). Elle est constituée de deux chaînes lourdes (H, Heavy) et de deux chaînes légères (L ; Light), identiques entre elles (Edelman, 1973) (**Figure 1**). L'Ig peut être membranaire, et en association avec l'hétérodimère  $Ig\alpha/Ig\beta$ , forme le BCR (*B cell receptor*). Elle peut également être sécrétée sous forme soluble et est alors appelée anticorps.

Les chaînes lourdes sont associées entre elles par deux ponts disulfures et chaque chaîne lourde est également associée à une chaîne légère par un pont disulfure. Chacune des chaînes lourdes est composée d'un domaine peptidique variable (VH) et d'au moins trois domaines constants (CH1, CH2 et CH3). Une région charnière est présente entre les domaines CH1 et CH2 des Ig, excepté dans le cas des IgM et des IgE. Ces dernières conservent tout de même des propriétés flexibles grâce à la présence d'un domaine CH additionnel (CH4) (Padlan, 1994) (**Figure 1**). Les chaînes légères sont également constituées d'un domaine peptidique variable (VL) mais ne contiennent qu'un seul et unique domaine constant (CL).

Les domaines variables de l'Ig sont impliqués dans la reconnaissance antigénique, notamment *via* des régions peptidiques hypervariables nommées CDR1, CDR2 et CDR3 (CDR pour *Complementary Determining Region*). Les CDR sont elles-mêmes cernées par quatre régions charpentes FR1, FR2, FR3 et FR4 (FR pour *Framework*). Les domaines CL ne porteraient qu'une fonction structurale alors que les domaines CH sont, quant à eux, porteurs des propriétés effectrices de l'Ig. Ces propriétés effectrices dépendent de la classe d'Ig, qui peut donc être de cinq types différents : IgM, IgD, IgG, IgA et IgE.

Les IgM sont les Ig produites en première intention suite à une stimulation antigénique, elles possèdent un haut pouvoir agglutinant des micro-organismes et constituent une première ligne de défense à part entière. Les IgD sont produites notamment lors d'infection bactérienne alors que les IgE sont plutôt produites lors d'infestation parasitaire et sont également incriminées dans certains processus allergiques. Les IgA jouent un rôle important dans l'immunité mucoale. Enfin, les IgG sont le sous-type le plus fréquent et sont impliquées dans la défense contre de multiples antigènes. Elles sont produites lors de réponses antivirales, antibactériennes et antiparasitaires (Schroeder and Cavacini, 2010).

Les différentes fonctions sont permises par la variabilité des domaines CH qui composent les différents isotopes. Ces domaines CH peuvent être responsables du recrutement de certaines cellules immunitaires, comme les macrophages, qui possèdent à leur surface des récepteurs (nommés FcR)



**Figure 2 : Représentation schématique des loci des immunoglobulines**

- A. Loci des chaînes légères
- B. Locus des chaînes lourdes
- C. Détail d'un segment CH, ici exemple de  $C\gamma 2b$  constitué de 3 exons CH, suivi de l'exon de sécrétion S et de deux exons de membranes M

reconnaissant ces domaines (Davies and Metzger, 1983). Les cellules peuvent ainsi jouer leur rôle et initier l'élimination du pathogène reconnu par l'anticorps. Le système du complément peut également être activé par les Ig : il permet une attaque directe membranaire du pathogène *via* l'action d'une cascade protéique.

## **I.2. Les gènes d'Ig**

Une Ig est codée par des gènes complexes localisés sur des chromosomes distincts. Ainsi la chaîne lourde est codée par le locus *IgH* alors que la chaîne légère peut être codée par les loci *Igκ* ou *Igλ*. Ces loci possèdent des points communs, notamment la présence de segments *V* (Variabilité) et *J* (Jonction), des segments *C* (Constants) ainsi que des éléments *cis*-régulateurs. Le locus *IgH* se distingue néanmoins des deux autres par la présence de segments *D* (Diversité). La structure germinale de ces loci est détaillée ci-après.

### **I.2.1. Loci des chaînes légères (IgL)**

#### **I.2.1.1. Le locus *Igλ***

Chez la souris, le locus *Igλ* est localisé sur le chromosome 16. Il occupe 200 kilobases (kb) et contient 3 segments *Vλ*, 4 segments *Jλ* qui, réarrangés, forment la région VL de l'Ig et 4 segments *Cλ* (dont *Cλ4* qui n'est pas fonctionnel). Ces segments *λ* sont organisés de façon particulière avec une scission en deux groupes contenant chacun des segments *Vλ*, *Jλ* et *Cλ* (**Figure 2A**) :

- *Vλ2 – Vλ3 – Jλ2 – Cλ2 – Jλ4 – Cλ4*
- *Vλ1 – Jλ3 – Cλ3 – Jλ1 – Cλ1*

Souvent, les réarrangements opérés lors de la recombinaison VJ se font entre des segments *Vλ* et *Jλ* proches, notamment entre *Vλ2 / Vλ3* et *Jλ2* ou *Vλ1* et *Jλ3* (Carson and Wu, 1989; Storb et al., 1989).

Les deux groupes de gènes sont séparés par deux éléments *cis*-régulateurs : *Eλ2-4* et *λRE3*. Après le deuxième groupe de gènes, trois autres éléments *cis*-régulateurs analogues sont également présents: *Eλ3-1*, *λRE2* et *λRE1* (Hagman et al., 1990; Predeus et al., 2014). Chez la souris, seulement 5% des cellules B expriment une chaîne légère de type *λ* (Arakawa et al., 1996; Engel et al., 1999).

#### **I.2.1.2. Le locus *Igκ***

Le locus *Igκ* est beaucoup plus grand que le locus *Igλ* et s'organise différemment. En effet, il s'étend sur 3 Mégabases (Mb) sur le chromosome 6 chez la souris. Il se compose d'une succession de 176 segments *Vκ* dont un peu moins d'une centaine sont fonctionnels, de 5 segments *Jκ* dont 4 fonctionnels et d'un seul et unique segment constant *Cκ* (Giudicelli et al., 2005; Schlissel, 2004). Ce

locus est aussi caractérisé par la présence de divers éléments *cis*-régulateurs (**Figure 2A**). Initialement, ce sont trois régions régulatrices qui ont été décrites. De 5' en 3', la première, localisée en amont de  $C\kappa$ , est  $MI E\kappa$ . Elle se compose d'une région d'attachement à la matrice (MAR pour *Matrix Attachment Region*) et de l'enhancer  $E\kappa$  (Emorine et al., 1983; Xu et al., 1996). En aval de  $C\kappa$  sont situés l'enhancer  $3'E\kappa$  et l'élément distal  $Ed$  (Inlay et al., 2002; Liu et al., 2002). Cinq autres éléments *cis*-régulateurs ont été décrits un peu plus tardivement. Il s'agit de  $\kappa RE1$ ,  $E88$  et  $E34$  localisés au sein des segments  $V\kappa$  ainsi que  $Cer$  et  $Sis$  (également nommés  $hs1.2$  et  $hs3.4.5.6$  respectivement, *hs* pour hypersensitivité à la DNase I), situés entre le dernier segment  $V\kappa$  et  $J\kappa1$  (Barajas-Mora et al., 2023, 2019; Predeus et al., 2014; Xiang et al., 2013, 2011) (**Figure 2A**).

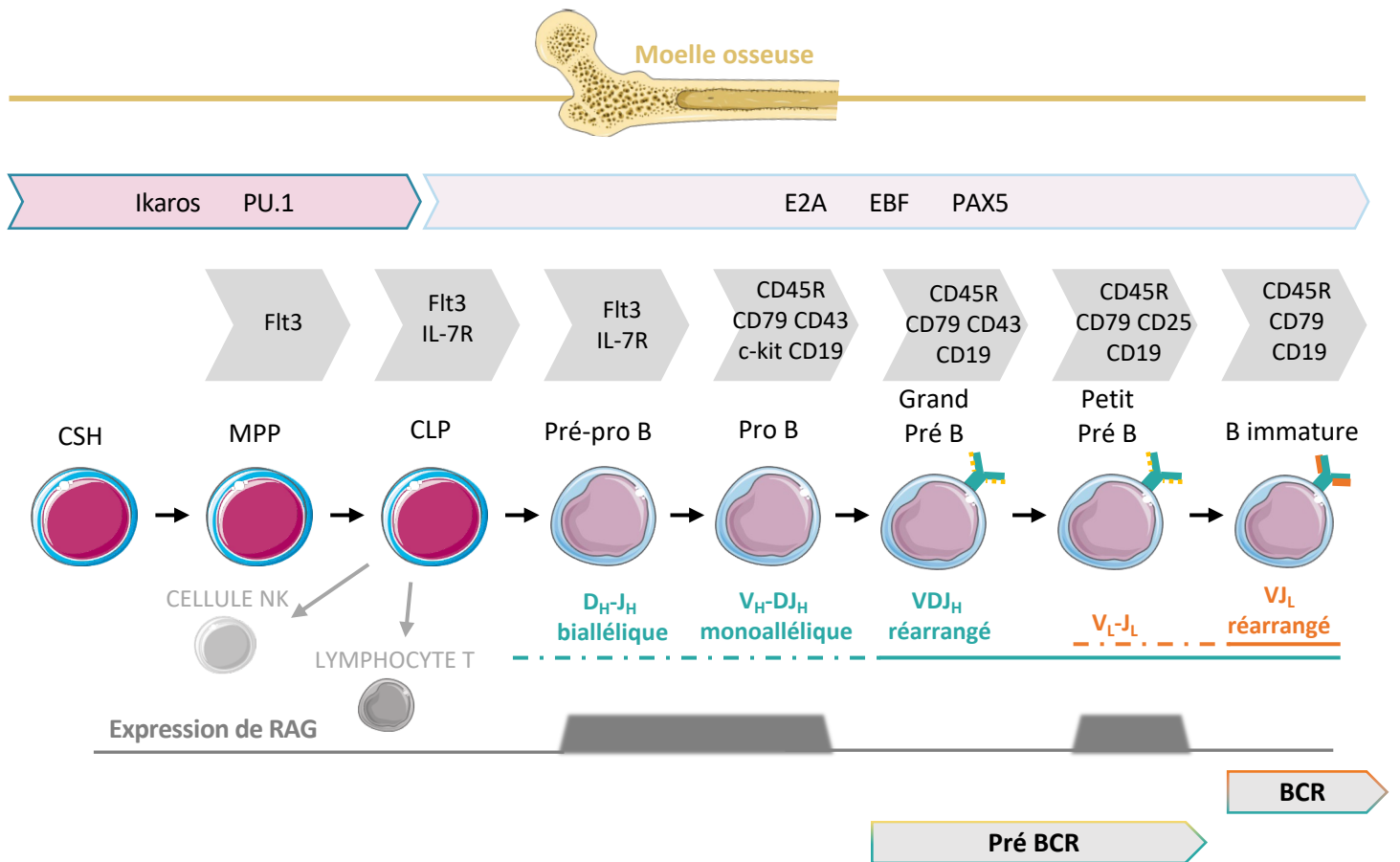
### 1.2.2. Locus des chaînes lourdes (*IgH*)

Chez la souris, le locus *IgH* est localisé sur le chromosome 12. Etendu sur 3Mb, il arbore une organisation très similaire au locus *IgK*. Il est composé d'une succession de 160 segments  $V_H$ , dont une centaine sont fonctionnels, nommés « distaux, intermédiaires et proximaux » (selon leur position sur le locus). La nomenclature actuelle identifie 15 familles  $V_H$  différentes (*Ighv1* à *Ighv15*, avec *Ighv1* majoritaire) (Johnston et al., 2006). Chacun des segments  $V_H$  fonctionnels est précédé par son promoteur  $pV_H$ . Les segments  $V_H$  sont suivis par 20 segments  $D_H$  (dont 14 sont fonctionnels) répartis dans 4 familles distinctes (*DFL16*, *DSP2*, *DST4* et *DQ52*). Ces segments de diversité sont également précédés de promoteur mais un seul d'entre eux présente une activité fonctionnelle importante :  $pDQ52$ . Enfin, 4 segments  $J_H$ , tous fonctionnels, clôturent la partie codant le domaine variable de l'Ig (Giudicelli et al., 2005). Au-delà de cette importante région génomique se situent les 8 segments constants codant les 8 classes d'Ig nommés (de 5' vers 3') :  $C\mu$ ,  $C\delta$ ,  $C\gamma3$ ,  $C\gamma1$ ,  $C\gamma2b$ ,  $C\gamma2a$  (ou  $C\gamma2c$  selon la lignée murine),  $C\epsilon$  et  $C\alpha$  (Morgado et al., 1989) (**Figure 2B**). Chaque gène constant, à l'exception de  $C\delta$ , est précédé d'une région switch « S » ainsi que de l'exon I (I pour *intervening*). Chaque gène  $Cx$  est composé de différents exons  $CH$  (entre 1 et 4 selon l'isotype) suivi de l'exon S de sécrétion et des exons membranaires  $M$  (**Figure 2C**).

Divers éléments *cis*-régulateurs sont décrits dans ce locus. En amont des segments  $V_H$  se situe une région isolatrice composée de 4 sites *hs*, appelée  $5'hs123ab$  (Pawlitzky et al., 2006). Entre les segments  $V_H$  et  $D_H$ , une région intergénique de contrôle (*IGCR1*) formée par 6 sites *hs* ( $hs1$  à  $hs6$ ) est présente (Alessandrini and Desiderio, 1991). Les segments  $Cx$  sont, quant à eux, encadrés par les deux éléments *cis*-régulateurs majeurs du locus : la région  $E\mu MARS$  et la  $3'RR$ . La région  $E\mu MARS$  prend place entre le dernier segment  $J_H$  et la région  $S\mu$  alors que la  $3'RR$  s'étend en aval de  $C\alpha$  (Banerji et al., 1983; Gillies et al., 1983; Lieberson et al., 1991). L'extrémité 3' du locus *IgH* est caractérisé par la présence de 4

sites *hs* (*hs5* à *hs8*) contenant 10 sites de fixation de la protéine CTCF (CCCTC binding Factor), appelés *3'CBEs* (pour *3' CTCF Binding Elements*) (**Figure 2B**).

Ces éléments cis-régulateurs jouent des rôles cruciaux pendant les événements de recombinaison effectués par la cellule B au cours de son développement. La description détaillée de leurs rôles sera faite dans la partie III.



**Figure 3 : Phase précoce de l'ontogénie B, indépendante de l'antigène**

Cette figure représente la phase précoce de l'ontogénèse B qui prend place dans la moelle osseuse. Elle récapitule les principaux stades, les marqueurs de surface associés, les facteurs nécessaires à la progression à travers tout les stades et l'occurrence des remaniements géniques des chaînes lourdes et légères.



## Partie II. Ontogenèse B normale et Lymphomagenèse

---

### II.1. Ontogenèse B normale

Le développement lymphocytaire B correspond à l'ensemble des stades à travers lesquels évolue un précurseur jusqu'au stade de différenciation ultime qu'est le plasmocyte, cellule spécialiste de la production d'anticorps de haute affinité. Ce processus est divisé en deux phases, distinctes par leur chronologie, lieu et événements occurrents : la phase précoce et la phase tardive.

La phase précoce du développement prend place dans la moelle osseuse et est définie par l'acquisition de l'expression du récepteur spécifique des cellules B, le BCR, grâce à la recombinaison V(D)J. La phase tardive se tient en périphérie dans les organes lymphoïdes secondaires et est quant à elle caractérisée par l'hypermutation somatique (SHM pour *Somatic HyperMutation*) et le switch isotypique (CSR pour *Class Switch Recombination*) permettant respectivement d'augmenter l'affinité de l'anticorps pour son antigène et d'adapter les fonctions effectrices de l'Ig.

#### II.1.1. Phase précoce de l'ontogenèse B : indépendance à l'antigène (Figure 3)

##### II.1.1.1. Cellules souches et précurseurs non B

Le développement précoce du lymphocyte B prend place dans la moelle osseuse à partir d'une cellule souche hématopoïétique (CSH). Les CSH, caractérisées par leur multipotence et leur haute capacité d'auto-renouvellement, vont se différencier en progéniteurs multipotents (MPP pour *MultiPotent Progenitors*) grâce notamment à l'expression de deux facteurs : PU.1 et Ikaros (**Figure 3**) (Hardy et al., 2007; Monroe and Dorshkind, 2007; Yoshida et al., 2006).

Les MPP vont ensuite se différencier en précurseurs lymphoïdes communs (CLP pour *Common Lymphoid Precursor*) (Chao et al., 2008). Cette différenciation est en partie médiée par la signalisation induite par le récepteur à tyrosine kinase Flt3 exprimé à la surface des MPP dont le ligand est exprimé par les cellules stromales de la moelle (Adolfsson et al., 2001; Sitnicka et al., 2002). Le facteur Ikaros est aussi indispensable à la progression des MPP en CLP puisque sa perte entraîne l'incapacité des MPP à se différencier en CLP (Yoshida et al., 2006). Cette dépendance à Ikaros et PU.1 peut s'expliquer par le fait qu'ils sont régulateurs de l'expression de Flt3 (DeKoter et al., 2002; Nichogiannopoulou et al., 1999).

Les CLP sont capables de s'engager dans différents lignages selon les signaux qu'ils perçoivent dans leur environnement médullaire et sont à l'origine notamment des cellules NK, des lymphocytes T et des lymphocytes B. Les CLP expriment à leur membrane le récepteur à l'interleukine (IL) 7 (IL-7R), dont l'expression est régulée par PU.1 et la signalisation en aval de Flt3 (DeKoter et al., 2002; Medina et al., 2004). Suite à une stimulation de ce récepteur par l'IL-7, le CLP va exprimer le facteur de transcription

E2A (E box protein binding 2A). E2A est essentiel à la progression dans le lignage B puisqu'un blocage complet du développement B est observé chez des souris dépourvues de ce facteur (Zhuang et al., 1994). L'action combinée de E2A, PU.1 et la signalisation en aval de IL-7R va initier l'expression d'EBF (Early B cell Factor), un facteur de transcription qui permet l'expression de différents facteurs tels que PAX5 (Paired box 5), protéine indispensable aux phases précoces du développement B (Medina et al., 2004; Seet et al., 2004).

#### **II.1.1.2. Stade pré-pro B : porte d'entrée du lignage B**

L'expression d'EBF signe l'entrée dans le lignage B puisqu'en son absence un défaut majeur du développement B est observé chez la souris (Lin and Grosschedl, 1995). La protéine PAX5 est aussi un acteur critique du développement B, avec à la fois un rôle activateur de gènes typiques du lignage B comme *Cd19* et un rôle répresseur de certains gènes caractéristiques de la différenciation T tel que *Notch1* (Busslinger, 2004; Chiang and Monroe, 2001, 1999; Cobaleda et al., 2007; Delogu et al., 2006; Kozmik et al., 1992; Nutt et al., 1999; Souabni et al., 2002). En effet, le rôle de PAX5 a été démontré dans un modèle murin présentant une délétion conditionnelle de *Pax5* dans lequel les cellules pro-B gardent le potentiel de différenciation vers différents lignages (Mikkola et al., 2002). L'expression concomitante de ces facteurs marque alors la transition de CLP vers le premier stade de la différenciation lymphocytaire B : le stade pré-pro B (**Figure 3**). Ce stade est aussi défini par l'initiation des premiers remaniements géniques inhérents à la recombinaison V(D)J dont le mécanisme sera détaillé dans un prochain paragraphe (II.1.3.1). En effet, la cellule pré-pro-B va dans un premier temps réaliser le réarrangement entre un segment  $D_H$  et un segment  $J_H$ , formant ainsi un unique segment  $DJ_H$ , et ce sur chaque allèle.

#### **II.1.1.3. Stade pro-B : acquisition d'une chaîne lourde réarrangée**

Ce stade est caractérisé par l'expression de différents marqueurs à la surface de la cellule tels que CD45R, CD79a, CD79b, CD43 ainsi que c-kit (ou CD117) et CD19 (Hardy et al., 2007). La cellule pro-B va initier un deuxième réarrangement génique afin de recombiner un segment  $V_H$  avec le segment  $DJ_H$  préalablement réarrangé (Alt et al., 1992, 1984; Jung et al., 2006). Ce réarrangement  $V_H-DJ_H$  est mono-allélique : la cellule pro-B va d'abord réaliser la recombinaison  $V_H-DJ_H$  sur un allèle, si ce premier évènement conduit à un réarrangement fonctionnel, l'allèle est « productif » ; il n'y aura alors pas d'autres réarrangements. En revanche, si la recombinaison conduit à un réarrangement non fonctionnel, la cellule pro-B va initier le réarrangement de son deuxième allèle. Cet ordonnancement et la séquentialité des évènements de recombinaisons précoces sont à l'origine de l'exclusion allélique : il n'y a alors qu'une seule chaîne lourde unique exprimée par la cellule (Corcoran, 2005). Si

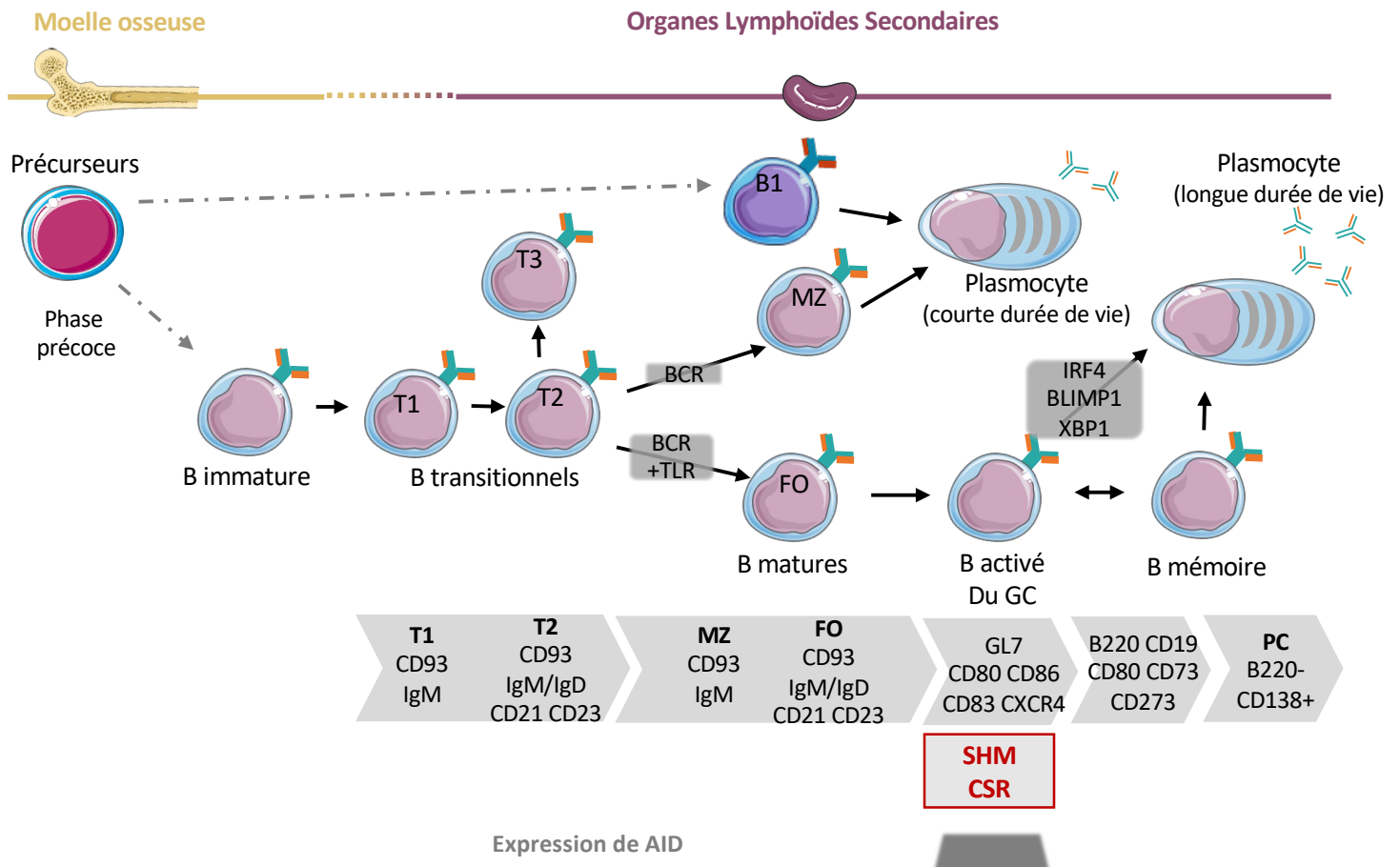
le deuxième réarrangement s'avère non productif, la cellule meurt par apoptose. De ce fait, la recombinaison V(D)J au locus *IgH* est à l'origine de l'élimination de nombreux progéniteurs.

#### **II.1.1.4. Stade pré-B : acquisition du BCR**

Lorsque le précurseur pro-B a terminé son réarrangement VDJ il devient une cellule « grand pré-B » exprimant alors dans son cytoplasme la chaîne lourde ainsi que le pré-BCR à sa surface. Le pré-BCR est constitué de la chaîne lourde préalablement réarrangée assemblée à une pseudo-chaîne légère en association avec l'hétérodimère de signalisation CD79a/CD79b connu aussi sous le nom  $Ig\alpha/Ig\beta$ . La pseudo-chaîne légère est constituée des protéines  $\lambda 5$  et  $V_{pr\beta}$  qui possèdent respectivement des propriétés communes avec la partie constante de la chaîne légère  $\lambda$  et le domaine variable des chaînes légères, permettant leur association avec la chaîne lourde (Mårtensson and Ceredig, 2000; Melchers et al., 1993). Après un essor prolifératif, la cellule « grand pré-B » devient une cellule « petit pré-B » perdant l'expression du marqueur de surface CD43 au profit du gain de l'expression de CD25 (Hardy et al., 2007). Les réarrangements  $V_L-J_L$  s'opèrent à ce stade de façon ordonnée (**Figure 3**). En premier lieu les réarrangements ciblent le loci  $Ig\kappa$ , puis si les réarrangements sont improductifs, les loci  $Ig\lambda$  peuvent être réarrangés toujours sur le premier allèle puis le deuxième (Arakawa et al., 1996). A l'inverse de la recombinaison VDJ au locus *IgH*, de nombreuses cellules passent ce cap de par le nombre de possibilité augmenté (deux loci et deux allèles). La cellule va alors exprimer une chaîne lourde réarrangée unique associée à une unique chaîne légère, de type  $\kappa$  ou  $\lambda$ , réarrangée : c'est l'exclusion isotypique, contrôlée grâce aux réarrangements ordonnés et séquentiels (Arakawa et al., 1996). Cet ensemble chaîne lourde/chaîne légère, en association avec l'hétérodimère  $Ig\alpha/Ig\beta$ , forme le BCR, exprimé à la surface de la cellule et marque le passage au stade B immature (**Figure 3**).

#### **II.1.1.5. Cellule B immature : acquisition de la tolérance centrale**

L'acquisition de l'expression d'une IgM de surface par la cellule s'accompagne de la vérification de la tolérance au soi avant de quitter la moelle osseuse afin de s'assurer que les IgM produites ne reconnaîtront pas d'éléments de l'organisme en tant qu'antigène. Le processus d'acquisition de la tolérance au soi est un processus de sélection négative conduisant à la mort par apoptose des cellules exprimant un BCR auto-réactif. Il existe néanmoins un mécanisme de sauvetage permettant à la cellule de réaliser de nouvelles recombinaison VJ sur les loci des chaînes légères, c'est le processus de « BCR editing » (Halverson et al., 2004). La nouvelle chaîne légère résultante de ces nouveaux réarrangements géniques est exprimée en association avec la chaîne lourde initiale formant ainsi un nouveau BCR qui sera de nouveau testé pour son auto-réactivité. S'il s'avère toujours auto-réactif une anergie ou une mort par apoptose de la cellule est induite (Norvell et al., 1995; Pike et al., 1982).



**Figure 4 : Phase tardive de l'ontogénie B**

Cette figure représente la phase tardive de l'ontogénèse B qui prend place dans les organes lymphoïdes secondaires. Elle récapitule les principaux stades, les marqueurs de surface associés ainsi que les facteurs nécessaires à la progression à travers les différents stade.

Les cellules B non auto-réactives vont alors sortir de la moelle osseuse, emprunter la circulation sanguine et rejoindre les organes lymphoïdes secondaires en périphérie, tels que la rate ou les ganglions, où elles pourront subir des événements de différenciation tardive.

### **II.1.2. Phase tardive du développement : dépendance à l'antigène (Figure 4)**

Les cellules B immatures qui rejoignent les organes lymphoïdes secondaires, comme la rate, les plaques de Peyer ou les ganglions, présentent un phénotype différent des B immatures de la moelle osseuse ainsi que des B matures, ils portent alors le nom de B transitionnels.

#### **II.1.2.1. Cellules B transitionnels : phénotype intermédiaire**

Les B transitionnels sont subdivisés en deux grandes sous-populations : les transitionnels de niveau 1 (T1) et de niveau 2 (T2). Cette subdivision repose sur des marqueurs de surface ainsi que des caractéristiques fonctionnelles différentes (Chung et al., 2002; Su and Rawlings, 2002). En effet, les T1 expriment à leur surface l'IgM mais sont dépourvus d'IgD, de CD21 et de CD23 et présentent un phénotype plutôt immature (Allman et al., 1992). En revanche, les T2 sont caractérisés par l'expression de CD21, CD23 ainsi que l'IgM et l'IgD de surface (Chung et al., 2002; Loder et al., 1999). Les B transitionnels possèdent néanmoins un point commun qui les distingue du compartiment mature : l'expression du marqueur CD93 (**Figure 4**). Lorsque les T1 gagnent la rate dans le manchon péri-artériolaire dans la pulpe blanche, ils sont de nouveau soumis à une sélection : le processus de tolérance périphérique au soi. Leur BCR, et donc leur Ig, est mis à l'épreuve vis-à-vis d'antigènes qui étaient absents dans la moelle osseuse. Les cellules avec un BCR auto-réactif sont alors éliminées par apoptose (Petro et al., 2002). Les cellules qui ne sont pas auto-réactives deviennent des T2 qui migrent vers la zone folliculaire de la rate où les antigènes exogènes sont présentés. Les T2, *via* la stimulation du BCR, deviendront des B matures. Leur devenir en B mature folliculaire (FO) ou de la zone marginale (MZ) est dépendant des signaux qu'ils reçoivent (Niiri and Clark, 2002). Ainsi, un T2 qui sera stimulé uniquement par son BCR deviendra préférentiellement un B MZ alors qu'un T2 qui est stimulé à la fois par son BCR ainsi qu'un autre récepteur, notamment un TLR (Toll Like Receptor) se différenciera majoritairement en B FO (Pillai and Cariappa, 2009). Les TLR sont des récepteurs de l'immunité impliqués dans la reconnaissance de motifs associés à des agents pathogènes bactériens, fongiques ou viraux. En plus des deux sous-populations majeures de B transitionnels T1 et T2, une troisième (les T3), dont l'origine et les fonctions sont moins documentées, est également physiologiquement présente (Allman et al., 2001). Les T3 ressemblent phénotypiquement aux T2 mais expriment faiblement l'IgM. Ce point suggère que les T3 dériveraient de T2 ayant perçu une stimulation antigénique. Les T3 sont aussi décrits comme étant des cellules auto-réactives et anergiques, et de ce fait, ne sont pas à l'origine des B matures (Liubchenko et al., 2012; Merrell et al., 2006; Teague et al., 2007).

### **II.1.2.2. Cellules B matures**

Les B matures sont constitués de trois populations majeures : les cellules B2, folliculaire ou de la zone marginale et les cellules B1 (**Figure 4**).

#### **II.1.2.2.1. Cellules B1**

Les cellules B1, localisées dans les cavités pleurales et péritonéales, sont de deux types, distincts par l'expression du marqueur CD5 : B1-a (CD5<sup>+</sup>) et B1-b (CD5<sup>-</sup>). Ces cellules B1 ont une origine différente des cellules B2. Dérivant du foie fœtal, elles sont les premières cellules B produites au cours du développement et de façon importante tout au long de la vie embryonnaire (Hayakawa et al., 1985). Ces lymphocytes B1 sécrètent des anticorps naturels poly-spécifiques et poly-réactifs et constituent donc une première ligne de défense face à différents pathogènes (Baumgarth, 2017, 2011).

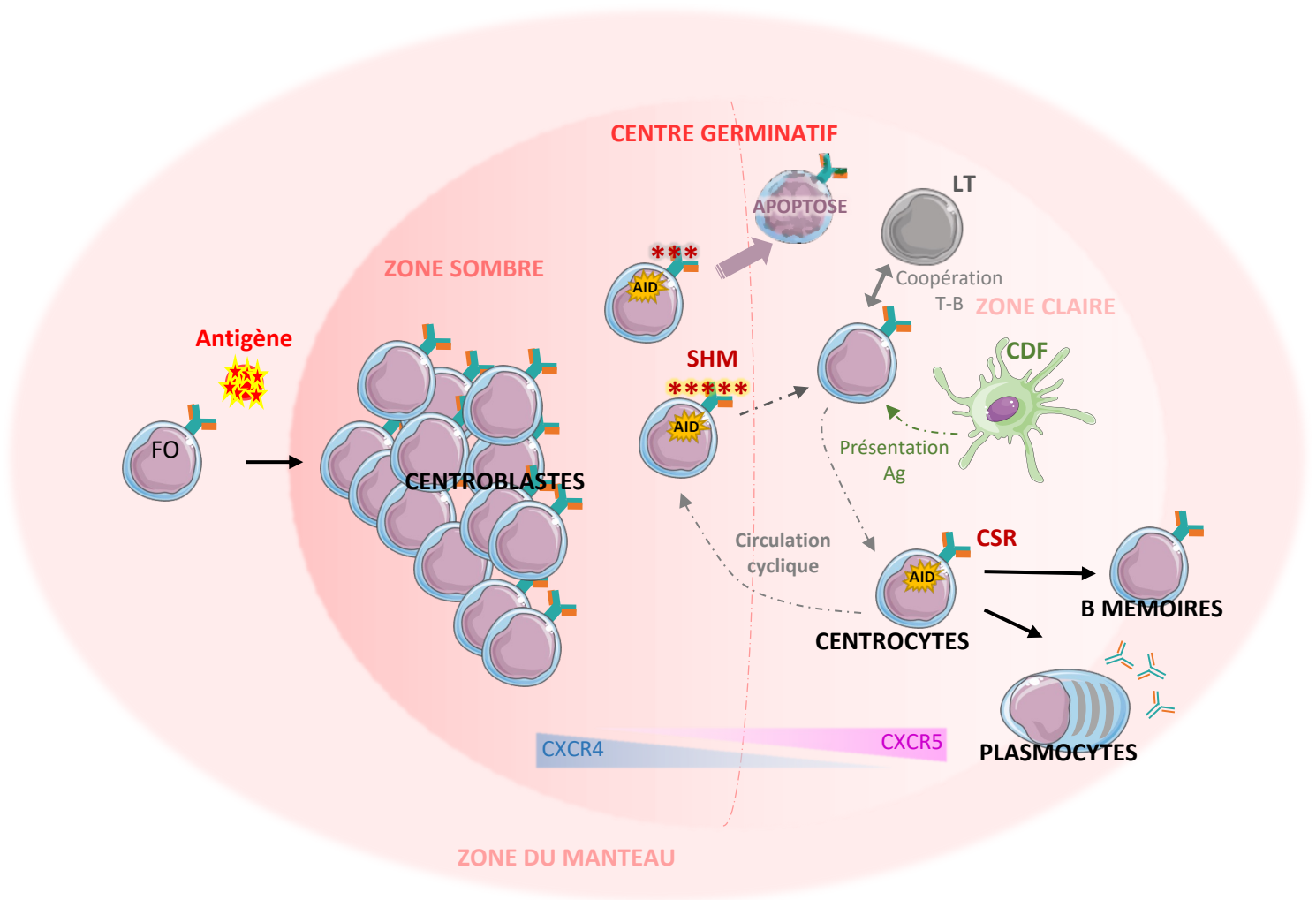
#### **II.1.2.2.2. Cellules B2**

Les cellules B de la zone marginale (MZ) sont, comme leur nom l'indique, localisées dans la zone de jonction entre la pulpe blanche et la pulpe rouge de la rate ou d'autres structures ganglionnaires. Ces cellules, caractérisées par une forte expression d'IgM membranaire et de CD21 associé à une absence de CD23 (Cerutti et al., 2013), possèdent une capacité de réponse rapide mais non spécifique à une stimulation antigénique, grâce à leur BCR poly-réactif capable de reconnaître des motifs communs à plusieurs micro-organismes (Pillai et al., 2005). Ces cellules auraient aussi un rôle primordial dans le transport des antigènes vers les follicules (Cinamon et al., 2008; Ferguson, 2004). Elles n'exercent généralement pas de coopération avec les lymphocytes T contrairement aux cellules B FO. Cependant, il a été montré que les B MZ peuvent aussi initier des réponses à des antigènes T-dépendants (Song and Cerny, 2003).

Les cellules B FO qui sont CD23<sup>high</sup> et CD21<sup>low</sup>, sont les cellules majoritaires des follicules des organes lymphoïdes secondaires. Dans la rate, elles sont localisées dans la pulpe blanche. Ces cellules, suite à un contact avec un antigène et en réponse aux signaux délivrés par les lymphocytes T helper (T<sub>H</sub>), génèrent les follicules lymphoïdes secondaires.

### **II.1.2.3. Activation des cellules B : formation du centre germinatif**

Les cellules B sont activées par un antigène présenté par des cellules présentatrices d'antigène (CPA) présentes à l'intérieur des follicules : les cellules dendritiques folliculaires (CDF). Les cellules B migrent alors vers la zone inter-folliculaire particulièrement riche en lymphocytes T dont certains ont été activés par le même antigène. Un dialogue s'établit alors entre lymphocytes B et T qui vont se fournir des signaux nécessaires à leur activation complète, c'est la coopération B-T. Les lymphocytes T deviennent des Lymphocytes T<sub>H</sub> et migrent vers le centre des follicules. L'ensemble de ces événements conduit à la formation d'une structure particulière constituée majoritairement de cellules B activées :



**Figure 5 : Structure schématique d'un centre germinatif**

Les deux zones (sombre et claire) du centre germinatif sont représentées. Les cellules retrouvées au sein de ces zones, ainsi que les événements génétiques qu'elles subissent sont indiqués. Le devenir des centrocytes et centroblastes est également schématisé ici. Les facteurs majeurs permettant la rétention des cellules au sein d'une zone particulière du GC sont également mentionnés.

le centre germinatif (GC pour *Germinal Center*) (**Figure 5**). Les cellules B non activées par l'antigène sont écartées vers l'extérieur du GC et forment la zone du manteau du GC. Le GC est formé de deux zones distinguées par les cellules qui les composent : une zone sombre (ZS), qui contient de nombreux centroblastes, et une zone claire (ZC) où sont retrouvés les centrocytes.

Les centroblastes sont caractérisés par une prolifération intense et expriment fortement les marqueurs GL7 et CXCR4 et de façon très limitée CD86 et CD83. Ces cellules sont aussi caractérisées par la réduction considérable de l'expression de leurs IgM et IgD de surface (Victoria et al., 2010). Ces cellules subissent le processus d'hypermutation somatique (SHM, mécanisme décrit au paragraphe II.1.3.2.2) dont l'objectif est d'augmenter l'affinité de l'Ig pour l'antigène. La SHM conduit ces cellules à exprimer une Ig membranaire mutée, elles sont alors nommées centrocytes et rejoignent la ZC. Les centrocytes subissent alors une sélection basée sur la capacité de leur Ig de surface à reconnaître l'antigène. Si les cellules B présentent un BCR de meilleure affinité, elles capturent l'antigène présenté par les CDF de façon très efficace. Ces lymphocytes B peuvent ensuite se comporter à leur tour comme une CPA et sont capables de présenter à leur surface des portions d'antigène associées à une molécule du complexe majeur d'histocompatibilité (CMH) de classe II. Les lymphocytes T environnant, dont le récepteur (TCR) est spécifique de cet antigène, vont interagir avec la cellule B présentant l'antigène et lui délivrer les signaux nécessaires à sa survie et à la poursuite de son développement. Au contraire, si l'Ig composant le BCR n'est pas capable d'engager efficacement l'antigène présenté par la CDF, cela signifie que l'Ig n'est pas assez affine ou spécifique de l'antigène. Une cellule B porteuse d'un tel BCR meurt par apoptose. Des cellules ayant été positivement sélectionnées peuvent retourner dans la ZS pour subir un nouveau cycle prolifératif, incluant la SHM, afin d'augmenter d'avantage l'affinité de l'Ig. Cette circulation cyclique des cellules B au sein du GC est gérée notamment par l'expression périodique de CXCR4 et CXCR5 dont les ligands, produits par les cellules stromales, sont respectivement CXCL12 et CXCL13. Ainsi, CXCL12 retient les centroblastes dans la ZS tandis que l'expression progressive de CXCR5 va conduire à la migration des cellules B vers la ZC (Allen et al., 2004).

Les centrocytes présentent une capacité proliférative moins importante, ils expriment eux aussi le marqueur GL7, fortement les marqueurs CD80, CD86 ainsi que l'IgM/IgD de surface, sans exprimer CXCR4 (Victoria et al., 2010). Les centrocytes, exécutent, quant à eux, la recombinaison de classe (CSR) dont le mécanisme sera aussi détaillé dans un prochain paragraphe (II.1.3.2.3).

Plusieurs facteurs sont indispensables à la formation du GC, c'est le cas de la protéine MYC. En effet, sa délétion dans des cellules B activées conduit à une perte de capacité à former les GC (Calado et al., 2012). Néanmoins, le principal régulateur du GC reste BCL6 (B Cell Lymphoma 6). Comme MYC, sa délétion entraîne une perturbation de la formation du GC (Cattoretti et al., 1995). BCL6 est aussi à l'origine de la répression de certains facteurs de différenciation : BCL6 inhibe la transcription de *Pdrn1*,



gène codant la protéine BLIMP-1, facteur clé de la différenciation plasmocytaire (Shaffer et al., 2000; Tunyaplin et al., 2004).

A l'issue de la réaction du GC, deux grands devenir s'offrent à la cellule B : la cellule B mémoire ou le plasmocyte, cellule hautement spécialisée dans la sécrétion d'anticorps (**Figure 5**).

#### II.1.2.4. Différenciation plasmocytaire

La différenciation des cellules B du GC en plasmocytes est majoritairement dictée par le facteur de transcription IRF4 ; ce dernier possédant un rôle ambivalent selon son niveau d'expression. Ainsi, une expression modérée de IRF4 est responsable de l'expression de *Bcl6* et *Pax5*, gènes codant des protéines indispensables à la formation du GC et responsables de l'inhibition du programme de différenciation plasmocytaire (Nera and Lassila, 2006; Tunyaplin et al., 2004). En revanche, l'expression prolongée de IRF4 favorise la transcription de *Prdm1* (Angelin-Duclos et al., 2000; Low et al., 2019). L'expression de BLIMP-1 (codée par *Prdm1*) qui en résulte conduit à l'inhibition de l'expression de *Bcl6*, *Myc* et *Pax5* alors qu'elle favorise l'expression de XBP1, protéine clé de la différenciation en plasmocytes. La transition du programme transcriptionnel B vers un programme transcriptionnel plasmocytaire conduit la cellule B vers un stade intermédiaire, nommé plasmablaste, caractérisé par l'expression concomitante de marqueurs B (BCR, CD45R) et plasmocytaires (CD138) (Oracki et al., 2010; Smith et al., 1996). Le plasmablaste perd ensuite l'expression de marqueurs B et devient plasmocyte (**Figure 4**). Il acquiert alors une morphologie adaptée à sa fonction de cellule sécrétrice d'anticorps, avec un réticulum endoplasmique et un appareil de Golgi de taille importante (Nutt et al., 2015).

Il existe deux types de plasmocytes, les plasmocytes à courte durée de vie et les plasmocytes à longue durée de vie, qui diffèrent par leur origine (**Figure 4**). Les plasmocytes à courte durée de vie sont issus des cellules B1 et des cellules B MZ après leur activation indépendante des lymphocyte T (Balázs et al., 2002; Martin et al., 2001). Ces plasmocytes sécrètent des IgM de faible affinité. Ils peuvent aussi être générés après une activation T-dépendante dans le but de fournir rapidement les premiers anticorps en attendant les anticorps de haute affinité. Les plasmocytes à longue durée de vie sont quant à eux générés à partir des cellules B folliculaires ayant subi une activation T-dépendante dans les GC. Ils sont capables de produire des anticorps très affins et de classe adaptée à l'antigène. Ces plasmocytes peuvent rejoindre la moelle osseuse au sein de niches spécifiques où se trouvent des facteurs essentiels à leur survie (Calame et al., 2003; Tangye, 2011).

De façon concomitante à la différenciation plasmocytaire, les cellules B issues du GC peuvent également devenir des cellules B mémoires.

### II.1.2.5. Cellules B mémoires

Les cellules B mémoires sont générées dès la réponse primaire à un antigène (Kurosaki et al., 2015). Elles circulent dans le sang et séjournent dans les follicules des organes lymphoïdes secondaires. Elles sont caractérisées par une prolifération lente et l'expression d'Ig de surface d'isotype variable. Elles présentent également des marqueurs caractéristiques tels que CD19, CD80, CD73 et CD273 (ou PD-L2) (**Figure 4**). Lors d'un deuxième contact avec l'antigène dont elles sont spécifiques, ces cellules sont capables de se différencier rapidement en plasmocytes ou d'entrer à nouveau dans le GC (Dogan et al., 2009). L'interaction lymphocyte B-lymphocyte T influencerait la différenciation en cellule B mémoire dépendante ou indépendante du GC *via* l'engagement de la molécule CD40 (Taylor et al., 2012). En effet, si le CD40 est fortement engagé lors de la coopération T-B, les cellules B mémoires sont indépendantes du GC. A l'inverse, un faible engagement du CD40 est à l'origine de cellules B mémoires dépendantes du GC. Les cellules B mémoires peuvent aussi être générées en réponse à une stimulation T indépendante à partir de cellules B1 (Yang et al., 2012).

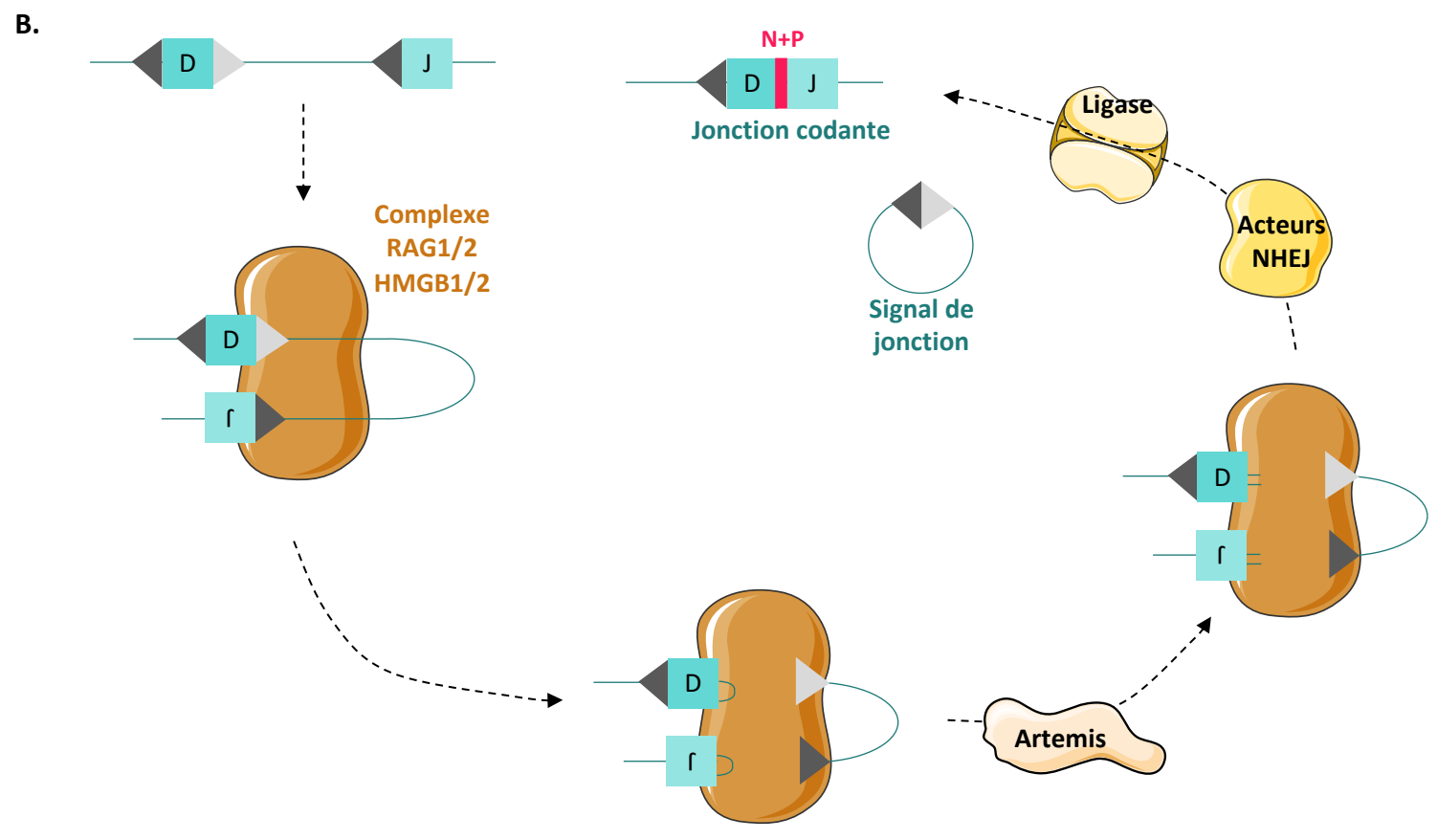
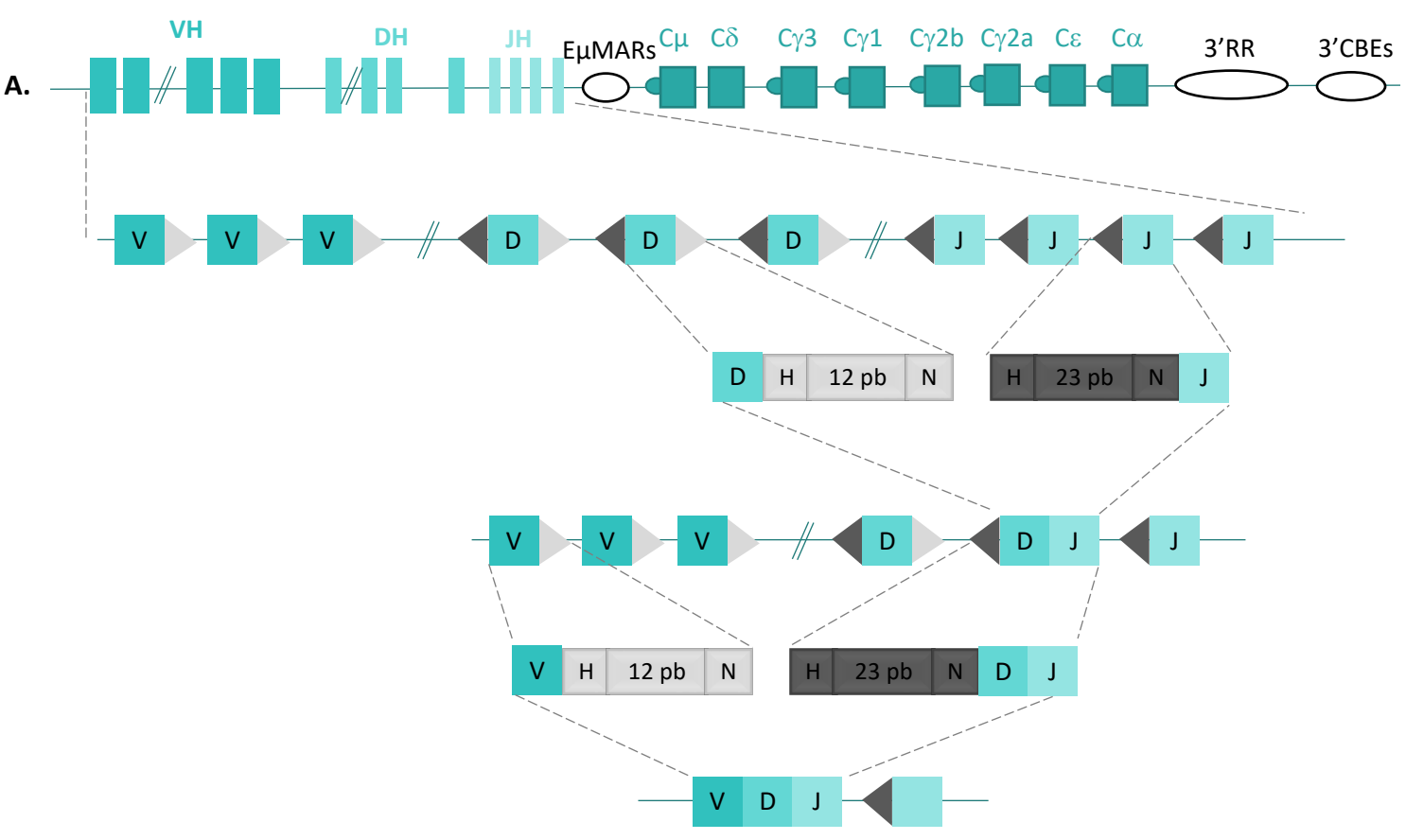
### II.1.3. Remaniements géniques associés au développement B

#### II.1.3.1. La recombinaison V(D)J, évènement précoce

Comme décrit dans le chapitre précédent, en configuration germinale les gènes d'Ig sont regroupés en cluster de façon linéaire : le cluster des segments de Variabilité (V), Diversité (D) et Jonction (J). Afin de former le domaine variable de l'Ig, composé d'un segment VDJ pour la chaîne lourde et d'un segment VJ pour la chaîne légère, les différents segments subissent des réarrangements stochastiques: c'est la recombinaison V(D)J. Ce processus est réalisé en deux étapes hautement ordonnées. Une étape concerne le clivage de l'ADN par le complexe RAG (*Recombination Activating Gene*) composé des recombinases RAG1 et RAG2 (Oettinger et al., 1990; Schatz et al., 1989) tandis que la seconde phase consiste en la réparation de la cassure précédemment induite par les acteurs de la voie de réparation de jonction d'extrémités non homologues (NHEJ pour *Non Homologous End Joining*). Le mécanisme ne sera pas décrit en détail puisqu'il n'a pas été objet d'étude lors de ma thèse.

##### II.1.3.1.1. Mécanisme (**Figure 6**)

Dans leur disposition germinale, les segments à recombinaison, à savoir *V* ; *D* et *J* pour le locus *IgH* et *V* et *J* pour les loci *IgL*, sont flanqués par des séquences spécifiques appelées séquence signal de recombinaison (*RSS* pour *Recombination Signal Sequence*) (Early et al., 1980). Chacune des *RSS* contient des séquences très conservées : un heptamère (5'-CACAGTC-3') et un nonamère (5'-ACAAAACC-3') séparés par un espaceur dont la longueur, mais pas la séquence, est conservée : 12 ou 23 paires de base (pb) (Schatz and Swanson, 2011). Une recombinaison efficace ne peut être réalisée qu'entre des segments ayant respectivement des espaceurs de 12 et 23 pb : c'est la règle 12/23 (Schatz



**Figure 6 : Mécanisme de la recombinaison V(D)J**

- A. Représentation du mécanisme séquentiel de la recombinaison VDJ et de la règle 12/23
- B. Intervention de RAG

and Ji, 2011). Cette règle assure que les bons segments soient recombinaison entre eux. Au locus *IgH*, cela permet la recombinaison entre un segment  $V_H$  et un segment  $D_H$  ou entre un segment  $D_H$  et un segment  $J_H$  et évite l'occurrence de réarrangement direct entre un segment  $V_H$  et un segment  $J_H$  (**Figure 6A**).

Les RSS sont reconnues et liées par les enzymes RAG1 et RAG2, ainsi que les protéines HMGB1 (high mobility group Box 1) et HMGB2 (Fugmann et al., 2000; Shinkai et al., 1992). L'ensemble de ces protéines, formant le complexe recombinase, permet le rapprochement des RSS à recombinaison grâce à la formation d'une synapse (**Figure 6B**) (Lovely et al., 2015; Pavlicek et al., 2008). Le clivage double brin de l'ADN au niveau des RSS est initié par le complexe RAG et génère une structure en épingle à cheveux (McBlane et al., 1995). Les portions d'ADN contenant les RSS sont ligaturées et éliminées sous une forme circulaire, appelée signal de jonction. Quant aux extrémités codantes, elles sont prises en charge par la voie NHEJ. En premier lieu, l'hétérodimère KU, composé de KU70 associée à KU80, va lier les extrémités qui sont toujours maintenues par le complexe recombinase. La protéine KU recrute alors la protéine kinase dépendante de l'ADN (DNA-PK) ainsi qu'Artemis et ATM (Bredemeyer et al., 2006; Malu et al., 2012; Nussenzweig et al., 1996). Artemis, activée par la DNA-PK, clive de façon aléatoire la structure en épingle à cheveux conduisant à la libération d'une extrémité simple brin constituée de nucléotides complémentaires dans la configuration germinale, formant ainsi des courtes séquences palindromiques « P » (Ma et al., 2002). Le caractère hasardeux de la coupure apporte de la diversité à la séquence. Celle-ci est augmentée par l'ajout de nucléotides « N » à l'extrémité des brins à ligaturer par l'enzyme Terminal désoxynucléotidyl transférase (TdT) (Desiderio et al., 1984). Pour finir, la ligase IV est recrutée par le facteur XRCC4 afin de catalyser la réaction de ligature entre les extrémités codantes (Li et al., 1995). L'ensemble de ces événements, effectués de façon aléatoire, conduit à une multitude de réarrangements au locus *IgH* parmi lesquels seulement un tiers est productif et conduit donc à la synthèse d'une chaîne lourde fonctionnelle.

Le déroulement de la recombinaison V(D)J faisant intervenir des cassures de l'ADN suivies de processus de réparation doit être régulé afin d'éviter toute occurrence d'évènement illégitime.

#### **II.1.3.1.2. Régulation**

La régulation de la recombinaison V(D)J s'opère à plusieurs niveaux afin que les réarrangements aux loci des chaînes lourdes et légères se déroulent de manière ordonnée et séquentielle en lien avec le stade de différenciation.

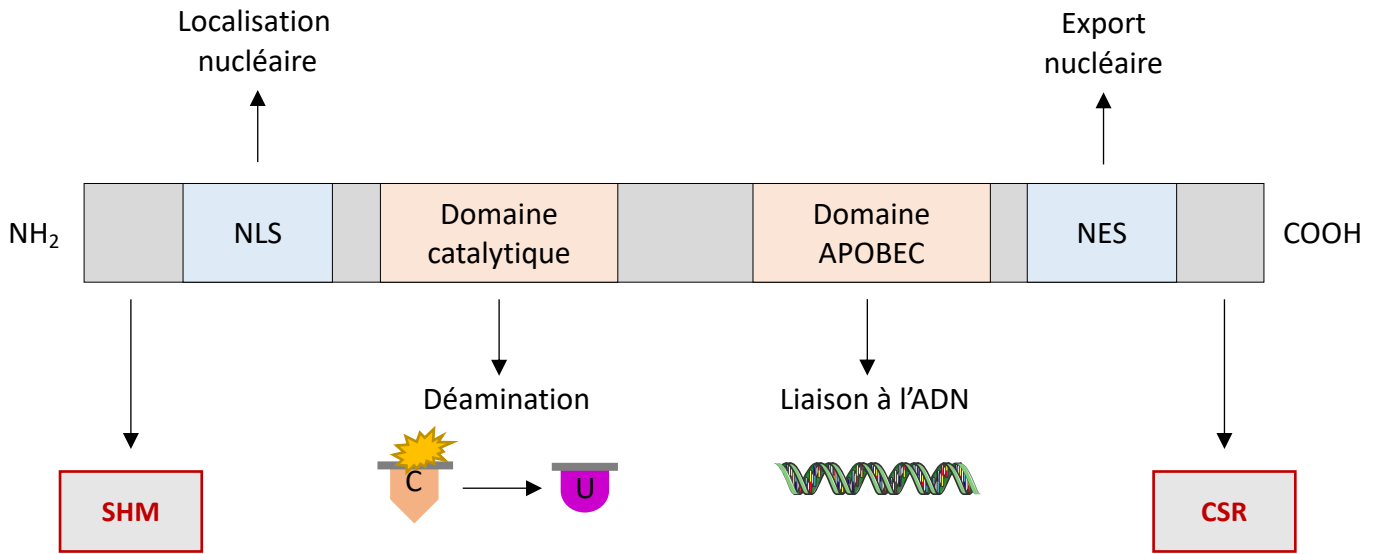
Le premier niveau de régulation s'effectue vis-à-vis des protéines RAG et TdT dont l'expression est limitée aux cellules effectuant physiologiquement la recombinaison V(D)J : les cellules T et B (Oettinger et al., 1990). De plus, l'expression des protéines RAG1 et RAG2 est considérablement diminuée au stade de cellule « grand pré-B » après l'occurrence de l'ensemble des réarrangements V(D)J puis elles

sont de nouveau exprimées au stade de cellule « petit pré-B » pour le réarrangement  $V_L-J_L$  (**Figure 3**) (Grawunder et al., 1995).

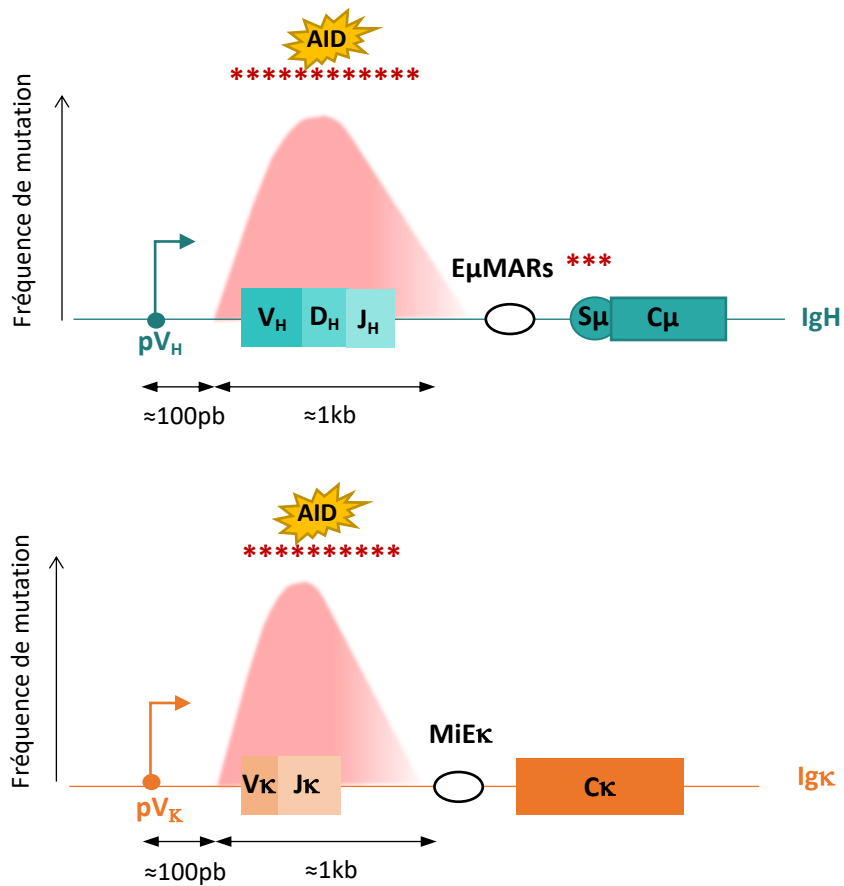
Un autre type de régulation de la recombinaison V(D)J s'opère au niveau de l'accessibilité du locus *IgH* via les marques de chromatine. En effet, les marques de la chromatine sont modifiées en accord avec les étapes ordonnées et séquentielles de la recombinaison V(D)J (Sen and Oltz, 2006). Au stade pré-pro B, au moment où la recombinaison  $D_H-J_H$  s'opère, la chromatine des régions  $D_H$  et  $J_H$  arbore des marques actives alors que la chromatine de la région  $V_H$  est associée, quant à elle, à des marques inactives (Chakraborty et al., 2007; Morshead et al., 2003; Tsukada et al., 1990). Au stade pro-B, lorsque la recombinaison  $D_H$  vers  $J_H$  est terminée, les marques d'histones évoluent. Effectivement, dans la région  $V_H$  les histones sont hyperacétylées, traduisant une accessibilité élevée de la chromatine. En revanche, dans les cellules pré-B, chez lesquelles la recombinaison V(D)J est achevée, les histones de la région  $V_H$  sont hypoacétylées et rendent ainsi la région inaccessible (Chowdhury and Sen, 2003, 2001).

Un niveau supplémentaire de la régulation s'opère au niveau transcriptionnel. Ainsi, préalablement au réarrangement  $D_H-J_H$ , la transcription germinale sens est initiée à partir de l'enhancer intronique  $cE\mu$ , dont le transcrit est appelé  $I\mu$ , ainsi que depuis le promoteur  $pDQ52$  avec le transcrit  $\mu 0$ , qui s'étend vers  $C\mu$  (Alessandrini and Desiderio, 1991; Lennon and Perry, 1985). Des transcrit antisens sont également retrouvés au sein de la région  $D_H$  (Bolland et al., 2007). Après la recombinaison  $D_H-J_H$ , la transcription sens est initiée en amont des segments  $V_H$  à partir de leur promoteur  $pV_H$  (Bolland et al., 2004). Des transcrits antisens sont également initiés à partir de ces promoteurs. La présence de deux sites de fixation de la protéine CTCF au sein de la région intergénique  $V_H-D_H$  empêche la transcription antisens initiée au sein de la région  $D_H$  de se propager à la région  $V_H$  durant la recombinaison  $D_H-J_H$ ; cette région intergénique joue donc un rôle majeur dans la régulation ordonnée des réarrangements au locus *IgH* (Chaumeil and Skok, 2012; Degner et al., 2009; Featherstone et al., 2010). Effectivement, la délétion de la région intergénique conduit à une augmentation de la transcription antisens  $D_H$  qui s'étend alors vers  $V_H$  (Giallourakis et al., 2010). La présence de ces sites *CBEs* indique également qu'il existe un niveau de régulation supplémentaire médié par le mécanisme d'extrusion de boucle dont les principaux acteurs sont CTCF et la cohésine (Degner et al., 2011; Guo et al., 2011). Ce mécanisme sera détaillé dans la partie IV.

Ces éléments pointent la corrélation entre la transcription germinale, les marques de la chromatine et le processus ordonné de recombinaison V(D)J.



**Figure 7 : Représentation schématique de la structure d'AID**  
 Les principaux domaines de la protéine, ainsi que leur fonction, sont représentés.



**Figure 8 : Localisation de l'hypermutation somatique**

L'introduction des mutations au sein des régions variables réarrangées commence à une 100<sup>aine</sup> de pb du site d'initiation de la transcription et s'étend sur environ 1kb. L'action d'AID s'apparente à une gaussienne.

### **II.1.3.2. Hypermutation somatique et commutation de classe : évènements caractéristiques du développement tardif**

#### **II.1.3.2.1. AID, enzyme essentielle**

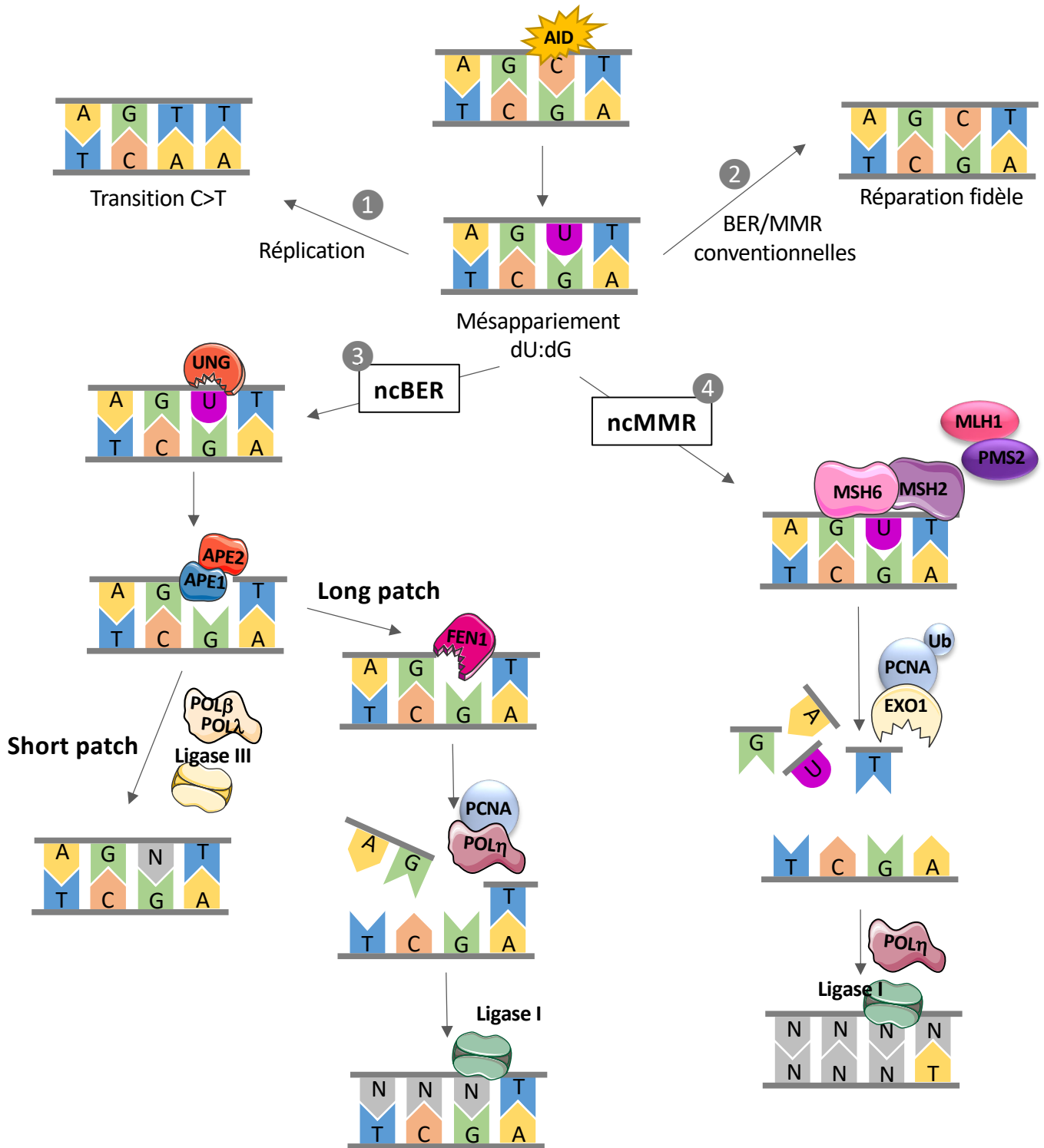
La phase tardive du développement B est caractérisée par l'occurrence de deux évènements géniques au sein du GC : la SHM et la CSR. Ces deux évènements sont initiés par une seule et même enzyme : AID (*Activation Induced Deaminase*) (Muramatsu et al., 2000). Cette protéine, longue de 198 acides aminés et de poids moléculaire de 24 kDa, est codée par le gène *Aicda* localisé sur le chromosome 6 chez la souris. Elle est exprimée uniquement au stade B activé. La protéine est composée de différents domaines : de N- vers C-terminal, AID comprend un signal de localisation nucléaire suivi par son domaine catalytique lui conférant son activité de cytidine déaminase, puis un linker ainsi qu'un domaine APOBEC et pour terminer le signal d'export nucléaire (**Figure 7**).

Il semblerait que les parties N et C terminales d'AID aient des fonctions distinctes. En effet, une mutation unique (G23S, Y13H ou encore V18R) en partie N-terminale conduit à un défaut du processus de SHM mais n'entraîne aucun défaut de CSR (Shinkura et al., 2004). A l'inverse, une délétion d'acides aminés ( $\Delta$ 189-198) en partie C-terminale affecte la CSR alors que le mécanisme de SHM, lui, reste normal (Barreto et al., 2003). AID exerce son activité sur les cytidines présentes au sein de motifs WRCY (où WA/T ; R=A/G ; Y = C/T), préférentiellement dans l'ADN simple brin (Dickerson et al., 2003; Larijani and Martin, 2007; Rogozin and Kolchanov, 1992). La transcription est un prérequis à la déamination de l'ADN par AID. Ainsi, lors de la SHM, la transcription de la région VDJ est nécessaire à l'action d'AID. Pour la CSR, la transcription des régions switch, donneuse et acceptrice(s) est également un préalable indispensable à la déamination des cytidines par AID.

Dans chacun des processus de SHM et de CSR, l'action d'AID est suivie de mécanismes de réparation conduisant respectivement à l'introduction de mutations ou la réparation de cassures doubles brins (CDB).

#### **II.1.3.2.2. Hypermutation somatique**

La SHM est un processus au cours duquel des mutations ponctuelles s'accumulent au sein des régions variables des chaînes lourdes et légères. Cette accumulation est particulièrement importante au niveau des CDRs des segments V qui codent les régions de l'Ig directement en contact avec l'antigène. Comme déjà décrit auparavant, la SHM s'opère dans les centroblastes au sein de la ZS du GC d'organes lymphoïdes secondaires. Ce processus ne fait pas la distinction entre les mutations « favorables » et les mutations « non favorables ». Ainsi, ce mécanisme aléatoire peut mener à la production d'Ig de plus haute affinité, d'affinité diminuée ou encore d'affinité égale (Bemark and Neuberger, 2003). Les mutations introduites peuvent aussi conduire à la production d'une Ig non fonctionnelle, lié par



**Figure 9 : Illustration des différentes prises en charge du mésappariement dU:dG initié par AID**

Le mésappariement dU:dG est initié par la déamination d'une cytosine en uracile, sous l'action d'AID

- 1) Le mésappariement n'est pas détecté, l'uracile est considérée comme une thymine par l'ADN polymérase, ceci engendre une mutation de type transition C>T
- 2) Le mésappariement est détecté et corrigé par les voies conventionnelles de réparation, ne conduisant pas à l'introduction de mutations
- 3) Le mésappariement est pris en charge par la voie BER non conventionnelle avec comme premier acteur UNG, suivi par l'action de polymérase infidèles conduisant à la modification d'1 pb (Short Patch) à 10pb (Long patch) autour du mésappariement.
- 4) La reconnaissance du mésappariement est effectuée par MSH2/MSH6, acteurs de la voie MMR non conventionnelle, suivi par l'action de l'exonucléase EXO1 qui excise entre 10 et 100 pb qui seront comblées par une polymérase infidèle.



exemple à un mauvais repliement de la protéine ou encore à l'insertion d'un codon stop prématuré. Dans ces derniers cas, la cellule est éliminée.

#### *II.1.3.2.2.1. Mécanisme*

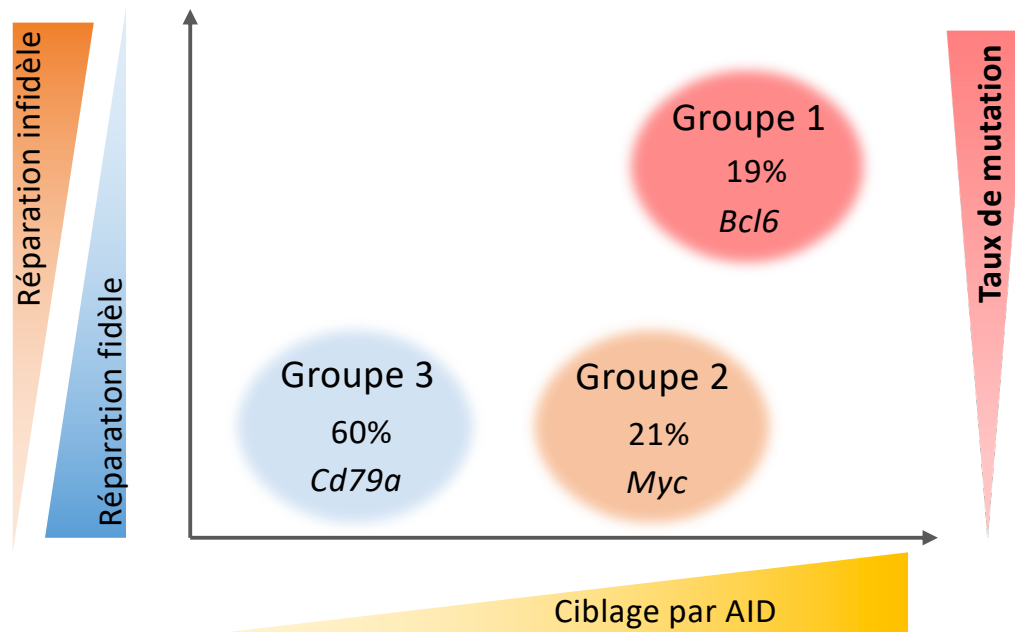
Le mécanisme de SHM est réalisé en deux étapes, une première étape de déamination par l'enzyme AID est suivie par la réparation permettant l'introduction des mutations (Rada et al., 1998). L'activité d'AID s'étale sur une gaussienne d'1kb amorcée à environ 100pb du site d'initiation de la transcription ( $pV_H$  et  $pV_K$ ) (**Figure 8**). Dans un premier temps, durant la phase G1 du cycle cellulaire, l'enzyme AID va déaminer une cytosine donnant lieu alors à sa conversion en uracile. Dans un second temps, le mésappariement dU:dG résultant est pris en charge afin d'être réparé. Cette prise en charge peut s'effectuer de différentes façons. Dans un cas, le mésappariement dU:dG n'est pas détecté, l'ADN polymérase considérera alors l'uracile comme une thymine lors de la réplication (**Figure 9**, voie 1) (Neuberger et al., 2005). Ceci entraîne une mutation de type transition C vers T. Dans un autre cas, le mésappariement dU:dG peut également être pris en charge par les voies de réparation d'excision de base (BER pour *Base Excision Repair*) et de mésappariement (MMR pour *MisMatch Repair*). Les voies BER et MMR ainsi recrutées peuvent être les voies dites « conventionnelles » et ne sont donc pas à l'origine d'introduction de mutations car les mécanismes de réparation associés sont fidèles (**Figure 9**, voie 2). Pendant le processus de SHM, ce sont plutôt les voies dites « non conventionnelles » qui sont recrutées : ces voies font appel à des acteurs « error-prone » ou infidèles et sont donc responsables de l'introduction de mutations.

#### ➤ Voie ncBER (**Figure 9**, voie 3)

Cette voie de réparation est initiée par l'enzyme UNG (Uracile N Glycosylase) qui va exciser l'uracile générée par l'action d'AID, menant à l'apparition d'un site abasique. Ce site dépourvu de base est reconnu par les endonucléases APE1 et APE2 (pour *apuriniques/apirimidiques endonuclease*) qui clivent le brin d'ADN du côté 5' du site abasique (Stavnezer et al., 2014). Ce clivage permet le recrutement de différents acteurs pour réparer l'excision. Ainsi, deux modèles de réparation sont retrouvés : « short patch » et « long patch » selon les acteurs recrutés en aval du clivage (Fortini and Dogliotti, 2007).

Dans le modèle « short patch », un seul nucléotide est remplacé par les ADN polymérases  $\beta$  et  $\lambda$  puis le nucléotide est ligaturé grâce à l'action de la ligase III. Une autre polymérase, REV1, serait aussi recrutée dans ce modèle et favoriserait les transversions C vers G (Di Noia and Neuberger, 2002).

Le modèle « long patch » est caractérisé par l'intervention de FEN1, une endonucléase capable d'enlever entre 2 et 10 nucléotides qui seront remplacés par l'ADN polymérase  $\eta$  dont le recrutement est médié par PCNA (pour *proliferating cell nuclear antigen*) (Fortini and Dogliotti, 2007; Krokan and



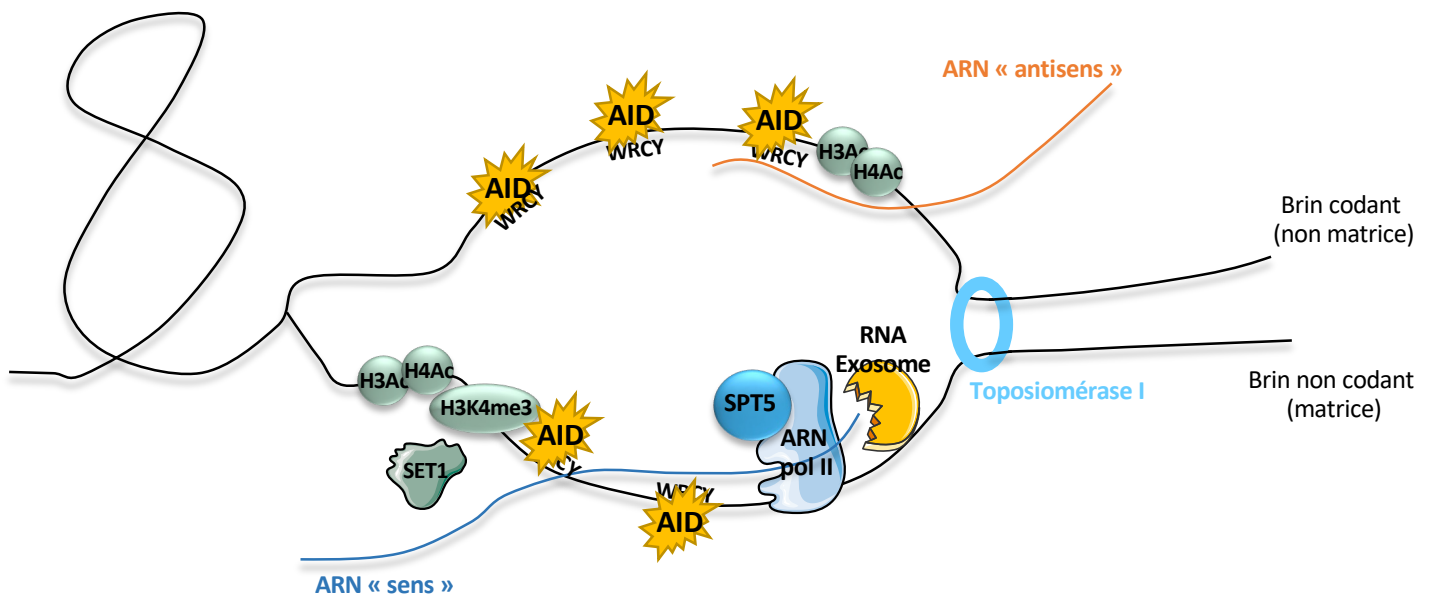
**Figure 10 : Classement des gènes AID « off-target » en trois groupes**

Adapté depuis Liu *et al*, 2008

Le groupe 1 est très ciblé par AID et est réparé de façon infidèle, ce qui induit un fort taux de mutation

Le groupe 2 est également très ciblé par AID mais ne subit pas de réparation infidèle, ce qui conduit à un très faible taux de mutation

Le groupe 3 n'est pas muté car il n'est pas une cible préférentielle d'AID et est réparé de façon fidèle.



**Figure 11 : Régulation du ciblage par AID lors de la SHM**

La régulation du ciblage par AID lors du processus de SHM est complexe et requiert la coopération entre différents acteurs qui sont représentés ici : l'ARN polymérase II, le facteur d'élongation SPT5, la topoisomérase I, l'enzyme AID, la méthyltransférase SET1 et le RNA exosome. La synthèse des ARN « sens » et « antisens » permet d'offrir le substrat d'ADN simple brin à AID.

Bjoras, 2013; Robertson et al., 2009). Les nucléotides ainsi ajoutés sont joints par l'action de l'ADN ligase I, fréquemment en association avec son cofacteur XRCC1, bien que XRCC1 serait un facteur limitant la SHM puisque sa délétion à l'état hétérozygote augmenterait la fréquence d'hypermutation (Saribasak et al., 2011).

➤ Voie ncMMR (Figure 9, voie 4)

Dans cette voie, le mésappariement dU:dG est détecté par le complexe hétérodimérique formé par les protéines MSH2 et MSH6 (MSH pour MutS Homolog). Ce complexe interagit avec MLH1 (MutL Homologie 1) et PMS2 (Postmeiotic segregation 2) au niveau du mésappariement. Cette interaction permet le recrutement de EXO1 associée à PCNA. Dans la voie ncMMR, PCNA est mono-ubiquitinylé sur la lysine 164 et permet le recrutement de polymérases infidèles (Langerak et al., 2007; Roa et al., 2008). De par son activité d'exonucléase, EXO1 excise entre 10 et 100pb qui sont ensuite remplacées par l'ADN polymérase  $\eta$ . Finalement, la ligase I intervient pour lier les extrémités. Cette voie est majoritairement à l'origine de mutations sur les bases A/T (Rada et al., 2004).

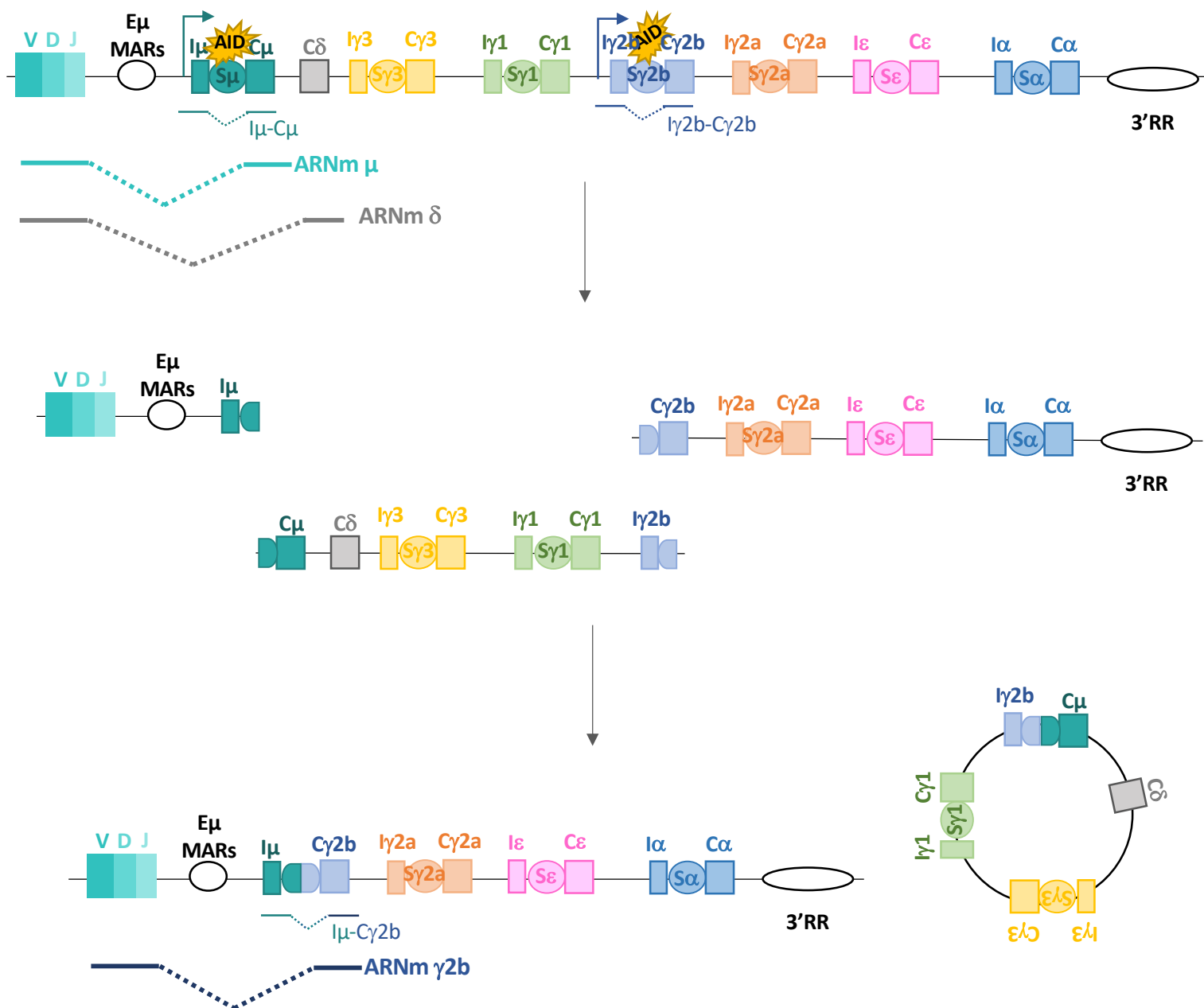
De façon plus générale, lors de la SHM, les mutations sur les bases A/T sont majoritaires et représentent 60% de l'ensemble des mutations alors que les mutations sur les bases G/C représentent 40% ; ce constat indique que les voies BER et MMR non conventionnelles ne contribuent pas de façon équivalente à la SHM (Bahjat and Guikema, 2017; Peled et al., 2008).

#### *II.1.3.2.2.2. Dommages collatéraux*

La SHM est spécifique des loci d'Ig mais, bien que le reste du génome soit « protégé » des mutations, il peut arriver que certains gènes se retrouvent ciblés par AID et subissent alors la SHM (Gordon et al., 2003; Pasqualucci et al., 2001, 1998). Ces gènes sont appelés « AID off targets » et sont classés en trois groupes distincts basés sur la capacité d'AID à les cibler ainsi que sur le type de réparation (fidèle ou infidèle) qu'ils reçoivent (**Figure 10**) (Liu et al., 2008).

- le groupe 1 (19% des gènes) est caractérisé par un ciblage actif par AID suivi de l'action des voies de réparation infidèles, des gènes tels que *Bcl6*, *Cd83* ou encore *Pim1* se trouvent dans ce groupe.
- le groupe 2 (21% des gènes) contient les gènes ciblés par AID mais réparés de façon fidèles, un gène emblématique de ce groupe est l'oncogène *Myc*.
- enfin, le groupe 3 rassemble le reste des gènes dans lesquels des mutations ne sont pas détectées, et ce même en fond déficient pour les voies de réparation de SHM.

La fine régulation du processus de SHM à différents niveaux permet, entre autres, de limiter la survenue de mutations au sein de ces cibles illégitimes.



**Figure 12 : Mécanisme de la recombinaison de classe**

Chaque unité transcriptionnelle des gènes constants, excepté  $\delta$ , est composée de l'exon I, la région switch S et le segment constant C.

Les flèches indiquent le site d'initiation de la transcription. Les transcrits germinaux pré-switch et post switch sont représentés.

La transcription offre le substrat à AID pour exercer son activité de déaminase. Dans les régions S, les nombreux mésappariements induits par AID génèrent des cassures double brins. Les régions switch donneuse ( $S_{\mu}$ ) et acceptrice (ici  $S_{\gamma 2b}$ ) sont ensuite recombinaison ensemble. La portion qui était initialement entre ces deux régions est éliminée sous forme d'épisome.

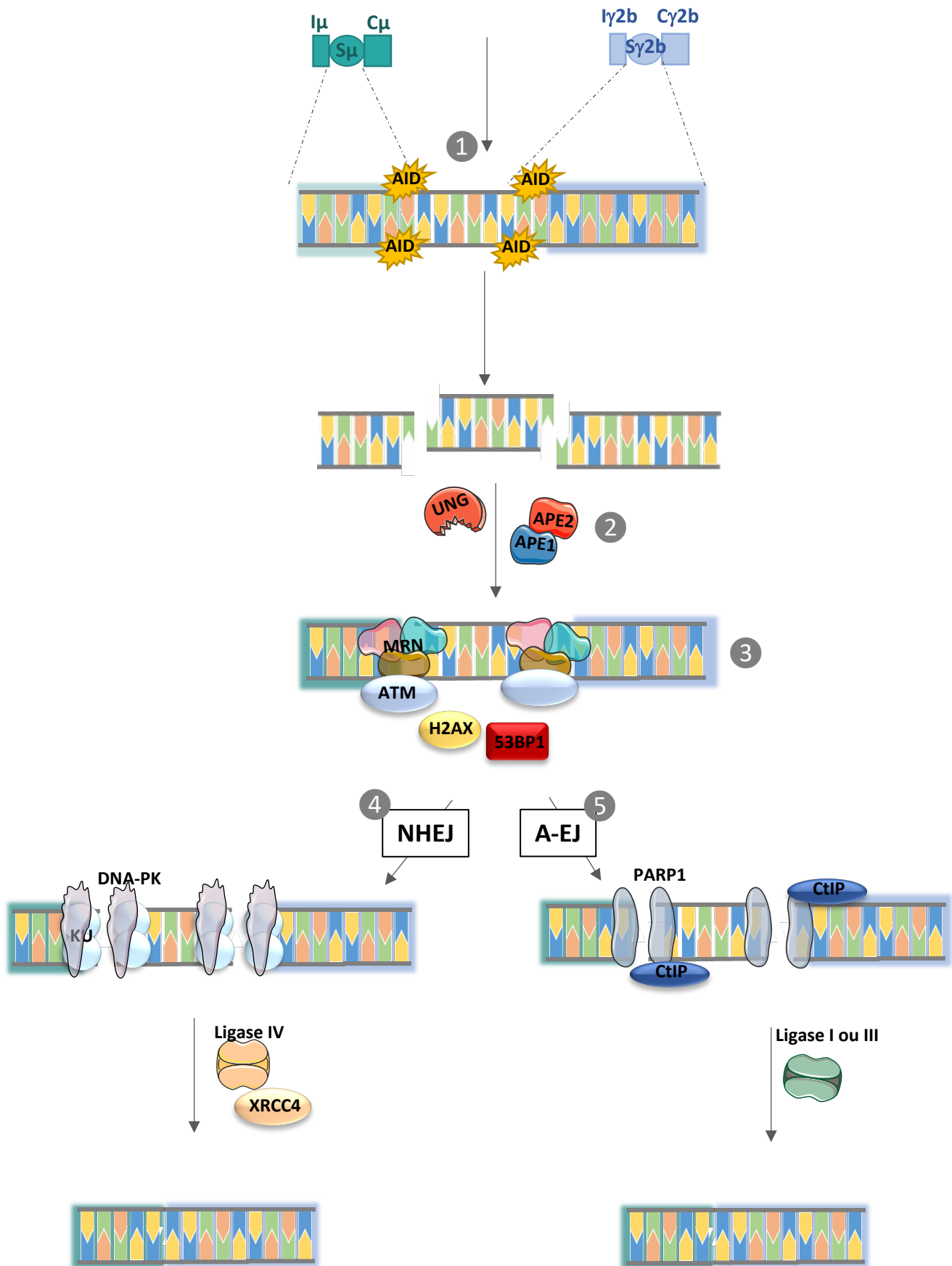
#### II.1.3.2.2.3. Régulation (Figure 11)

Le premier élément régulateur du processus de SHM est la transcription (Fukita et al., 1998). Une hypothèse fréquemment proposée est que la transcription fournit le substrat d'ADN sous forme simple brin, requis pour l'activité de déamination d'AID (Storb, 2014). Lors de la SHM, la transcription est initiée par l'ARN polymérase II au niveau du promoteur du segment V réarrangé, permettant l'exposition du brin non matrice à l'enzyme AID (Bransteitter et al., 2003). Le mécanisme de SHM est également optimisé par l'initiation de la transcription antisens permettant d'exposer le brin matrice d'ADN à l'action d'AID (Perlot et al., 2008). La dégradation de l'ARN nouvellement synthétisé par les acteurs du complexe RNA exosome permet d'augmenter l'efficacité de la SHM. L'hypothèse actuelle suggère que cette dégradation favorise l'exposition du brin matrice (Pefanis et al., 2014; Pefanis and Basu, 2015). L'ARN polymérase II progresse grâce au facteur d'élongation de la transcription SPT5 (*suppressor of ty5 homolog*) qui favorise également le recrutement d'AID (Maul et al., 2014; Nambu et al., 2003; Pavri et al., 2010).

L'ARN polymérase II est également à l'origine de modification d'histone : elle recrute la méthyl transférase SET1 qui est à l'origine de la tri-méthylation de l'histone 3 au niveau de la lysine 4 (H3K4me3). Cette marque persiste après la fin de la transcription et donne ainsi une mémoire moléculaire d'une activité transcriptionnelle récente (Ng et al., 2003). La réduction de la marque H3K4me3 est alors associée à une diminution de la fréquence d'hypermutation somatique. Les gènes « AID off-targets » sont également enrichis de cette marque épigénétique (Kato et al., 2012). De plus, les histones H3 et H4 sont hyperacétylées dans la région V, indiquant un rôle dans le processus de SHM (Sheppard et al., 2018; Woo et al., 2003).

En plus des modifications d'histone, la structure de l'ADN lui-même serait impliquée dans le bon déroulement du processus de SHM. En effet, la formation de structure spécifique, en G-quadruplex, au niveau de séquences enrichies en G, constituerait un socle solide de recrutement pour AID. Bien que le rôle de ces structures en G-quadruplex est surtout décrit pour le mécanisme de CSR, il a été montré dans un modèle murin présentant une délétion du motif de reconnaissance des G-quadruplex au sein de l'enzyme AID une diminution de la fréquence de SHM (Yewdell et al., 2020).

D'autres facteurs sont également impliqués dans le bon déroulement de la SHM, notamment le facteur d'épissage SRF1-3 (Serine/Arginine splicing factor 1-3) qui favorise l'accumulation d'AID dans le noyau et contribue donc à augmenter la fréquence de mutation (Kanehiro et al., 2012; Kawaguchi et al., 2017). SRF1-3 est également capable de recruter la topoisomérase I, une enzyme qui facilite le relâchement des contraintes topologiques de l'ADN (Kumar Singh et al., 2019). Cette topoisomérase génère des cassures temporaires sur un brin et favorise alors l'accès du simple brin à AID (Baranello et al., 2013). Contre toute attente, l'inhibition de la topoisomérase I est à l'origine d'une augmentation de la fréquence de SHM. Ceci pourrait s'expliquer par le fait que l'accumulation des torsions de l'ADN



**Figure 13 : Voies de réparation recrutées au niveau des cassures doubles brins pendant la CSR**

(1) La déamination est initiée par AID au niveau de nombreux sites WRCY au sein des régions switch, conduisant à la génération de cassures doubles brins (CDB). (2) Les CDB sont reconnues par le complexe MRN (MRE11, RAD50 et NSB1) qui recrute ATM, qui, à son tour, recrute H2AX et 53BP1 (3). La réparation des CDB est largement médiée par la voie NHEJ (4) pendant la CSR alors que la voie A-EJ (5) semble recrutée plutôt en second choix.

diminue l'efficacité d'avancement de l'ARN polymérase II, l'ADN simple brin libéré pendant la transcription serait alors accessible plus longtemps à la machinerie mutationnelle (Maul et al., 2015). Enfin, des séquences spécifiques du recrutement de AID au sein des loci d'Ig ont été recherchées. Il a d'abord été montré que les éléments *cis*-régulateurs des loci *Ig*, notamment les promoteurs, ne contiennent pas de motifs nécessaires au recrutement d'AID, ils joueraient d'avantage leur rôle via la transcription. Néanmoins, il a été mis en évidence la présence de séquences DIVAC (*DIVersification ACTivator*) au sein des loci d'Ig (Blagodatski et al., 2009; Kothapalli et al., 2011). Leur présence serait nécessaire à l'introduction de mutations, en effet l'insertion ectopique de séquences DIVAC au niveau de loci qui ne sont normalement pas ciblés par la SHM conduit à une augmentation significative de leur taux de mutations (Senigl et al., 2019a).

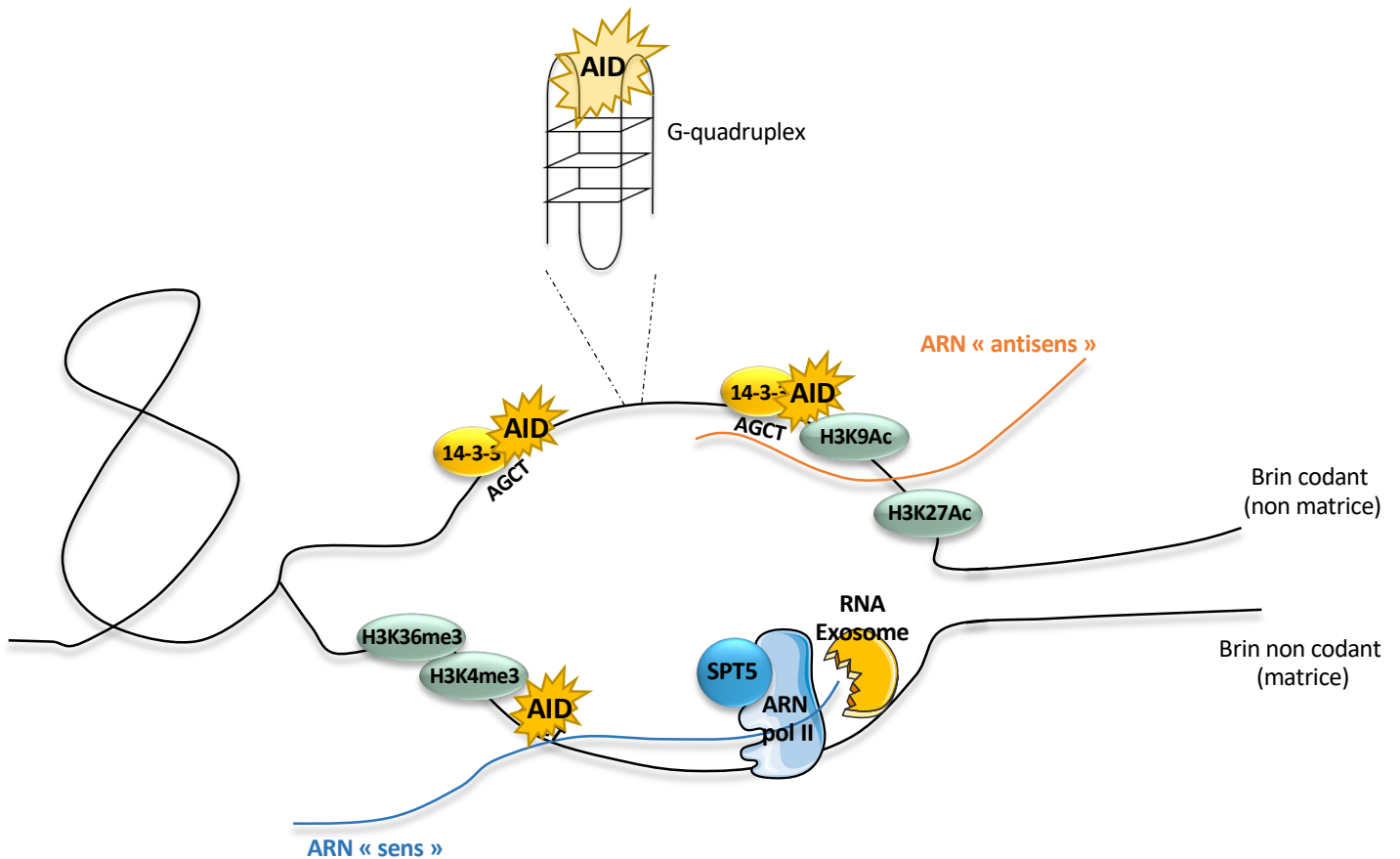
### **II.1.3.2.3. Commutation de classe**

La commutation de classe, également appelée switch isotypique, est un processus initié lui aussi par l'enzyme AID dans les centrocytes de la ZC du GC. Ce mécanisme permet le remplacement de la chaîne constante  $\mu$  par un autre isotype ( $\gamma$ ,  $\varepsilon$  ou  $\alpha$ ) afin d'adapter les fonctions effectrices de l'Ig à l'antigène rencontré (**Figure 12**). Chaque gène constant  $x$ , à l'exception de  $\delta$ , est composé de trois parties distinctes :

- l'exon  $I_x$  (pour intervening), précédé d'un promoteur  $I$
- la région switch  $x$  ( $S_x$ ) dont la taille varie entre 1 et 14kb selon le gène constant et composée de motifs répétés, riche en G, de type GAGCT ou GGGGGT
- le gène constant  $C_x$

#### *II.1.3.2.3.1. Mécanisme*

Comme lors de l'hypermutation somatique, l'enzyme AID va déaminer une cytosine en uracile. Contrairement à la SHM, la CSR nécessite l'introduction de CDB au sein des régions *S* (Stavnezer and Schrader, 2014). La présence abondante de motifs WRCY répétés au sein des régions *S* permet à AID d'initier son activité catalytique à plusieurs endroits très proches les uns des autres. Cette particularité permet la génération de cassures simples brins suffisamment proches sur les deux brins d'ADN pour former des CDBs (Han et al., 2011; Stavnezer and Schrader, 2014; Xu et al., 2012). L'intervention d'UNG contribue à la formation des CDBs, notamment par l'excision des U générés et le recrutement sous-jacent des endonucléases APE. Les CDBs ainsi formées sont reconnues par le complexe MRN formé des protéines MRE11, RAD50 et NSB1. Le complexe MRN active les voies de réparation des CDBs de l'ADN, notamment en recrutant ATM (*ataxia telangiectasia mutated*). Cette kinase recrute à son tour des acteurs impliqués dans la prise en charge des CDBs : H2AX et 53BP1 (Lee and Paull, 2007). La présence



**Figure 14 : Régulation du recrutement de AID aux régions switch pendant la CSR**

Le recrutement de AID au niveau des régions switch dépend de la présence de plusieurs éléments tels que l'ARN Pol II et le facteur SPT5 qui permettent la synthèse d'ARN libérant, qui, en coopération avec le RNA exosome, permet la libération du substrat ADN simple brin pour AID. La présence de motifs AGCT au sein des régions switch permet le recrutement de la protéine adaptatrice 14-3-3, responsables du recrutement de AID. La présence de motifs répétés riches en G permet la formation de structure typique appelée G-quadruplex qui forment une plateforme de recrutement d'AID bien plus efficient que de l'ADN linéaire.



de l'ensemble de ces acteurs au niveau des CDBs stabilise les extrémités. La région donneuse  $S\mu$  et la région acceptrice  $Sx$  sont rapprochées afin d'être ligaturées et le fragment d'ADN initialement compris entre les deux régions  $S$  est excisé (**Figure 12**). La prise en charge des CDBs peut être effectuée par deux voies : la jonction d'extrémité non homologue (NHEJ) ou la jonction d'extrémité alternative (A-EJ) (**Figure 13**) (Boboila et al., 2012). Lors de la CSR, la voie majoritairement empruntée est la voie NHEJ, la voie A-EJ serait plutôt une voie de second choix, utilisée lorsque les acteurs NHEJ sont absents. Comme pour la recombinaison VDJ, le complexe KU, formé de KU70 et KU80, maintient les extrémités des CDBs et recrute la DNA-PK. Finalement, les deux régions  $S$  sont jointes par l'action de la ligase IV, en coopération avec son cofacteur XRCC4 (Lieber, 2010). Comme cité précédemment, la voie A-EJ peut également être recrutée afin de clôturer le processus de CSR. Cette voie fait appel à des acteurs différents. Ainsi, les CDBs sont reconnues par PARP1 (*Poly-ADP Ribose polymerase 1*) puis CtIP (*Ct-interacting protein*) est recrutée et finalement la jonction entre les extrémités libres est réalisée par la ligase I ou la ligase III (Boboila et al., 2010; Hwang et al., 2015).

#### II.1.3.2.3.2. Régulation (Figure 14)

Comme dit précédemment, AID intervient sur l'ADN simple brin, rendu accessible par la transcription. La production de transcrits germinaux constitue donc un prérequis majeur pour la CSR : ces transcrits sont initiés au promoteur  $I\mu$  ainsi qu'au promoteur  $Ix$ . Cependant, ces transcrits germinaux ne sont pas régulés de façon équivalente : la transcription de  $S\mu$  est constitutive, toujours présente, alors que la transcription de la région  $Sx$  acceptrice apparaît en réponse à un stimulus. Par exemple, *in vitro* la transcription de  $S\gamma3$  et  $S\gamma2b$  apparaît en réponse au lipopolysaccharide (LPS) bactérien alors que la transcription à  $S\gamma1$  et  $S\epsilon$  est initiée suite à un stimulus par l'IL-4 (l'isotype est sélectionné en fonction de la concentration en IL-4). Bien que ces transcrits ne codent pas de protéines, ils sont essentiels à la CSR. En effet, leur association au brin matrice, formant ainsi des hybrides ADN-ARN aussi appelés R-loop, optimise l'accès du brin non matrice à l'activité catalytique d'AID notamment en favorisant la pause du complexe ARN polymérase II -SPT5 responsable notamment du recrutement d'AID (Pavri and Nussenzweig, 2011; Yu et al., 2003; Zhang et al., 2014). L'accessibilité du brin matrice est quant à elle favorisée par l'intervention du RNA exosome qui dégrade les transcrits néosynthétisés (Basu et al., 2011; Laffleur et al., 2017).

Les modifications épigénétiques au niveau des histones semblent aussi jouer un rôle important, notamment pour l'accessibilité des régions  $S$  à la machinerie transcriptionnelle et donc dans l'initiation de la transcription. En effet, des marques actives, telle que l'acétylation des lysines 9 et 27 de l'histone 3 (H3K9Ac et H3K27Ac) ainsi que la tri-méthylation des lysines 4 et 36 (H3K4me3 et H3K36me3) au niveau de l'histone 3 sont particulièrement enrichies au niveau de la région  $S\mu$  (Sheppard et al., 2018; Wang et al., 2009; Wu et al., 2018).

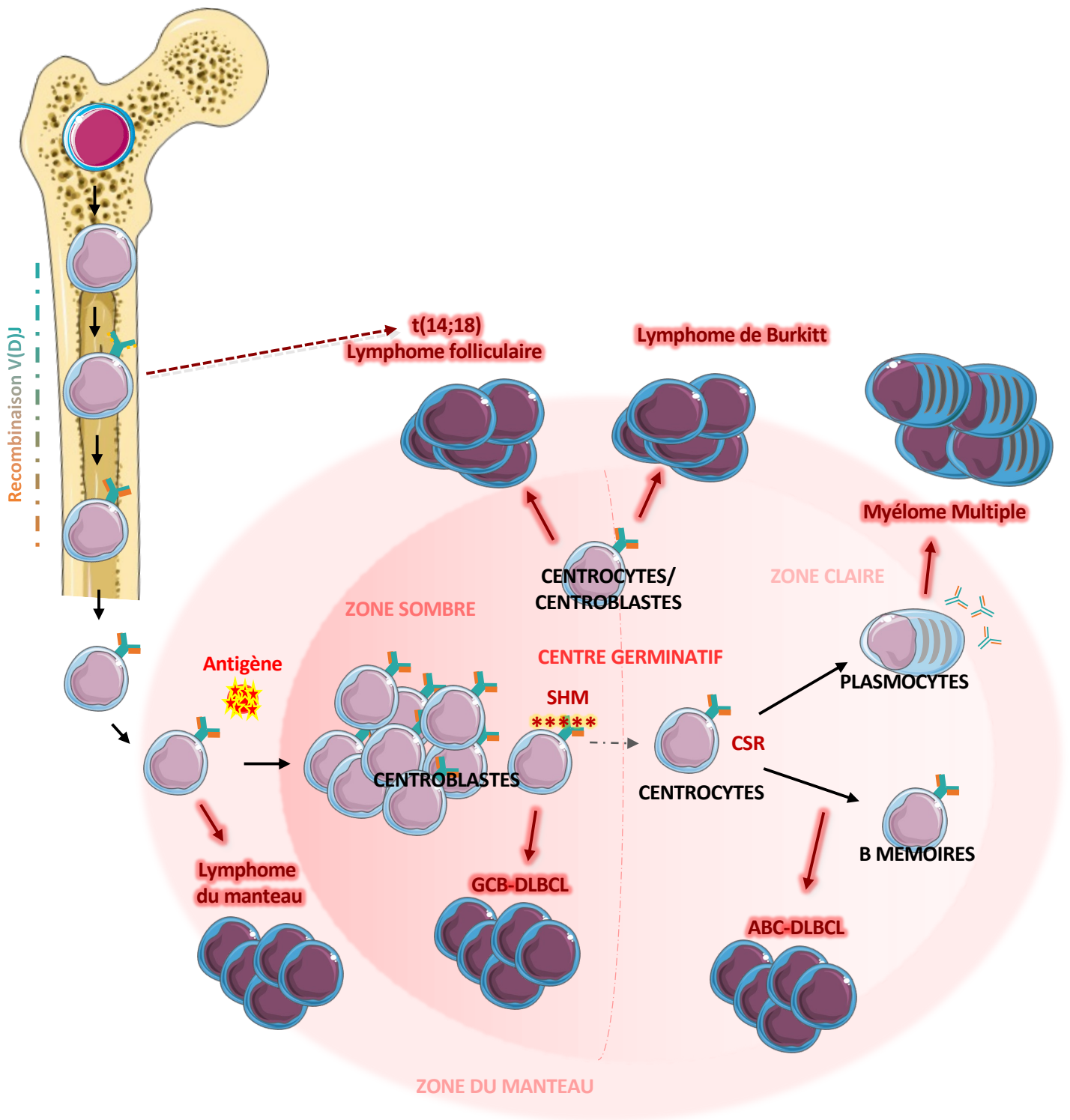


Figure 15 : Représentation des contreparties physiologiques des principaux lymphomes non hodgkiniens et du myélome multiple

Les différents lymphomes abordés dans cette partie et le myélome multiple sont représentés avec leur contrepartie cellulaire physiologique.

Enfin, la séquence des régions *S* en elle-même est importante pour le recrutement d'AID. En effet, la présence de motifs répétés particulièrement riche en G au sein de ces régions permet la formation de structure appelée G-quadruplex. Ces structures, très largement retrouvées au niveau des régions *S*, sont à l'origine d'un recrutement plus efficace de l'enzyme AID, en comparaison à une structure linéaire, et favoriseraient la CSR (Pucella and Chaudhuri, 2017; Qiao et al., 2017). De plus, les régions *S* sont aussi très riches en séquences cible d'AID, notamment le motif AGCT qui est reconnu par la protéine adaptatrice 14-3-3 qui est en partie responsable du recrutement d'AID (Lam et al., 2017; Xu et al., 2010). En outre, les séquences répétées au sein des régions *S* contribuent à stopper la progression de l'ARN polymérase II, favorisant ainsi le recrutement d'AID en quantité importante par le complexe formé de l'ARN polymérase II associée à SPT5 (Pavri et al., 2010).

La structure du locus *IgH* constitue également un niveau de régulation clé de la CSR. En effet, la topologie tridimensionnelle du locus, médiée par le mécanisme d'extrusion de boucle, a été montrée comme nécessaire à la recombinaison entre la région donneuse  $S\mu$  et la région acceptrice  $Sx$  (X. Zhang et al., 2019). De fait, le processus d'extrusion de boucle (décrit dans la partie IV), permet le rapprochement des régions *S* à recombiner dans une même synapse et l'élimination du segment génique intermédiaire à ces régions switch sous la forme d'épisome (Wuerffel et al., 2007).

## II.2. Lymphomagenèse B

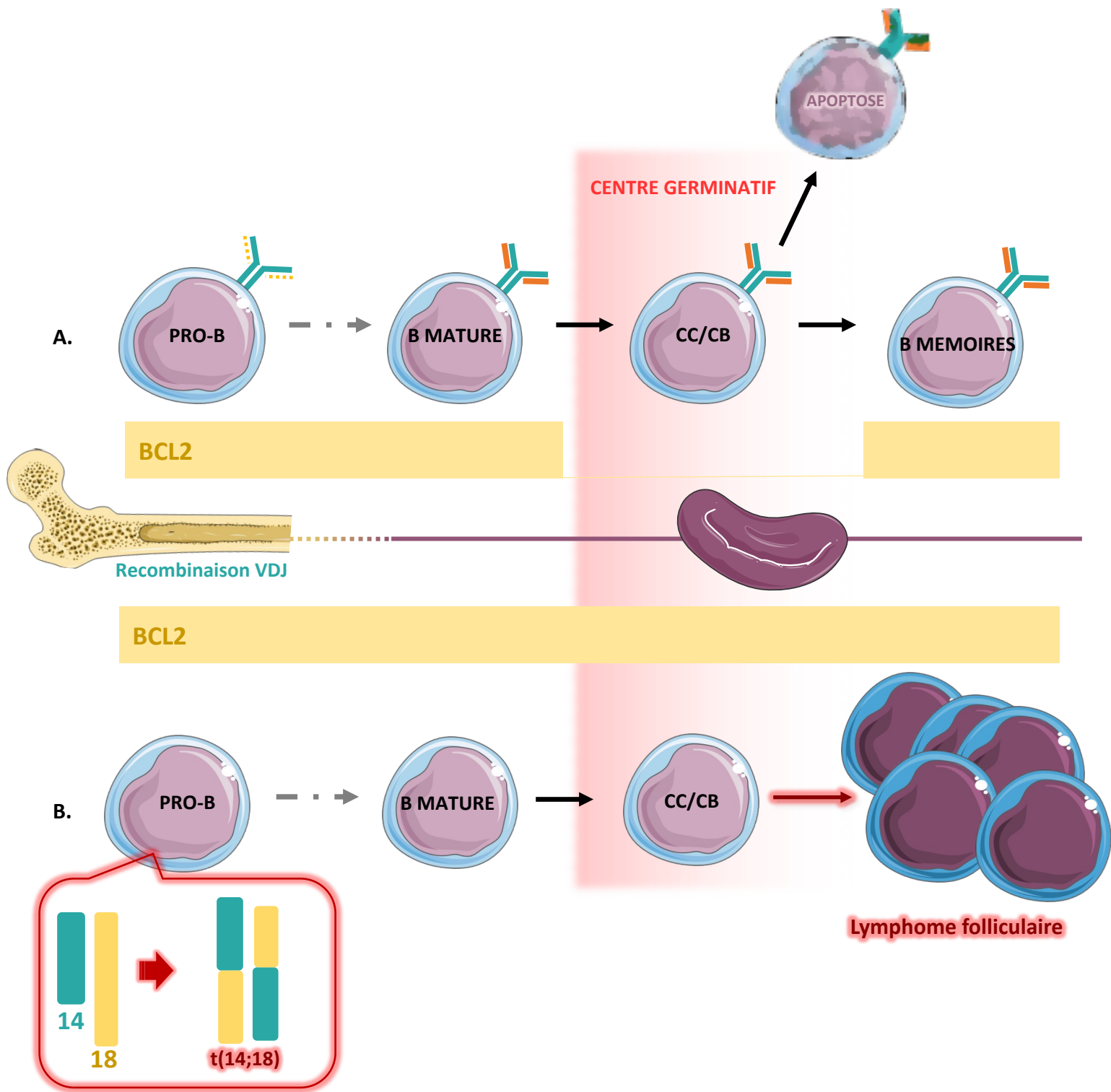
### II.2.1. Généralités

Les lymphomes se réfèrent à des types de cancers qui affectent les cellules B, T ou NK, qui sont toutes des cellules lymphoïdes. La classification des lymphomes, selon l'organisation mondiale de la santé (OMS), est établie selon plusieurs critères hématologiques, biologiques et génétiques (Swerdlow et al., 2016).

La classification de l'OMS répartit les lymphomes B en deux grands groupes distincts : les lymphomes de Hodgkin et les lymphomes non hodgkiniens (LNH). Les lymphomes de Hodgkin se distinguent par la présence de cellules de Reed-Sternberg qui sont des cellules anormalement géantes envahissant les tissus lymphoïdes affectés. Les LNH sont, quant à eux, les lymphomes les plus fréquents puisqu'ils représentent entre 80% et 90% des cas de lymphomes.

Les LNH sont distingués selon la vitesse de prolifération des cellules malignes et peuvent alors être indolents ou au contraire agressifs. Toutefois, cette distinction n'est pas exclusive puisqu'un LNH indolent peut évoluer vers une forme agressive.

La contrepartie physiologique des cellules cancéreuses intervient aussi dans le classement des lymphomes B (**Figure 15**). De même, le statut mutationnel des gènes d'Ig et l'isotype d'Ig porté par la cellule lymphomateuse peut donner une information sur la cellule « d'origine ». Ainsi, avec ces



**Figure 16 : Translocation pendant la recombinaison VDJ, exemple de la translocation t(14;18) dans le lymphome folliculaire**

La translocation t(14;18) impliquant IgH et Bcl2 se produit pendant la recombinaison VDJ et conduit à une résistance à l'apoptose des cellules B dans le centre germinatif. Cet évènement est caractéristique du lymphome folliculaire.

**A.** Développement et pattern d'expression de *Bcl2* normaux

**B.** Conséquence de la translocation t(14;18)

CB : Centroblaste ; CC : Centrocyte

différentes informations, il est admis que près de 80% des LNH dérivent de cellules B du GC ou post-GC. Parmi les exemples de LNH qui dérivent de cellules B GC ou post-GC, peuvent être cités le lymphome folliculaire qui est le LNH indolent le plus fréquent ou encore le lymphome diffus à grande cellules B (DLBCL) dont le sous-type GCB (pour *Germinal Center B cell*) dérive des centrocytes et le sous-type ABC (pour *Activated B cell*) de cellules sorties du GC, le plus courant des LNH agressif.

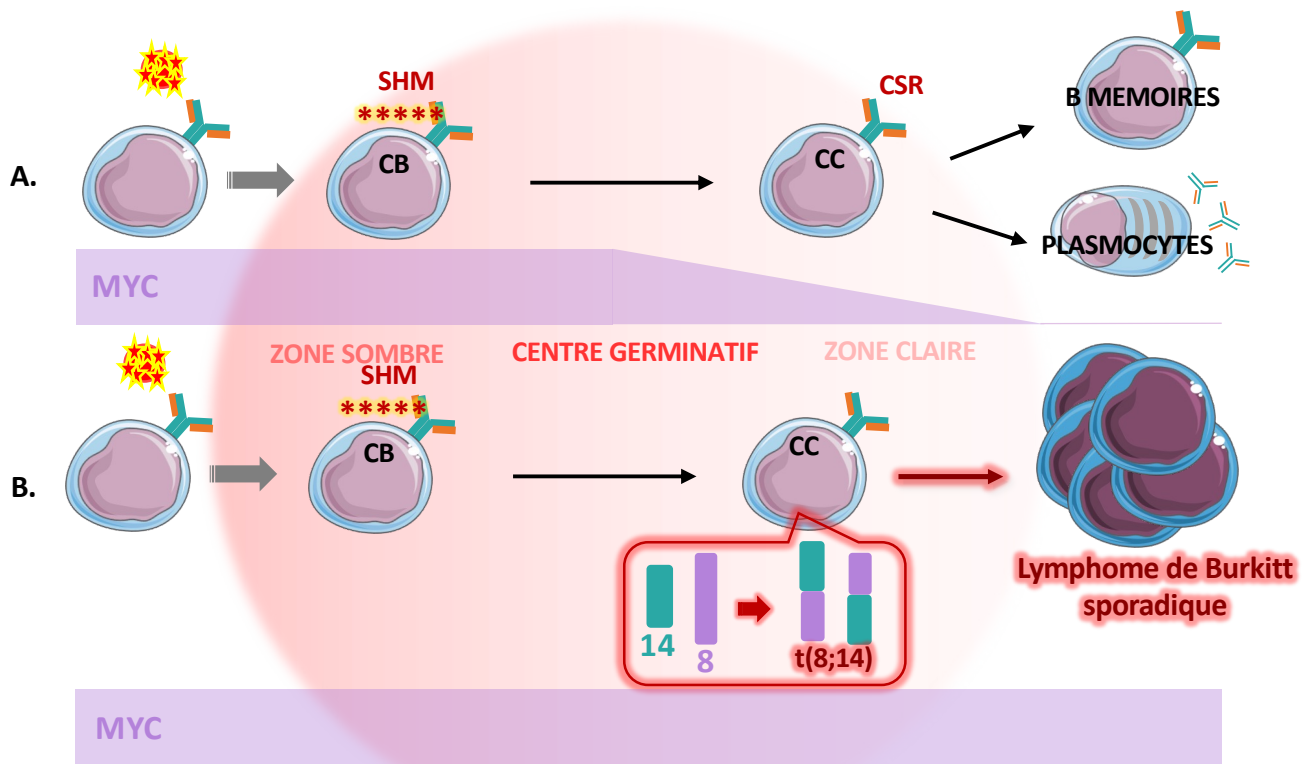
### II.2.2. Altérations liées à la maturation des gènes d'Ig

Le développement physiologique des cellules B est articulé autour de trois grandes étapes de maturation des gènes d'Ig : la recombinaison V(D)J, la SHM et la CSR, permettant *in fine* l'expression d'une Ig de haute spécificité et affinité. L'occurrence de chacun de ces événements est régulée à plusieurs niveaux *via* des modifications génétiques et épigénétiques ainsi que par des communications entre la cellule B et l'environnement cellulaire dans lequel elle se trouve (moelle osseuse, organes lymphoïdes secondaires). Toute dérégulation d'un de ces processus peut alors initier la lymphomagenèse B. Seuls les événements géniques liés à la recombinaison V(D)J, la SHM et la CSR seront abordés ici.

#### II.2.2.1. Évènements de translocation pendant la recombinaison V(D)J

Un exemple typique d'altération génique lors de la recombinaison V(D)J est la translocation t(14 ;18) impliquant le locus *IGH*, localisé sur le chromosome 14, et l'oncogène *BCL2* présent sur le chromosome 18 chez l'Homme. L'insertion de l'oncogène entre les segments  $D_H$  et  $J_H$  ainsi que la présence de résidus « N » témoignent de l'occurrence de cette translocation pendant la recombinaison VDJ même si une autre étude suggère un dysfonctionnement de la SHM comme pouvant être responsable de cette translocation (Bakhshi et al., 1987; Nadel et al., 2001; Tsujimoto et al., 1985). Cette translocation, place l'oncogène *BCL2* sous le contrôle des enhancers forts du locus *IGH*, conduisant à la surexpression du produit du gène *BCL2*. Bien que cet événement se produise au cours de la phase précoce de l'ontogenèse B, il constitue un événement primaire fréquent au développement de lymphome folliculaire dont les cellules sont issues du GC. Ceci s'explique par le rôle physiologique que tient la protéine BCL2 dans les cellules B. En effet, en condition normale, l'expression de *BCL2* varie au cours du développement B pour être négativement régulée dans les cellules B du GC, et ce afin de permettre l'apoptose des cellules ayant un BCR moins affin (**Figure 16A**). Lorsque les cellules possèdent la translocation t(14 ;18), la répression de BCL2 ne s'opère pas, les cellules deviennent alors résistantes à l'apoptose et prolifèrent de façon anormale (**Figure 16B**) (Shen et al., 2004).

Un autre exemple est celui de la translocation t(11 ;14), retrouvée dans le lymphome du manteau, qui juxtapose le gène codant la cycline D1 (*CCND1*) avec le locus *IGH*. Le point de cassure dans le locus *IGH* est principalement localisé entre les segments  $D_H$  et  $J_H$ , suggérant là aussi la survenue de cet événement pendant la recombinaison VDJ. La surexpression de la cycline D1 résultant de cette translocation



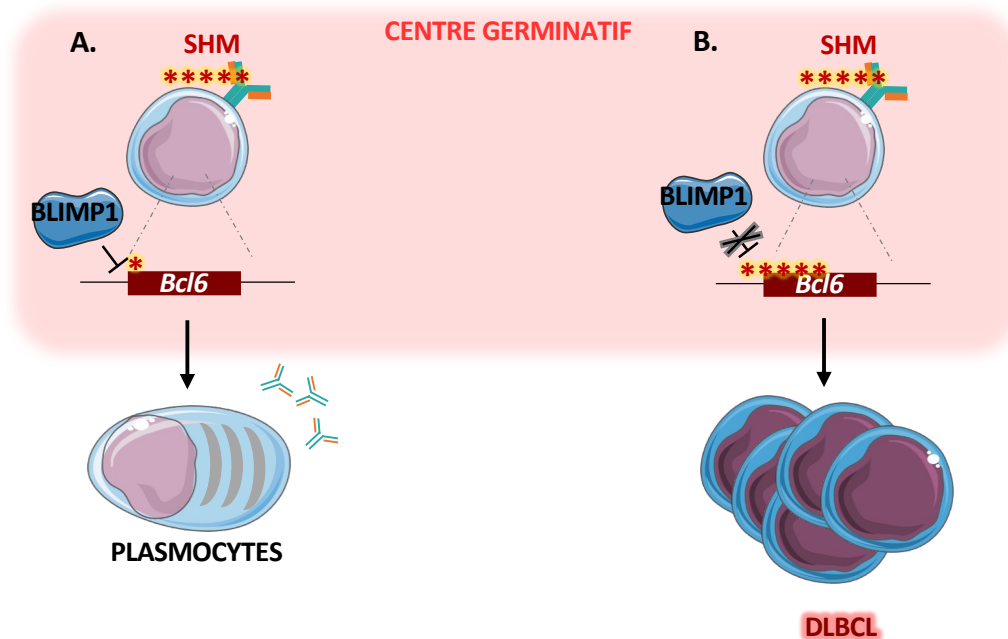
**Figure 17 : Translocation pendant la CSR, exemple de la translocation t(8;14) dans le lymphome de Burkitt**

La translocation t(8;14) impliquant Myc et IgH se produit pendant la CSR et conduit à une prolifération intense des cellules B dans le centre germinatif. Cet évènement est caractéristique de la forme sporadique du lymphome de Burkitt.

**A.** Développement et pattern d'expression de *Myc* dans des cellules sans la translocation t(8;14)

**B.** Conséquence de la translocation t(8;14)

CB : Centroblaste ; CC : Centrocyte



**Figure 18 : Mutation d'oncogènes pendant le processus de SHM, exemple de *Bcl6* dans le DLBCL**

L'introduction de mutation dans la partie 5' de *Bcl6* empêche l'inhibition de son expression par BLIMP1 et conduit à une prolifération intense. *Bcl6* est muté dans 75% des DLBCL.

**A.** Contexte physiologique

**B.** Conséquence des mutations dans la partie 5' de *Bcl6*

conduit à une dérégulation du cycle cellulaire et par conséquent une prolifération anormale des cellules (Stamatopoulos et al., 1999; Welzel et al., 2001).

### **II.2.2.2. Altérations géniques lors du développement tardif**

De la même façon que lors de la recombinaison V(D)J, les remaniements géniques tardifs peuvent aussi être à l'origine d'altérations géniques initiant ou participant à la lymphomagenèse.

#### **II.2.2.2.1. Associées à la CSR**

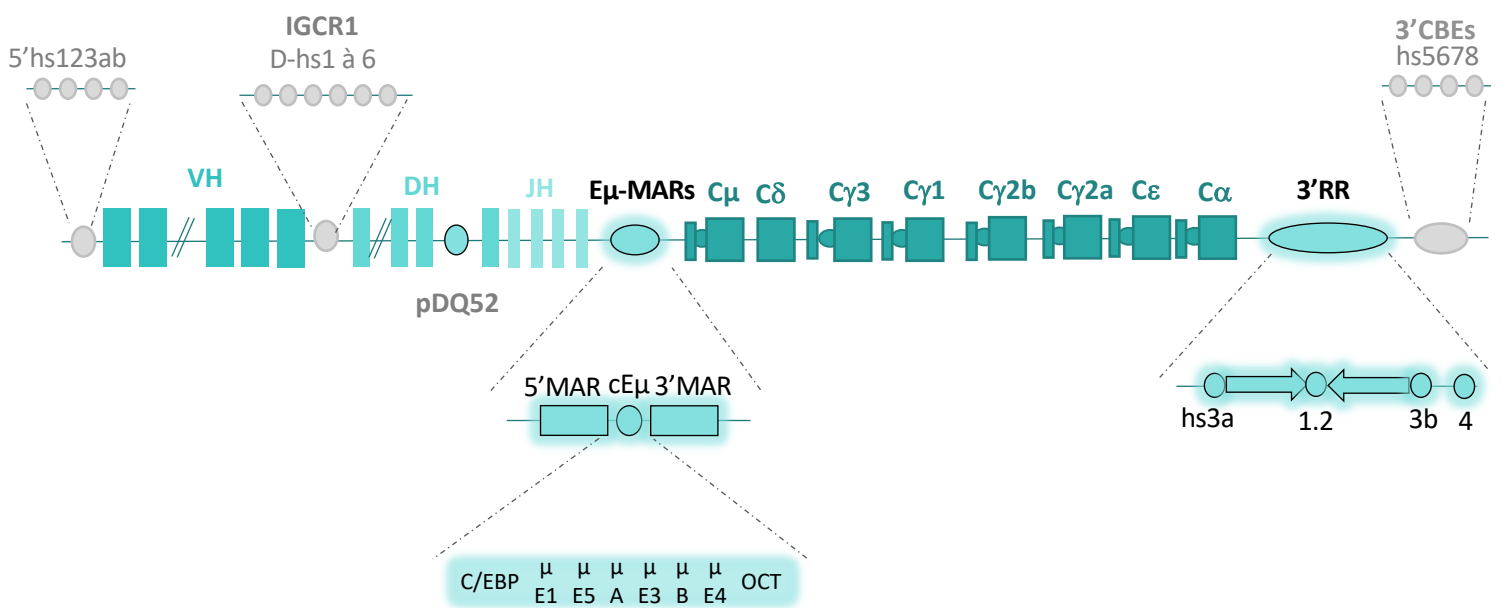
Du fait de l'introduction de CDB, la CSR constitue, au même titre que la recombinaison VDJ, un facteur de risque de translocation. Un exemple typique est la translocation t(8 ;14) dans le lymphome de Burkitt sporadique (Basso and Dalla-Favera, 2015; Schmitz et al., 2012). Cette translocation place l'oncogène *c-MYC* sous le contrôle des enhancers (*cE $\mu$*  et les enhancers de la 3'RR) du locus *IGH* (Truffinet et al., 2007; Wang and Boxer, 2005). Ceci conduit à son expression ectopique dans les cellules du GC qui n'expriment normalement ce gène que sur une très courte durée, conférant ainsi, aux cellules un haut pouvoir prolifératif. Cette translocation constitue un évènement primaire dans le cas du lymphome de Burkitt sporadique. (**Figure 17**).

De même, dans le myélome multiple, cancer des plasmocytes, de nombreuses translocations ont été identifiées au niveau des régions switch du locus *IGH* (Bergsagel et al., 1996). Certaines de ces translocations sont « silencieuses » en analyse cytogénétique conventionnelle, notamment la translocation t(14 ;16) qui implique *c-MAF* (Chesi et al., 1998).

#### **II.2.2.2.2. Associées à la SHM**

Durant la SHM, deux types d'altérations peuvent se produire : des translocations du fait de l'induction de cassure de l'ADN et l'introduction de mutations au sein de gènes non Ig. Parmi ces gènes, l'une des cibles pouvant donner lieu à un état pathologique est *BCL6* (Pasqualucci et al., 1998; Shen et al., 1998). En effet, dans environ 75% des cas de DLBCL, tout sous-types confondus, la partie 5' du gène *BCL6* présente un taux anormalement élevé de mutation (Iqbal et al., 2007). En conditions physiologiques, *BCL6* est exprimé dans le GC puis réprimé par le facteur clé de la différenciation plasmocytaire : BLIMP-1 (**Figure 18A**). Les mutations anormalement introduites en 5' vont altérer les sites de fixation de BLIMP1, ce qui entraîne l'expression continue de *BCL6* (Bal et al., 2022). Une survie et une prolifération intense des cellules résultent de cette non répression, pouvant ainsi conduire à la lymphomagenèse (**Figure 18B**).

Les évènements géniques inhérents au développement de la cellule B doivent donc être strictement régulés. Cette régulation est en partie médiée par la transcription, elle-même contrôlée par les éléments *cis*-régulateurs présents aux loci codant les Ig. Le prochain chapitre décrit les éléments *cis*-régulateurs du locus *IgH* et leurs fonctions.



**Figure 19 : Éléments *cis*-régulateurs du locus IgH**

Cette figure positionne les différents éléments *cis*-régulateurs mentionnés dans le chapitre III. Elle détaille aussi leur composition.

Les régions isolatrices sont représentées en gris et les promoteurs/enhancers en bleu. Les flèches correspondent au quasi-palindrome de la 3'RR.



## Partie III. Éléments *cis*-régulateurs du locus *IgH*

---

Plusieurs éléments *cis*-régulateurs jonchent le locus *IgH* (**Figure 19**). Ces éléments sont essentiels au bon déroulement du développement de la cellule B, compte tenu des réarrangements géniques opérés tout au long de celui-ci. Comme cité dans la partie I, ces éléments sont de différentes natures et exercent donc des rôles différents. Les régions isolatrices, les promoteurs et les activateurs transcriptionnels (aussi appelés enhancers) sont alors distingués.

### III.1. Régions isolatrices et promoteurs

#### III.1.1. Régions isolatrices

De part et d'autre du locus sont retrouvées des régions isolatrices, chacune composée de quatre sites *hs*. Ces deux régions permettent ainsi de cantonner au sein du locus *IgH* les événements géniques inhérents au développement B. Du côté 5' du locus, les éléments *5'hs123ab* permettent principalement de protéger les gènes adjacents des événements de recombinaison VDJ alors que du côté 3', les éléments *hs5678*, communément appelée *3'CBE*, ont pour fonction de limiter la localisation des événements de CSR au locus *IgH* (Pawlitzky et al., 2006; Volpi et al., 2012; Zhang et al., 2021). Plus récemment, il a été démontré que cette région participe également à la recombinaison VDJ, en permettant d'établir le centre de recombinaison où RAG effectue son « balayage » (Liang et al., 2023). Une autre région de type isolatrice, appelée région de contrôle intergénique (*IGCR1*), localisée entre les segments  $V_H$  et  $D_H$ , joue un rôle prépondérant dans la régulation des événements de la recombinaison VDJ : cette région limite la progression des transcrits germinaux antisens vers les segments  $V_H$  lors des premiers réarrangements  $D_H$  vers  $J_H$  et participe à l'établissement du centre de recombinaison (Guo et al., 2011; Liang et al., 2023).

#### III.1.2. Promoteurs

##### III.1.2.1. Promoteurs $pV_H$

De nombreux promoteurs sont aussi présents au sein du locus. Situés à 250pb en amont des gènes  $V_H$ , se trouvent les promoteurs nommés  $pV_H$  (Buchanan et al., 1997). Ces derniers jouent, eux aussi, un rôle important dans l'initiation des réarrangements  $V_H-DJ_H$  puisqu'à l'origine des transcrits germinaux régulateurs de la région  $V_H$ . Plus spécifiquement, le  $pV_H$  du segment  $V_H$  réarrangé conserve un rôle important durant toute la vie de la cellule B, depuis le stade pré-B, où la chaîne lourde est réarrangée jusqu'au stade ultime plasmocyte, puisqu'il est à l'origine du transcrit de la chaîne lourde fonctionnelle. Les  $pV_H$  contiennent des séquences spécifiques à l'origine de leurs rôles prépondérants. Premièrement, la majorité des  $pV_H$  contiennent une « TATA Box » localisée à environ 30 pb en amont du site d'initiation de la transcription ainsi qu'un élément *Inr* qui contribue à l'initiation de la

transcription. En amont de cette TATA Box, se situe un octamère (ATGCAAT), site de liaison de différents facteurs de transcription (Johnston et al., 2006; Parslow et al., 1984; Roy et al., 2011). Enfin, les  $pV_H$  les plus actifs contiennent des éléments supplémentaires : *DICE* (pour Downstream Immunoglobulin Control Element) et *PAIR* (pour Pax-Activated Intergenic Repeat), capables de recruter des facteurs clés comme E2A et PAX5 (Ebert et al., 2011; Tantin et al., 2004).

### III.1.2.2. Promoteur $pDQ52$

Un autre promoteur important du locus *IgH* est  $pDQ52$ , situé en amont du dernier gène  $D_H$  (Alessandrini and Desiderio, 1991). C'est le seul promoteur actif localisé en amont d'un segment  $D_H$  avec un rôle notoire pour la régulation des événements de recombinaison  $D_H-J_H$ , il est alors qualifié de promoteur/enhancer (Kottmann et al., 1994). Ce promoteur est à l'origine de la genèse du transcrit  $\mu 0$  qui s'étend donc depuis  $DQ52$  jusqu'à  $C\mu$  en passant par l'ensemble des segments  $J_H$ , ce qui les rend accessibles au complexe recombinaise impliqué dans la recombinaison VDJ.

### III.1.2.3. Promoteurs I

Les promoteurs I constituent également des régulateurs spécifiques du locus *IgH*, ils sont localisés en amont de l'unité formée par la succession  $I_x - S_x - C_x$ , pour chaque gène constant excepté  $C\delta$ . Ces promoteurs sont à l'origine des transcrits germinaux essentiels à l'action d'AID pendant la CSR (Nambu et al., 2003).

Pour résumer, ces éléments permettent à la cellule B de contrôler les événements de remaniements géniques qu'elle opère au cours de sa différenciation. Deux autres régions régulatrices majeures sont impliquées dans la régulation de ces processus, il s'agit des régions  $E\mu$ MARs et 3'RR. Mon travail de thèse s'est particulièrement articulé autour de ces deux régions régulatrices majeures, leurs rôles respectifs sont décrits en détails ci-après.

## III.2. La région $E\mu$ MARs

Localisée entre le dernier segment  $J_{H4}$  et  $S\mu$ , la région  $E\mu$ MARs occupe une position centrale au sein du locus *IgH* et possède une activité spécifique au lignage B. Cette région est composée de l'enhancer core  $E\mu$  ( $cE\mu$ ), de 220 pb, flanqué par deux MARs. La MAR située en 5' occupe 310 pb et celle en 3' s'étend sur 350 pb.

### III.2.1. L'enhancer $cE\mu$

L'enhancer  $cE\mu$  comporte plusieurs sites de fixation pour plusieurs facteurs de transcription. Ces sites de fixation sont ordonnés de la façon suivante, selon le sens 5'-3' :  $C/EBP$ ,  $\mu E1$ ,  $\mu E5$ ,  $\mu A$ ,  $\mu E3$ ,  $\mu B$ ,  $\mu E4$  ainsi qu'une séquence octamérique similaire à celle retrouvée dans les  $pV_H$  (Ernst and Smale, 1995).

Tous ces éléments constituent des sites de liaisons pour différents facteurs tels que PU.1, E2A ou ceux de la famille OCT ainsi que la protéine ZEB. Il a été montré que la fixation de E2A, dans des cellules B, sur  $\mu E4$  et  $\mu E5$  est activatrice du locus *IgH*, alors que dans des cellules non-B ces sites sont occupés par la protéine ZEB dans le but d'inhiber l'expression d'*IgH* (Genetta et al., 1994).

Plusieurs modèles différents, d'insertion, de remplacement ou de délétion, ont mis en évidence l'implication de l'enhancer *cE $\mu$*  dans la recombinaison VDJ et par conséquent son rôle dans le développement précoce. Une première étude a ainsi montré que l'insertion de *cE $\mu$*  dans un transgène initie des recombinaisons  $D_H-J_H$  dans des cellules lymphoïdes (Ferrier et al., 1990). Un peu plus tard, le remplacement de *cE $\mu$*  par un oligonucléotide de 20 pb dans des souris chimériques montre un blocage partiel de la recombinaison  $V_H-DJ_H$  (Serwe and Sablitzky, 1993). L'importance de *cE $\mu$*  a été confirmée par son remplacement chez la souris par une cassette pGK-Néo<sup>R</sup>, modèle dans lequel une diminution drastique des recombinaisons  $V_H-DJ_H$  et  $D_H-J_H$  est observée (Chen et al., 1993; Perlot et al., 2005; Sakai et al., 1999). Enfin les animaux KO pour *cE $\mu$*  (*cE $\mu$*  KO) arborent un défaut majeur de recombinaison VDJ, entraînant un blocage important des cellules au stade pro-B, avec pour conséquence une diminution significative du nombre de cellules B en périphérie (Perlot et al., 2005). Les effets observés dans ce modèle s'expliquent par la diminution des transcrits germinaux  $\mu 0$  et  $1\mu$ , normalement initiés avant les premiers réarrangements  $D_H-J_H$  (Perlot et al., 2005). Dans un modèle similaire combinant la délétion de *cE $\mu$*  à celle de *pDQ52*, les auteurs observent également une diminution de la transcription antisens de la région incluant les segments  $D_H$  et  $J_H$  (Bolland et al., 2007). *A contrario*, la transcription germinale sens et antisens au niveau des segments  $V_H$  n'est pas affectée par la délétion de *cE $\mu$* , cet effet corrobore le fait que les réarrangements les plus impactés sont ceux de  $D_H$  vers  $J_H$  et non ceux de  $V_H$  vers  $DJ_H$ . L'effet de la suppression de *cE $\mu$*  sur les marques de la chromatine étaye également l'accessibilité à la transcription. Les marques actives (H3K9Ac), normalement présentes au sein des régions à recombiner, sont absentes dans des modèles de délétion de *cE $\mu$*  alors que les marques répressives (H3K9me2), elles, restent présentes (Chowdhury and Sen, 2001). Dans ces modèles soulignant la fonction régulatrice de *cE $\mu$* , les cellules B parvenant à réaliser un réarrangement  $D_H-J_H$  ne rencontrent pas de problème dans la suite de leur maturation.

Enfin, le modèle de délétion totale de la région *E $\mu$ MARs* (*cE $\mu$*  + *MARs*, appelée *E $\mu$ FL KO* pour *E $\mu$  Full Length KO*), initialement développé dans des souris chimériques (Sakai et al., 1999) puis étudié en détail au laboratoire dans un modèle KO, vient compléter les données précédemment décrites. En effet, ce modèle confirme le défaut de recombinaison VDJ, sans pour autant présenter de biais d'utilisation du répertoire  $V_H$ . Le développement B est altéré par un blocage important des cellules au stade pro-B, avec pour conséquence, là encore, une diminution du nombre de cellules B en périphérie (Marquet et al., 2014). Dans ce modèle, la transcription de la chaîne lourde  $\mu$  est également diminuée aux stades précoces de développement dans la moelle osseuse. La conséquence majeure est une

diminution de l'expression du pré-BCR au stade pré-B qui limite l'expansion clonale caractéristique de ce stade ; ce déficit de prolifération contribue à diminuer significativement le nombre de cellules B en périphérie (Marquet et al., 2014).

En revanche, la délétion de *cEμ* n'affecte plus la transcription de la chaîne lourde d'Ig dans les cellules B matures ; elle ne compromet pas les transcrits germinaux des régions S et par conséquent n'entraîne pas de défaut significatif de CSR (Li et al., 2010; Perlot et al., 2005). En ce qui concerne la SHM, le rôle de *cEμ* a fait l'objet de controverses. En effet, la délétion de l'ensemble de la région *EμMARS* dans des hybridomes ne provoque pas de défaut de SHM alors que la délétion seule de *cEμ* ou des *MARS* entraîne une diminution de la fréquence de mutation (Ronai et al., 2005). Ces observations ne corroborent pas avec les observations réalisées dans des modèles murins *cEμ* KO ou *EμMARS* KO (associées ou non à un allèle *IgH* préalablement réarrangé) dans lesquels la SHM n'est que faiblement diminuée (Li et al., 2010; Perlot et al., 2005). Ces données suggèrent que *cEμ* ne joue pas de rôle prépondérant dans la SHM.

### III.2.2. Les régions d'attachement à la matrice (MAR)

Comme décrit précédemment, l'enhancer *cEμ* est cerné par deux régions *MARS*, appelées *MARS<sub>Eμ</sub>*. Les *MARS* sont des régions d'ADN de 200 à 1000pb, riches en bases A/T, réparties dans le génome. Leur localisation est souvent télomérique ou intronique, proches d'enhancers et/ou promoteurs (Blasquez et al., 1989). Ces régions constituent des éléments régulateurs importants, notamment par leur composition en sites de fixation de nombreuses protéines modulatrices de l'expression des gènes. Deux types de *MARS* sont décrites (Hakes and Berezney, 1991) :

- Les *MARS* constitutives qui sont liées en continu avec la matrice nucléaire
- Les *MARS* facultatives dont la liaison avec la matrice nucléaire est dynamique

Les propriétés de liaison des *MARS* à la matrice nucléaire font de ces régions des participantes importantes pour l'organisation du noyau, notamment en permettant la formation de boucle de chromatine (Heng et al., 2004).

Bien que de nombreux modèles et études aient comparé l'effet de la délétion des divers éléments de la région *EμMARS*, l'implication des *MARS<sub>Eμ</sub>* dans les mécanismes liés au développement B n'était pas clairement établi. Alors que les modèles animaux de délétion de l'enhancer intronique (*cEμ* KO) ou de l'ensemble de la région (*EμFL* KO) établissent clairement la fonction de *cEμ* comme un régulateur de la recombinaison VDJ (Marquet et al., 2014; Perlot et al., 2005; Sakai et al., 1999), les observations controversées sur la SHM dans les modèles d'animaux et les hybridomes maintenaient un doute sur l'implication des régions *MARS<sub>Eμ</sub>* dans ce mécanisme (Li et al., 2010; Marquet et al., 2014; Perlot et al., 2005; Ronai et al., 2005). Un modèle murin portant la délétion des deux régions *MAR* flanquant *cEμ* (*MARS<sub>Eμ</sub>* KO) a été développé au laboratoire dans le but de répondre à cette controverse. L'étude de

ce modèle a révélé que les  $MARS_{E\mu}$  n'ont pas de rôle critique dans le développement B précoce ni dans le processus de CSR mais sont, en revanche, impliquées dans le processus de SHM. En effet, la délétion des  $MARS_{E\mu}$  induit une diminution modérée mais significative de la fréquence d'hypermutation (d'environ un facteur 2) dans la région intronique  $J_H$  en amont de  $cE\mu$ . Dans cette étude, nous montrons que cette diminution n'est pas liée à un défaut de transcription primaire du locus  $IgH$  et de l'action de AID mais à un recrutement moins efficace des facteurs de réparation « error prone » agissant en aval d'AID (Martin et al., 2023). Je co-signe ce travail publié dans *Frontiers in Immunology* en quatrième position (**Annexe 1**).

Les mécanismes moléculaires, médiés par les régions  $MARS_{E\mu}$ , agissant sur la régulation du locus  $IgH$  ne sont pas encore totalement décrits. Il est cependant connu que ces régions sont des sites d'ancrages de nombreuses protéines, appelées MARs BP (BP pour *Binding Protein*) (Wang et al., 2010). L'équipe B-NATION s'attache à comprendre l'ensemble des mécanismes qui régulent l'expression du locus  $IgH$  ; elle a donc poursuivi l'étude de l'implication des régions  $MARS_{E\mu}$  par l'étude de modèles murins présentant des délétions de certaines MAR BP, notamment SATB1 (Special A/T Rinch Binding Protein 1) (Yokota and Kanakura, 2014). Notre équipe a évalué l'effet de la délétion conditionnelle de SATB1 dans un modèle murin dès le stade pro-B : cette étude a révélé la dualité fonctionnelle de cette MAR BP dans les cellules B. En effet, SATB1 possède une activité activatrice de la transcription d' $IgH$  dans les cellules B matures au repos alors qu'elle endosse une activité répressive de l'expression d' $IgH$  ainsi que du mécanisme de SHM dans les cellules B activées (Thomas et al., 2023). Je co-signe ce travail en révision dans *Cellular and Molecular Immunology* en qualité de deuxième auteure (**Annexe 2**).

### **III.3. La région régulatrice en 3' (3'RR)**

La description de la structure complexe de cette région et de son rôle au cours du développement B a notamment été facilitée par l'étude de différents modèles murins, porteurs de délétions totales ou partielles de la région. Notre laboratoire s'est particulièrement illustré dans de nombreuses études sur la compréhension de cette région. La revue ci-après, que je signe en tant que première auteure et publiée dans *Advances in Immunology*, fait l'état des lieux des différents modèles murins ayant permis de déchiffrer le rôle de la 3'RR aux différents stades du développement B (Bruzeau et al., 2021).



# Panorama of stepwise involvement of the IgH 3' regulatory region in murine B cells

Charlotte Bruzeau<sup>a,b,c</sup>, Jeanne Moreau<sup>a,b,c</sup>, Sandrine Le Noir<sup>a,b,c</sup>, and Eric Pinaud<sup>a,b,c,\*</sup>

<sup>a</sup>CNRS, Contrôle de la Réponse Immune B et des Lymphoproliférations, UMR 7276, Limoges, France

<sup>b</sup>INSERM, Contrôle de la Réponse Immune B et des Lymphoproliférations, UMR 1262, Limoges, France

<sup>c</sup>Université de Limoges, Contrôle de la Réponse Immune B et des Lymphoproliférations, UMR 7276, UMR 1262, Limoges, France

\*Corresponding author: e-mail address: eric.pinaud@unilim.fr

## Contents

1. Structure of the IgH 3' boundary and associated regulatory region	96
2. Is the 3'RR involved in early B cell development?	98
3. Involvement of the 3'RR in Ig $\mu$ heavy chain production and B cell receptor expression	99
4. Individual 3'IgH enhancers control CSR by providing accessibility of acceptor switch regions	101
5. Involvement of the quasi-palindromic structure in somatic hypermutation	104
6. Plasma cell Ig heavy chain production depends on the full palindromic IgH 3'RR	106
7. Concluding remarks: Panorama of IgH 3'RR stepwise activation	107
8. Remaining 3'RR dark zones and future perspectives	109
Acknowledgments	110
Conflict of interest	110
References	110

## Abstract

Among the multiple events leading to immunoglobulin (Ig) expression in B cells, stepwise activation of the Ig heavy chain locus (*IgH*) is of critical importance. Transcription regulation of the complex *IgH* locus has always been an interesting viewpoint to unravel the multiple and complex events required for IgH expression. First, regulatory germline transcripts (GLT) assist DNA remodeling events such as VDJ recombination, class switch recombination (CSR) and somatic hypermutation (SHM). Second, productive spliced transcripts restrict heavy chain protein expression associated either with the surface receptor of developing B cells or secreted in large amounts in plasma cells. One main transcriptional regulator for *IgH* lies at its 3' extremity and includes both a set of

enhancers grouped in a large 3' regulatory region (3'RR) and a cluster of 3'CTCF-binding elements (3'CBEs). In this focused review, we will preferentially refer to evidence reported for the murine endogenous *IgH* locus, whether it is *wt* or carries deletions or insertions within the *IgH* 3' boundary and associated regulatory region.

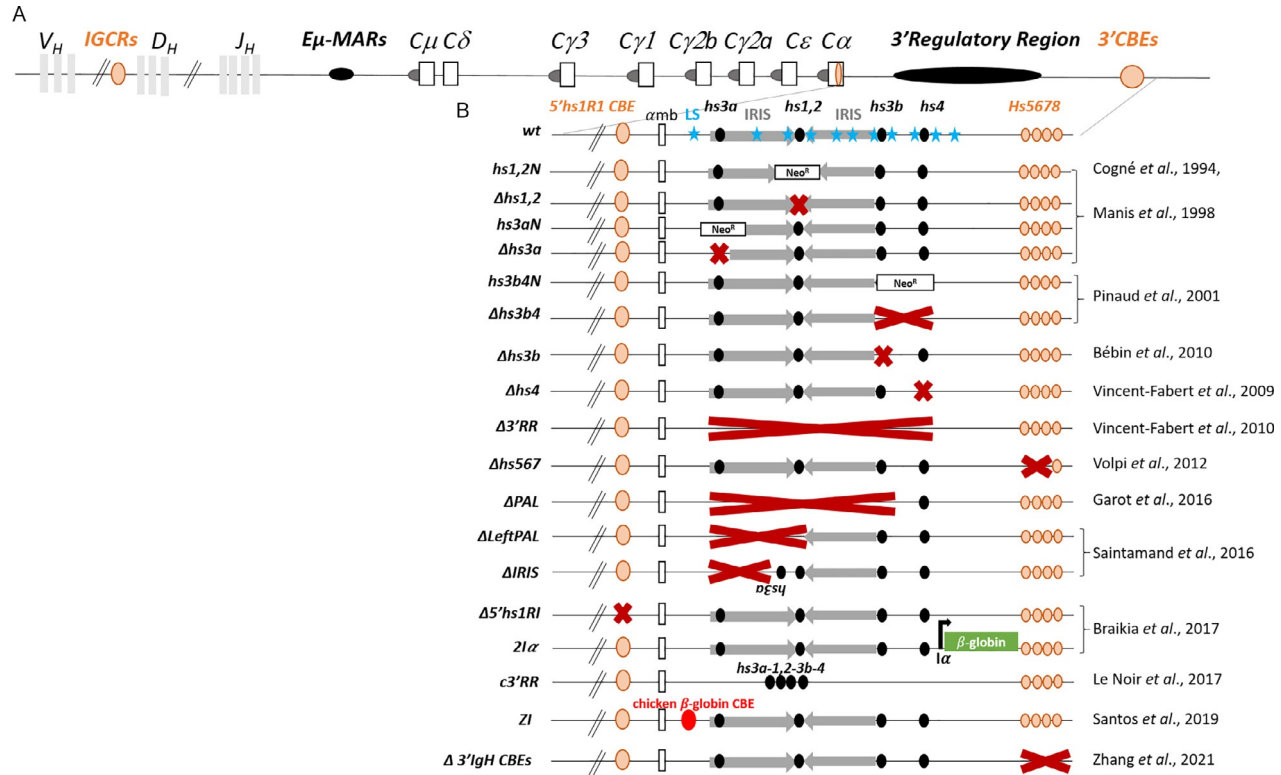


## 1. Structure of the *IgH* 3' boundary and associated regulatory region

Reflecting its tremendous complexity, the *IgH* locus spans approximately three megabases (Mb) on murine chromosome 12. Regions involved in transcriptional regulation, mainly bidirectional  $V_H$  promoters, switch region (*S*) I promoters, functional enhancers and CTCF-binding elements, stand between  $V_H$  to  $D_H$  clusters (IGCR elements), upstream from the  $C\mu$  gene ( $E\mu$  enhancer and *MARs* elements) and downstream from  $C\alpha$  (the 3'RR followed by the 3'CBEs) (Fig. 1A).

Initially identified as a locus control region stretching more than 30kb downstream from the  $C\alpha$  gene (Lieberson, Giannini, Birshstein, & Eckhardt, 1991), the 3'RR includes four transcriptional enhancers, named *hs3a*, *hs1-2*, *hs3b* and *hs4* (Giannini, Singh, Calvo, Ding, & Birshstein, 1993; Lieberson et al., 1991; Matthias & Baltimore, 1993; Michaelson, Giannini, & Birshstein, 1995). The central *hs1-2* element is flanked by inverted repeated intervening sequences (*IRIS*) that form a 25kb-long quasi-palindrome (Chauveau & Cogné, 1996). Although highly divergent in different species, *IRIS* sequences always stand as inverted copies on both sides of *hs1.2*, preserving a 3'RR quasi-palindromic symmetry (Sepulveda, Garrett, Pricewhelan, & Birshstein, 2005). Downstream from this palindrome, the *hs4* element is the last structure within the *IgH* locus carrying enhancer activity (Michaelson et al., 1995). Genomic regions that separate murine 3'enhancer elements also contain 11 “like switch  $\gamma$ ” stretches (*LS1* to *LS11*) that share sequence identity and DNA motifs with *S* $\gamma$  regions (Peron et al., 2012) (Fig. 1A). *LS* regions can, in germinal center B cells, be targeted by activation-induced cytidine deaminase (AID) leading to mutations (although to a lesser extent than in *S* regions) or to recombination (discussed further) (Peron et al., 2012).

Downstream from the 3'RR, the locus border is marked by four *hs* sites (*hs5-8*) that correlate with ~10 CTCF-binding elements (CBEs) (Garrett et al., 2005); to note, one cryptic CBE was identified upstream from the 3'RR enhancers in the intronic region between constant and membrane exons of the  $C\alpha$  gene (Thomas-Claudepierre et al., 2013) (Fig. 1A).



**Fig. 1** Immunoglobulin heavy chain loci and mouse models carrying deletions/insertions within the 3'RR. (A) *IgH* locus. (B) Mouse models. Light gray rectangles represent  $V_H$ ,  $D_H$  and  $J_H$  segment genes, white rectangles represent constant genes, dark gray circles represent switch regions and black ovals represent main regulatory elements of *IgH* loci. Orange circles represent CTCF Binding Elements (CBE). Gray arrows represent the inverted repeated intervening sequences (IRIS). Blue stars represent Like-Switch (LS) regions. Red crosses represent deletions in mouse models and white cassettes with "Neo<sup>R</sup>" represent resistant neomycin genes. Not to scale.



The  $3'$ CBE region was initially described as a border/insulator of the *IgH* regulatory domain (Garrett et al., 2005; Zhou et al., 2002). In light of genome-wide chromosome conformation capture (Hi-C) patterns, the  $3'$ CBE region has been definitively identified as the extremity of the topologically associating domain (TAD) that includes the complex *IgH* locus (Vian et al., 2018). The critical importance of this TAD boundary was illustrated by the deletion of the  $3'$ CBEs in the mouse that led to increased expression of genes located downstream from the *IgH* locus (Volpi et al., 2012); this assertion was confirmed in a recent study describing intense transcriptional activation of the 30 kb-downstream region upon the larger deletion of all  $\sim 10$   $3'$ CBEs (Zhang, Yoon, Chapdelaine-Williams, Kyritsis, & Alt, 2021).

The first disruption studies of the  $3'$ RR region in the mouse suggested its involvement in late B cell remodeling events such as CSR (Cogné et al., 1994; Manis et al., 1998), although growing evidence for stepwise activation of the  $3'$ RR (Garrett et al., 2005; Michaelson et al., 1995) encouraged scientists to decipher the function of this fascinating  $3'$ *IgH* region. For this purpose, multiple mouse knock-out (KO) or knock-in (KI) models that disrupted various parts or modules within the  $3'$ *IgH* region have been established over the last decades (Fig. 1B) and have greatly contributed to the description of its mechanism.



## 2. Is the $3'$ RR involved in early B cell development?

While transcriptional activity of the  $3'$ RR in developing B cells was initially questionable, multiple evidence supported the fact that the  $3'$ RR could be poised for transcriptional activity in bone marrow B-lineage cells. A mouse KI model inserting an ectopic transcription unit between *hs4* and the  $3'$ CBEs (*2I $\alpha$*  Fig. 1B), highlighted consistent expression of the reporter gene at the pre-B stage (Braikia et al., 2017). The presence of enhancer RNA transcripts (e-RNAs) initiated within the  $3'$ RR also attested to the accessibility and activity of this region. Indeed, discrete levels of e-RNAs at *hs3b*, *hs1–2* and *hs4* elements have been identified in pre-B cells (Braikia et al., 2015; Peron et al., 2012). Recently, global run-on sequencing (GRO-seq) in *Rag*-deficient pro-B cell lines (Zhang et al., 2019) indicated that such e-RNAs are mostly antisense transcripts.

In the *wt* context, the pioneering discovery of physical interactions between the  $3'$ *IgH* region and other parts of the locus (Wuerffel et al., 2007) has been of critical interest and opened the field of investigation in

developing B cells. In *Rag*-deficient pro-B cells, detailed 3C and 4C-seq experiments described flexible interactions involving the 3'CBEs and the whole *IgH* locus (Degner et al., 2011; Medvedovic et al., 2013). Multiple evidence for long distance interactions between the 3'RR/CBEs and the upstream part of the locus suggested a significant function of this regulatory region during *VDJ* recombination, especially regarding recent data connecting this process to chromatin loop extrusion (Zhang, Zhang, Ba, Liang, et al., 2019). So far, only a discrete defect in  $V_H$  gene contraction associated with increased proximal  $V_H$  segment usage was observed upon homozygous deletion of the 3'CBEs ( $\Delta hs5-7$  Fig. 1B) (Volpi et al., 2012). This suggests that other elements can serve as a boundary in the absence of *hs5-7*. Factually, homozygous deletion of the 3'RR in the mouse did not change flexible interactions in *Rag*-deficient pro-B cells (Medvedovic et al., 2013) nor drastically impair  $VDJ_H$  joining and the choice of  $V_H$  segment usage in pre-B cells (Rouaud et al., 2012). Beyond homozygous deletion models, mice carrying *IgH* alleles of distinct haplotypes have been useful to study the impact of regulatory elements. In such animals carrying heterozygous deletions, IgM-allotype expression in developing B-lineage cells reflects competition between the "a" mutant (from the 129 strain) and the "b" *wt* (from the C57BL/6 strain) *IgH* alleles. Thereby, with this strategy, the fact that newly formed/immature bone marrow B cells from F1 animals occurred in similar proportions when the functional H chain was driven by the *IgH<sup>b</sup>* *wt* allele competing with an *IgH<sup>a</sup>* allele devoid of the 3'RR strongly argued for minor involvement in *VDJ* reshuffling (Garot et al., 2016). However, the statement that the 3'RR was not involved in  $VDJ_H$  rearrangements has been elegantly challenged in a study detailing regulatory transcripts in *Rag*-deficient precursors (Braikia et al., 2015): while both the *D* to  $J_H$  event and the ordered recombination sequence remained normal, authors highlighted a repressive effect of the 3'RR on the *V* to  $DJ_H$  event. The origin of this intriguing silencing effect of the 3'RR is currently unknown, but this effect ceased once *V* to  $DJ_H$  recombination was completed (Braikia et al., 2015).



### 3. Involvement of the 3'RR in Ig $\mu$ heavy chain production and B cell receptor expression

Once a functionally recombined  $\mu$  heavy chain is expressed in large pre-B cells, the question of the involved enhancer(s) points to the 3'*IgH* region. Singularly, a decrease in BCR surface expression on resting B cells

was a phenotypic feature of all KO mouse models that deleted the *hs4* enhancer (Pinaud et al., 2001; Vincent-Fabert et al., 2010, 2009); this recurrent “BCR<sup>low</sup>” phenotype was assigned to a defect in  $\mu$  heavy chain transcription. This suggested that, in resting splenic B cells, *hs4* was a privileged partner of the  $V_H$  promoter belonging to the rearranged  $VDJ_H$  transcription unit. Since it is involved in long distance interactions in developing B cells (Degner et al., 2011; Medvedovic et al., 2013), it was also questionable whether the  $3'$ CBEs could be involved in optimizing functional  $\mu$ H chain expression. The answer is clearly not; this statement again came from mouse models carrying *IgH<sup>a</sup>* and *IgH<sup>b</sup>* haplotypes in which *wt* and *hs5–7* KO alleles produced identical BCR surface expression ( $\Delta$ *hs567* Fig. 1B (Volpi et al., 2012)). Evidence that *hs4* by itself was required for optimal BCR expression in developing B cells was also provided by heterozygous *a/b* mice that measured competition between *wt* and various KO alleles within the  $3'$ RR:  $\Delta$ *hs3b-4*,  $\Delta$ *PAL* and  $\Delta$  $3'$ RR. (Garot et al., 2016). In this study, IgM<sup>s</sup> surface expression was constantly decreased when the allele was devoid of *hs4* ( $\Delta$ *hs3b-4* and  $\Delta$  $3'$ RR) in B-lineage cell subsets that carry a functional IgM-BCR at their surface: newly-formed/immature and recirculating bone marrow B cells as well as splenic transitional, marginal zone, follicular and B1 B cells. As expected, such a decrease in IgM-BCR levels at the surface of developing B cells impaired peripheral B cell fate and limited their numbers (Garot et al., 2016; Ghazzoui et al., 2018; Saintamand et al., 2015). This decreased BCR expression, a common feature of naïve B cells lacking *hs4* (Garot et al., 2016; Vincent-Fabert et al., 2009), could predict impaired signaling in peripheral B cell populations. Factually, the impact was not drastic but a mild reduction in phosphorylated tyrosine targets was observed in resting and BCR-challenged splenocytes devoid of the  $3'$ RR (Saintamand, Rouaud, Garot, et al., 2015).

Peripheral B2 B cell subsets missing the  $3'$ RR displayed a normal  $VDJ_H$  repertoire (Ghazzoui et al., 2018); this was not the case for the peritoneal B1 cell population that, when lacking the  $3'$ RR, limited its H chain repertoire to the use of  $V_H1$  and  $V_H11$  segment subtypes (Ghazzoui et al., 2018). In line with locus wide loops and interactions that highlighted physical proximity of *IgH* elements (Vian et al., 2018; Wuerffel et al., 2007; Zhang, Zhang, Ba, Liang, et al., 2019), one undeniable conclusion was that the distal *hs4* enhancer combined with the  $3'$ CBEs are sufficient to drive optimal expression of the functional  $\mu$ H chain. Therefore, this distal part of the *IgH* locus (*hs4* to  $3'$ CBE) could be classified as a functional unit dedicated to regulation of IgM-BCR expression in immature and naïve B cells.



#### 4. Individual 3'IgH enhancers control CSR by providing accessibility of acceptor switch regions

It is manifest that some specific 3'RR activity was induced upon B cell activation (Oudinet, Braikia, Dauba, & Khamlichi, 2020), given that the region accumulated RNA pol II (Yamane et al., 2011) and engaged in sense and antisense transcription (Garot et al., 2016; Peron et al., 2012; Saintamand et al., 2017; Zhang et al., 2019).

More specifically, obvious strong CSR defects were initially observed as a common feature of the first mouse models that introduced a selection marker within the 3'RR and specifically upstream from the *hs4* element (*hs3aN*, *hs1–2N* and *hs3b4N* on Fig. 1B (Cogné et al., 1994; Manis et al., 1998; Pinaud et al., 2001)). Such a drastic effect, induced by the activation of a neomycin-selection cassette, can now be interpreted as a global disruption of the 3'RR unit since further studies describing individual “clean” deletions of single enhancer elements within this region did not significantly affect CSR (either *hs3a*, *hs1–2* (Manis et al., 1998), *hs3b* (Bébin et al., 2010) or *hs4* (Vincent-Fabert et al., 2010)). The suppression of at least two elements such as *hs3b* and *hs4* confirmed a functional role for 3'IgH enhancers on CSR by inducing a global decrease in GLT synthesis at *I* promoters of most *S* acceptor regions (except for *Sy1*) (Pinaud et al., 2001). The larger  $\Delta$ PAL deletion ( $\sim 20$  kb from *hs3a* to *hs3b*) also resulted in CSR and associated GLT defects (including *Sy1*) but maintained consistent class switching to IgG2b and IgA, suggesting a specific function for the distal *hs4* element in maintaining residual CSR (Garot et al., 2016). Suppression of the entire ( $\sim 30$  kb) 3'RR almost completely abrogated CSR events, including switching to IgG1, due to a drastic defect in GLT of all constant genes (Vincent-Fabert et al., 2010). The impact of the 3'RR on global *S* region chromatin patterns was also examined in the same model with the conclusion that this region was critical in providing efficient accessibility to transcription and AID-targeting of acceptor *S* regions but barely affected the donor *S $\mu$*  region (Saintamand et al., 2015). In agreement with the impaired targeting of all *S* acceptor regions, absence of the 3'RR curiously highlighted the occurrence of class switching to IgD as a discrete but physiological phenomenon (Issaoui et al., 2018; Rouaud et al., 2013). Whereas the increase in CSR defects upon gradual enhancer deletions confirmed functional redundancies of its components, it also aroused curiosity concerning a possible function for intervening regions that separate core-enhancers.

Several mouse models were created to answer this question by partially or completely removing the 3' palindromic architecture (Fig. 1): the large  $\Delta$ PAL deletion (20 kb from *hs3a* to *hs3b*) globally deleted the proximal 3'RR module (Garot et al., 2016), the  $\Delta$ leftPAL (11.5 kb from *hs3a* to *hs1-2*) suppressed the left half part of the palindrome, the  $\Delta$ IRIS model suppressed the same half part but reintroduced core *hs3a* and *hs1-2* while avoiding inverted structures (Saintamand et al., 2016) and lastly the *c*3'RR model minimized the region to its four core enhancers (Le Noir, Boyer, et al., 2017, Le Noir, Laffleur, et al., 2017). Comparison of these models helped clarify the function of core enhancers in CSR. The consistent CSR defects observed, at least for IgG3 and IgG1 subclasses, following  $\Delta$ PAL and  $\Delta$ leftPAL deletions confirmed involvement of the proximal *hs3a* and *hs1-2* enhancers in CSR to main isotypes (Garot et al., 2016; Saintamand et al., 2016). Interestingly, reintroduction of proximal *hs3a* and *hs1-2* in the  $\Delta$ IRIS model rescued CSR to IgG3 and IgG2b when compared to  $\Delta$ leftPAL, suggesting that core enhancers by themselves could improve CSR efficiency independently of the left palindromic part. This statement was nicely confirmed by the *c*3'RR model that almost restored CSR efficiency to normal for all isotypes upon reintroduction of the four core enhancers without any IRIS region. (Le Noir, Boyer, et al., 2017, Le Noir, Laffleur, et al., 2017). In the light of models harboring large deletions ( $\Delta$ PAL,  $\Delta$ IRIS,  $\Delta$ leftPAL), one striking observation is that CSR to IgA was never compromised when *hs4* was conserved, suggesting that the distal part of the 3'RR is sufficient for this particular event (Garot et al., 2016; Saintamand et al., 2016). Furthermore, in the particular B1a compartment, CSR to IgA even occurs independently of the 3'RR (Issaoui et al., 2018). By analyzing the multiple KO models cited above, the question of the specificity of activation of the different acceptor-S regions by distinct 3'RR enhancers is still difficult to unravel. Indeed, some compensatory mechanisms, a potential consequence of redundancy between elements and/or the structure of their surrounding chromatin, might obscure any potential clearcut effect.

Nonetheless, recent advances reported that chromatin loop extrusion could account for dynamic and region-specific optimal locus configurations for CSR (Zhang, Zhang, Ba, Kyritsis, et al., 2019). This model proposes the 3'RR as an *IgH* locus entry point for cohesin complexes. Once loaded, the complex promotes loop extrusion moving towards the constant regions until it becomes impeded by the  $E\mu$ - $S\mu$  region in resting B cells. The resulting  $E\mu$ -3'RR interaction, previously described by 3C (Wuerffel et al., 2007), isolates the *IgH* constant regions in a  $\sim$ 200 kb loop. The

3'RR of CSR-activated cells could load additional extrusion complexes to form internal loops towards a given and impeding acceptor-S region, leading to a CSR complex that juxtaposes AID targets and 3'enhancers (Zhang, Zhang, Ba, Kyritsis, et al., 2019). Using the 3'RR as an entry point for cohesin complexes, dynamic and successive loop extrusions turn out to be a convincing mechanism for GLT regulation during CSR. This hypothesis was strongly supported by a study published by the Alt group (during the review process of this manuscript) describing a consistent CSR defect to most isotypes (excepted for IgG1) in a mouse model devoid of all ~ 10 consecutive 3'IgH CBEs (Fig. 1) (Zhang et al., 2021), arguing that cohesin complexes entry point located within the 3' boundary. Although, some additional findings need to be reconciled with this model. A first example is the modest but striking increased CSR (to IgG1 mostly) obtained upon deletion of the *hs5-7* CBEs (Volpi et al., 2012) that could account for a potential negative effect of the 3'IgH boundary. A second interesting example, highlighted by the Khamlichi group, is the potential blocking function of the CBE located immediately upstream from the 3'RR (Thomas-Claudepierre et al., 2013). Since the mouse model lacking this 5'*hs1RI* displayed constitutive activation of *Sy3* GLT (Braikia et al., 2017) one could suppose that this CBE might prevent extrusion/looping of *C* genes in resting B cells. A third example, described by the same group, is the variable co-induction of GLT from two *I* promoters of the same allele: frequent *Sy3-Sy2b* GLT and rare *Sy1-Se* or *Sa-Sy2b* GLT occurring upon respectively LPS, IL4 and TGF- $\beta$  induction (Santos, Braikia, Oudinet, Dauba, & Khamlichi, 2019). Such events, described as a 3'RR-mediated coactivation of *I* promoters, imply that multiple internal loops juxtapose simultaneously one 3'RR and several *S*-acceptor regions. Another striking observation is that a functional 3'RR, located on an unexpressed *IgH* allele, could rescue significant CSR events on a 3'RR-defective allele (Dougier et al., 2006; Le Noir, Boyer, et al., 2017; Le Noir, Laffleur, et al., 2017). This 3'RR potency for trans-allelic complementation, formerly called transvection in drosophila (Lewis, 1954), implies physical crosstalk orchestrated by the 3' region; this hypothesis was supported by the consistent decrease in *IgH* allele pairing upon 3'RR deletion in activated B cells (Le Noir, Boyer, et al., 2017, Le Noir, Laffleur, et al., 2017).

Beyond transcription-induced accessibility for CSR, *IgH* cis-regulatory elements could potentially influence the outcome of DNA double strand breaks. Analysis of CSR junctions, facilitated by a dedicated bioinformatic pipeline (Boyer et al., 2017), is a convenient method to examine the involvement of protein partners involved in joining. CSR junctions

occurring in the  $\Delta 3'RR$  model displayed a similar pattern when compared to *wt* mice (Saintamand, Rouaud, Saad, et al., 2015), suggesting that the few CSR-induced DNA breaks occurring in the absence of the  $3'RR$  could be resolved by the predominant non-homologous end joining pathway (Stavnezer, Björkman, Du, Cagigi, & Pan-Hammarström, 2010; Yan et al., 2007). The non-conventional  $S\mu$ - $S\delta$  junctions, overrepresented in this same  $\Delta 3'RR$  model, were also identical to these rare physiological events since they displayed frequent microhomologies (Rouaud et al., 2013). This also suggests that the alternative-end joining (AEJ) pathway, involved to some extent in CSR, could also be efficiently recruited in the absence of the  $3'RR$  (Stavnezer et al., 2010; Yan et al., 2007). To note, AEJ has also been shown to be the preferential joining mechanism for the rare LSR events involving LS regions within the  $3'RR$  (Boutouil, Boyer, Cook-Moreau, Cogné, & Péron, 2019).

Moreover, the *IgH*  $3'RR$  is itself targeted by AID during locus suicide recombination (LSR). This event, in analogy with CSR, consists of the recombination between the  $S\mu$  donor and one of the 11 0.5- to 1-kb-long LS acceptor-regions spread throughout the  $3'RR$  (Peron et al., 2012). In the mouse, LSR leads to complete deletion of all constant segment genes and results in the loss of BCR expression when the targeted allele is functional. By depriving the B cell of its own surface receptor, such a drastic event abrogates survival signals and provokes cell death.

Studying such an event that leads to the disappearance of a targeted cell population is obviously a difficult task. Recently developed sensitive techniques, such as ligation amplification mediated high throughput genome-wide translocation sequencing (LAM-HTGTS), detected consistent LSR genetic hallmarks in activated B cells either *in vitro*, in datasets published by colleagues (Dong et al., 2015) or in our own datasets from chronically activated Peyer's patch B cells (Sandrine Le Noir, unpublished observations).



## 5. Involvement of the quasi-palindromic structure in somatic hypermutation

Within the *IgH* locus, somatic hypermutation (SHM) preferentially targets the coding-*VDJ* transcription unit and its immediate downstream non-coding intronic region. To a lesser extent, significant levels of base substitutions are observed in other *IgH* non-coding regions, such as donor and acceptor *S* regions and the  $3'RR$  LS regions for which the SHM-related process is still debatable. Since transcription is required for deamination



of these physiological targets by AID (Chaudhuri et al., 2003; Fukita, Jacobs, & Rajewsky, 1998; Peters & Storb, 1996), *IgH* regulatory regions play a crucial function in the process. Beyond cytidine to uracil conversion by AID, SHM mechanisms include specific processing of the resulting dU:dG mismatch by actors of the DNA base excision-(BER) and mismatch repair (MMR) pathways (for review Maul & Gearhart, 2010). BER- and MMR-associated factors promote unfaithful repair and efficiently increase mutations in *Ig* loci targeted regions; this singular process raises the question of the *cis*-acting elements involved in recruiting dedicated actors. In immunized mice devoid of the 3'CBE, SHM was tested in spleen GC B cells and only decreased slightly (Volpi et al., 2012). In other murine models, SHM was evaluated in intronic regions flanking rearranged *VDJ* segments of spontaneous GC B cells sorted from Peyer's patches. Single deletions of *hs3b* or *hs4* core enhancers (Bébin et al., 2010; Vincent-Fabert et al., 2009) did not impact the SHM rate in *VDJ* rearranged segments. Once more, by showing that individual enhancers are dispensable for SHM, such models highlighted some functional redundancies between elements of the 3'RR. Dual suppression of both *hs3* and *hs4* led to intermediate phenotypes. SHM was normal when evaluated in the mutated allele of heterozygous mice (Morvan, 2003), but reduced by twofold when SHM was assessed in homozygous animals by NGS using an efficient pipeline (Bruzeau and Le Noir, unpublished observations), that quantifies global SHM for all *IgH* alleles (Martin et al., 2018). This then potentially accounts for an increased proportion of unmutated alleles. By contrast, deletion of the entire region proved that the 3'RR is mandatory for SHM since mutations within *IgH V* and flanking regions were almost abrogated in this model (Rouaud et al., 2013). Indeed, a straightforward interpretation could assign the mutation defect to a two-to-three-fold drop in primary transcription initiated at the  $V_H$  promoter. In addition, a singular observation in these mice was that almost all *IgH* alleles in GC B cells remained unmutated, highlighting the incapacity of 3'RR deficient alleles to recruit an AID-associated SHM complex. However, when considering the few mutated alleles (about 6%), their mutation frequency was reduced by threefold (Rouaud et al., 2013). Strikingly, a similar threefold decrease in SHM frequency coupled to a two- to threefold decrease in *IgH* primary transcription was also observed in the model lacking the quasi-palindromic region  $\Delta PAL$  (Garot et al., 2016). When compared side by side, these two models suggest that the proximal quasi-palindromic region (containing *hs3a*, *hs1-2*, *hs3b* enhancers and their flanking regions) is required for optimal primary transcription and AID targeting in order to reach an optimal SHM level. While *hs4* does not enhance SHM in the



$\Delta$ PAL model, the presence of this enhancer increased the proportion of mutated alleles (up to 27%) among GC B cell populations (Garot et al., 2016). It was then tempting to speculate that *hs4* participates in the recruitment of SHM machinery on its own allele. More recently, studies in mouse models that conserved the distal 3'RR module ( $\Delta$ IRIS,  $\Delta$ leftPAL and  $c3'RR$ ) provided some clear-cut evidence for the critical role of this particular quasi-palindromic construction for SHM. First,  $\Delta$ LeftPAL and  $\Delta$ IRIS models indicated that disruption of the 5' part of the quasi-palindrome was sufficient to induce a consistent SHM defect, confirming the requirement of the 3'RR structure for the process (Saintamand et al., 2016). Second, once the palindromic structure was disrupted, reintroduction of *hs3a* and *hs1–2* elements in the  $\Delta$ IRIS model was not sufficient to restore SHM and primary transcription to normal levels (Saintamand et al., 2016). In agreement with such statements, a dramatic decrease in SHM was observed in the complementary  $c3'RR$  model, devoid of the intervening packaging DNA and preserving the four enhancers intact (Le Noir, Boyer, et al., 2017; Le Noir, Laffleur, et al., 2017). Displaying a SHM defect similar to that observed in  $\Delta$ 3'RR, this novel model attests that while 3'RR core enhancers weakly contribute to optimal SHM, the singular DNA structure of this region plays a major role in this process. Although core enhancers are mandatory for full activity of the 3'RR super-enhancer, the surrounding sequences do not constitute useless random DNA regions, but obviously upgrade the 3'RR activity for SHM.

Since the nature of base substitutions downstream from AID deamination is nonrandom but dictated by error prone repair pathways, base-substitution patterns in Ig gene regions not selected by antigen could constitute interesting hallmarks but remain difficult to decode. So far, almost all 3'RR mutant models described in Fig. 1B displayed normal base substitution patterns, suggesting that BER and MMR pathways could be efficiently recruited downstream from cytidine deamination by AID, even when, depending on the model, SHM frequency was decreased.



## 6. Plasma cell Ig heavy chain production depends on the full palindromic IgH 3'RR

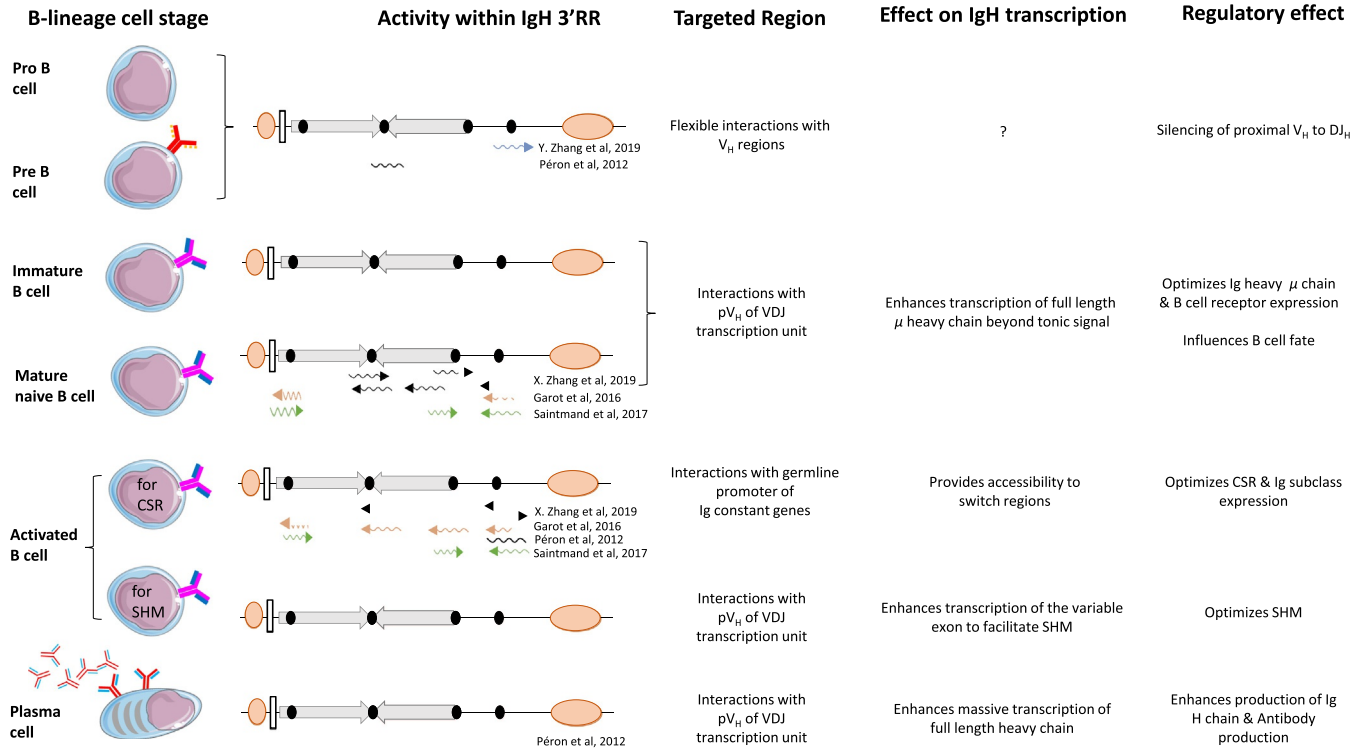
Being largely involved in CSR and SHM, the *IgH* 3'RR is also absolutely required for efficient Ig heavy chain production during the ultimate developmental stage of B cell development. Straightforward and simple readouts were provided by IgM levels in the serum of mouse models. Here again, while individual or dual enhancer deletions did not impact IgM secretion (Bébin et al., 2010; Manis et al., 1998; Pinaud et al., 2001;

Vincent-Fabert et al., 2010), larger deletions within the 3'RR led to severe decreases. Basic IgM serum production was lowered by 10-fold in  $\Delta 3'RR$  and  $\Delta PAL$  mice; such models also displayed an impaired antigen-driven IgM response (Garot et al., 2016; Vincent-Fabert et al., 2010) emphasizing a critical function of the proximal quasi-palindromic module of the 3'RR. The Ig secretion defect in plasma cells was demonstrated in heterozygous *a/b* mice in which intracellular  $IgM^d$  production was strongly decreased when driven by a  $\Delta PAL$  allele (Garot et al., 2016). Disorganization of the 3'RR structure, carried out in  $\Delta LeftPAL$  and  $\Delta IRIS$  mice, also led to a decrease in serum IgM production but a lesser extent (Saintamand et al., 2016), indicating that the full palindromic configuration is required for optimal antibody secretion. This statement was later confirmed in the *c3'RR* model lacking the 3'RR DNA packaging. Mice displayed a strong defect in IgM secretion (Le Noir, Boyer, et al., 2017, Le Noir, Laffleur, et al., 2017), comparable to that observed in  $\Delta 3'RR$  or  $\Delta PAL$  models (Garot et al., 2016; Le Noir, Boyer, et al., 2017; Le Noir, Laffleur, et al., 2017). This severe Ig secretion defect in the *c3'RR* model also indicated that core 3'enhancers are not sufficient to support optimal Ig production (Le Noir, Boyer, et al., 2017, Le Noir, Laffleur, et al., 2017). The consistent reduction in  $Ig\mu$  secreted transcripts in  $\Delta PAL$ , *c3'RR* and  $\Delta 3'RR$  homozygous mutants indicated obvious transcription failure in secreting B cells of these models (Garot et al., 2016; Le Noir, Boyer, et al., 2017; Le Noir, Laffleur, et al., 2017). Taken together, these elements pinpoint the essential function of the 3'RR in plasma cells: both enhancers and the full palindromic structure are required for optimal Ig transcription, potentially by optimizing the enhancing function of the core 3'RR elements. It should be noted that high levels of 3'RR eRNA transcripts around *hs1*- and *hs4* are detectable in plasma cells (Peron et al., 2012), although this does not compulsorily predispose to strong enhancer activity (Li, Notani, & Rosenfeld, 2016). In line with observations made by the Birshstein group in plasmocytoma cell lines (Ju et al., 2007), an attractive mechanistic picture could be that the singular palindromic structure favors long distance linking between the 3'RR and the *VDJ<sub>H</sub>* transcription unit.



## 7. Concluding remarks: Panorama of IgH 3'RR stepwise activation

The rich collection of mouse models mutated within the *IgH 3'RR* has been accumulated over years and has helped to establish quite a large panorama of specific activities of its multiple components during B cell development (Fig. 2). The concept of functional modules within this complex



**Fig. 2** Overview of B Cell Development and 3'RR implication. This figure shows the chronological order of B cell stages in the bone marrow and secondary lymphoid organs. The implication of 3'RR components is shown by highlighting them at the corresponding stage (blue represents a silencing effect and red an enhancing effect). Blue arrow represents transcripts identified by GRO-Seq in pro-B cells (Zhang, Zhang, Ba, Liang, et al., 2019). Black arrows represent transcripts identified by GRO-Seq in resting and αCD40 + IL4 stimulated spleen B cells (Zhang, Zhang, Ba, Kyritsis, et al., 2019). Orange arrows represent transcripts identified by RNA-Seq in resting and LPS stimulated splenic B cells (Garot et al., 2016). Black lines represent transcripts identified by qRT-PCR in activated B cells and plasma cells (Péron et al., 2012). Green arrows represent transcripts identified by RNA-Seq in resting and LPS stimulated B cells (Saintamand et al., 2017).

region, a notion shared by *IgH 3'RR* pioneers such as the Birshstein and Alt groups, has now become evident. Beyond the two simple modules that consist of the *3' CBE* and the distal *hs4* involved, respectively, in long range looping and Ig H chain expression in immature and naïve B cells, the proximal palindromic module can be functionally subdivided. The components of this palindromic module, including transcriptional enhancers and *IRIS* regions, could either act as separate modules during CSR and SHM or as a scaffolded full module in plasma cells. This conception, factually supported by mouse phenotypes, is in line with the former hypothesis of stepwise activation of this region during B cell ontogeny. Whether the *3'RR* acts as a superenhancer (SE) is probably no longer a matter of debate since some long-standing statements, formerly identifying this region as a locus control region (Madisen & Groudine, 1994), match with the established SE criteria. Composed of clustered enhancers, this region is only active in B-lineage cells and its transcriptional activity rises progressively throughout B cell development. *3'RR* enrichment for the typical H3K27Ac histone mark, shared with *E $\mu$* , *E $\kappa$*  and *3'Ek* regions in Ig light chain loci (Meng et al., 2014; Qian et al., 2014), confirms the pivotal function of this singular region as a master locus regulator.



## 8. Remaining *3'RR* dark zones and future perspectives

Several unsolved questions remain regarding the repetitive nature of the *IRIS* that structures the *3'RR*. The striking conservation of inverted repeated regions flanking *hs1–2* in the *3'RR* of humans and other mammals reflects a dominant function but remains mechanistically unsolved. The origin and the selection pressure that maintain such a singular *3'RR* architecture is up until now mysterious. One interesting analogy in mammalian genomes is the discovery of large palindrome structures (up to 3Mb in length) in male-specific portions of the Y chromosome (Rozen et al., 2003; Skaletsky et al., 2003), that have potentially contributed, throughout evolution, to select certain genes involved in male fertility. A recent Hi-C study highlighted frequent contacts within and between palindromes of Y chromosomes (Cechova et al., 2020), suggesting that such singular regions are dynamically evolving possibly by a mechanism close to gene conversion. Whether the *3'RR* can act, beyond transcription, on the recruitment of specific factors involved downstream from AID is still questionable.

The question of the potential silencing function of the *IgH 3'RR* at early developmental stages, a hypothesis proposed by the Khamlichi group, is definitely an interesting viewpoint and deserves to be studied further.

Remarkably, cells engaged in plasma cell differentiation undergo nuclear clustering of loci involved in antibody production (i.e., loci coding Ig heavy, light and J chains located on three different chromosomes) in close proximity to nuclear pore complexes (Park, Xiang, Feng, & Garrard, 2014). This pioneering study proposes that Ig gene regulatory regions, such as the *3'RR*, are important regulators of this clustering; the various mouse models listed in this review could constitute ideal tools to verify this hypothesis.

As a follow up to the models that delete palindromic *3'RR* sequences, a lingering question concerns the 3D structure adopted by this remarkable module. Is the DNA sequence composition by itself sufficient to confer a specific architecture or are protein folding factors required? These unanswered questions will no doubt be the object of intensive studies in the years to come.

## Acknowledgments

This work has been supported by the Fondation ARC pour la recherche sur le cancer, La Ligue contre le Cancer Comités du Limousin, Institut National du Cancer–Canceropole Grand Sud-Ouest, Comité d'Orientation Recherche Cancer en Limousin, Conseil Régional Nouvelle Aquitaine, Agence Nationale de la Recherche. C.B. was supported by a Ph.D. fellowship of the french Ministère de l'Enseignement Supérieur, de la Recherche et de de l'Innovation.

We regret, because of space restrictions, not being able to cite the numerous cellular and transgenic models published by all colleagues who strongly contributed to the elucidation of the function of this singular *IgH* regulatory region in the mouse. Indeed, this review does not cover *IgH* locus regulation in other species.

## Conflict of interest

The authors declare no conflict of interest.

## References

- Bébin, A.-G., Carrion, C., Marquet, M., Cogné, N., Lecardeur, S., Cogné, M., et al. (2010). In vivo redundant function of the *3' IgH* regulatory element HS3b in the mouse. *Journal of Immunology*, *184*, 3710–3717. <https://doi.org/10.4049/jimmunol.0901978>.
- Boutouil, H., Boyer, F., Cook-Moreau, J., Cogné, M., & Péron, S. (2019). *IgH* locus suicide recombination does not depend on NHEJ in contrast to CSR in B cells. *Cellular & Molecular Immunology*, *16*, 201–202. <https://doi.org/10.1038/s41423-018-0172-2>.
- Boyer, F., Boutouil, H., Dalloul, I., Dalloul, Z., Cook-Moreau, J., Aldigier, J.-C., et al. (2017). CSRreport: A new computational tool designed for automatic analysis of class switch recombination junctions sequenced by high-throughput sequencing. *The Journal of Immunology*, *198*, 4148–4155. <https://doi.org/10.4049/jimmunol.1601924>.

- Braikia, F.-Z., Conte, C., Moutahir, M., Denizot, Y., Cogné, M., & Khamlichi, A. A. (2015). Developmental switch in the transcriptional activity of a long-range regulatory element. *Molecular and Cellular Biology*, *35*, 3370–3380. <https://doi.org/10.1128/MCB.00509-15>.
- Braikia, F.-Z., Oudinet, C., Haddad, D., Oruc, Z., Orlando, D., Dauba, A., et al. (2017). Inducible CTCF insulator delays the *IgH* 3' regulatory region-mediated activation of germline promoters and alters class switching. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 6092–6097. <https://doi.org/10.1073/pnas.1701631114>.
- Cechova, M., Vegesna, R., Tomaszkiwicz, M., Harris, R. S., Chen, D., Rangavittal, S., et al. (2020). Dynamic evolution of great ape Y chromosomes. *Proceedings of the National Academy of Sciences of the United States of America*, *117*, 26273–26280. <https://doi.org/10.1073/pnas.2001749117>.
- Chaudhuri, J., Tian, M., Khuong, C., Chua, K., Pinaud, E., & Alt, F. W. (2003). Transcription-targeted DNA deamination by the AID antibody diversification enzyme. *Nature*, *422*, 726–730. <https://doi.org/10.1038/nature01574>.
- Chauveau, C., & Cogné, M. (1996). Palindromic structure of the IgH 3' locus control region. *Nature Genetics*, *14*, 15–16. <https://doi.org/10.1038/ng0996-15>.
- Cogné, M., Lansford, R., Bottaro, A., Zhang, J., Gorman, J., Young, F., et al. (1994). A class switch control region at the 3' end of the immunoglobulin heavy chain locus. *Cell*, *77*, 737–747. [https://doi.org/10.1016/0092-8674\(94\)90057-4](https://doi.org/10.1016/0092-8674(94)90057-4).
- Degner, S. C., Verma-Gaur, J., Wong, T. P., Bossen, C., Iverson, G. M., Torkamani, A., et al. (2011). CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the *Igh* locus and antisense transcription in pro-B cells. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 9566–9571. <https://doi.org/10.1073/pnas.1019391108>.
- Dong, J., Panchakshari, R. A., Zhang, T., Zhang, Y., Hu, J., Volpi, S. A., et al. (2015). Orientation-specific joining of AID-initiated DNA breaks promotes antibody class switching. *Nature*, *525*, 134–139. <https://doi.org/10.1038/nature14970>.
- Dougier, H.-L., Reynaud, S., Pinaud, E., Carrion, C., Delpy, L., & Cogné, M. (2006). Interallelic class switch recombination can reverse allelic exclusion and allow trans-complementation of an IgH locus switching defect. *European Journal of Immunology*, *36*, 2181–2191. <https://doi.org/10.1002/eji.200535529>.
- Fukita, Y., Jacobs, H., & Rajewsky, K. (1998). Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity*, *9*, 105–114. [https://doi.org/10.1016/S1074-7613\(00\)80592-0](https://doi.org/10.1016/S1074-7613(00)80592-0).
- Garot, A., Marquet, M., Saintamand, A., Bender, S., Le Noir, S., Rouaud, P., et al. (2016). Sequential activation and distinct functions for distal and proximal modules within the IgH 3' regulatory region. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 1618–1623. <https://doi.org/10.1073/pnas.1514090113>.
- Garrett, F. E., Emelyanov, A. V., Sepulveda, M. A., Flanagan, P., Volpi, S., Li, F., et al. (2005). Chromatin architecture near a potential 3J end of the *Igh* locus involves modular regulation of histone modifications during B-cell development and in vivo occupancy at CTCF sites. *Molecular and Cellular Biology*, *25*, 15.
- Ghazzau, N., Issaoui, H., Saintamand, A., Oblat, C., Carrion, C., & Denizot, Y. (2018). The immunoglobulin heavy chain 3' regulatory region superenhancer controls mouse B1 B-cell fate and late VDJ repertoire diversity. *Blood Advances*, *2*, 252–262. <https://doi.org/10.1182/bloodadvances.2017014423>.
- Giannini, S. L., Singh, M., Calvo, C. F., Ding, G., & Birshtein, B. K. (1993). DNA regions flanking the mouse Ig 3' alpha enhancer are differentially methylated and DNAase I hypersensitive during B cell differentiation. *The Journal of Immunology*, *150*, 1772–1780.

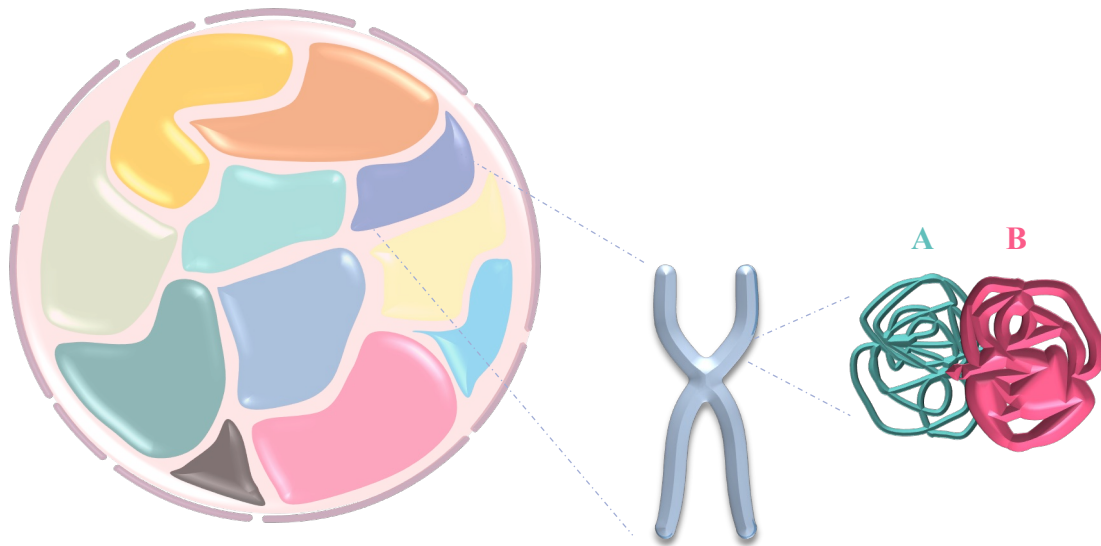
- Issaoui, H., Ghazzaoui, N., Saintamand, A., Carrion, C., Oblet, C., & Denizot, Y. (2018). The IgH 3' regulatory region super-enhancer does not control IgA class switch recombination in the B1 lineage. *Cellular & Molecular Immunology*, *15*, 289–291. <https://doi.org/10.1038/cmi.2017.103>.
- Ju, Z., Volpi, S. A., Hassan, R., Martinez, N., Giannini, S. L., Gold, T., et al. (2007). Evidence for physical interaction between the immunoglobulin heavy chain variable region and the 3' regulatory region. *The Journal of Biological Chemistry*, *282*, 35169–35178. <https://doi.org/10.1074/jbc.M705719200>.
- Le Noir, S., Boyer, F., Lecardeur, S., Brousse, M., Oruc, Z., Cook-Moreau, J., et al. (2017). Functional anatomy of the immunoglobulin heavy chain 3' super-enhancer needs not only core enhancer elements but also their unique DNA context. *Nucleic Acids Research*, *45*, 5829–5837. <https://doi.org/10.1093/nar/gkx203>.
- Le Noir, S., Laffleur, B., Carrion, C., Garot, A., Lecardeur, S., Pinaud, E., et al. (2017). The IgH locus 3' cis-regulatory super-enhancer co-opts AID for allelic transvection. *Oncotarget*, *8*, 12929–12940. <https://doi.org/10.18632/oncotarget.14585>.
- Lewis, E. (1954). The theory and application of a new method of detecting chromosomal rearrangements in *Drosophila melanogaster*. *The American Naturalist*, *88*, 225–239.
- Li, W., Notani, D., & Rosenfeld, M. G. (2016). Enhancers as non-coding RNA transcription units: Recent insights and future perspectives. *Nature Reviews. Genetics*, *17*, 207–223. <https://doi.org/10.1038/nrg.2016.4>.
- Lieberman, R., Giannini, S. L., Birshtein, B. K., & Eckhardt, L. A. (1991). An enhancer at the 3' end of the mouse immunoglobulin heavy chain locus. *Nucleic Acids Research*, *19*, 933–937. <https://doi.org/10.1093/nar/19.4.933>.
- Madisen, L., & Groudine, M. (1994). Identification of a locus control region in the immunoglobulin heavy-chain locus that deregulates c-myc expression in plasmacytoma and Burkitt's lymphoma cells. *Genes & Development*, *8*, 2212–2226. <https://doi.org/10.1101/gad.8.18.2212>.
- Manis, J. P., van der Stoep, N., Tian, M., Ferrini, R., Davidson, L., Bottaro, A., et al. (1998). Class switching in B cells lacking 3' immunoglobulin heavy chain enhancers. *The Journal of Experimental Medicine*, *188*, 1421–1431. <https://doi.org/10.1084/jem.188.8.1421>.
- Martin, O. A., Garot, A., Le Noir, S., Aldigier, J.-C., Cogné, M., Pinaud, E., et al. (2018). Detecting rare AID-induced mutations in B-lineage oncogenes from high-throughput sequencing data using the detection of minor variants by error correction method. *Journal of Immunology*, *201*, 950–956. <https://doi.org/10.4049/jimmunol.1800203>.
- Matthias, P., & Baltimore, D. (1993). The immunoglobulin heavy chain locus contains another B-cell-specific 3' enhancer close to the alpha constant region. *Molecular and Cellular Biology*, *13*, 1547–1553. <https://doi.org/10.1128/mcb.13.3.1547>.
- Maul, R. W., & Gearhart, P. J. (2010). AID and somatic hypermutation. *Advances in Immunology*, *105*, 159–191. [https://doi.org/10.1016/S0065-2776\(10\)05006-6](https://doi.org/10.1016/S0065-2776(10)05006-6).
- Medvedovic, J., Ebert, A., Tagoh, H., Tamir, I. M., Schwickert, T. A., Novatchkova, M., et al. (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. <https://doi.org/10.1016/j.immuni.2013.08.011>.
- Meng, F.-L., Du, Z., Federation, A., Hu, J., Wang, Q., Kieffer-Kwon, K.-R., et al. (2014). Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell*, *159*, 1538–1548. <https://doi.org/10.1016/j.cell.2014.11.014>.
- Michaelson, J. S., Giannini, S. L., & Birshtein, B. K. (1995). Identification of 3 $\alpha$ -hs4, a novel Ig heavy chain enhancer element regulated at multiple stages of B cell differentiation. *Nucleic Acids Research*, *23*, 975–981. <https://doi.org/10.1093/nar/23.6.975>.
- Morvan, C. L. (2003). The immunoglobulin heavy-chain locus hs3b and hs4 3' enhancers are dispensable for VDJ assembly and somatic hypermutation. *Blood*, *102*, 1421–1427. <https://doi.org/10.1182/blood-2002-12-3827>.



- Oudinet, C., Braikia, F.-Z., Dauba, A., & Khamlichi, A. A. (2020). Mechanism and regulation of class switch recombination by IgH transcriptional control elements. In *Advances in immunology* (pp. 89–137). Elsevier. <https://doi.org/10.1016/bs.ai.2020.06.003>.
- Park, S.-K., Xiang, Y., Feng, X., & Garrard, W. T. (2014). Pronounced cohabitation of active immunoglobulin genes from three different chromosomes in transcription factories during maximal antibody synthesis. *Genes & Development*, 28, 1159–1164. <https://doi.org/10.1101/gad.237479.114>.
- Peron, S., Laffleur, B., Denis-Lagache, N., Cook-Moreau, J., Tinguely, A., Delpy, L., et al. (2012). AID-driven deletion causes immunoglobulin heavy chain locus suicide recombination in B cells. *Science*, 336, 931–934. <https://doi.org/10.1126/science.1218692>.
- Peters, A., & Storb, U. (1996). Somatic hypermutation of immunoglobulin genes is linked to transcription initiation. *Immunity*, 4, 57–65. [https://doi.org/10.1016/s1074-7613\(00\)80298-8](https://doi.org/10.1016/s1074-7613(00)80298-8).
- Pinaud, E., Khamlichi, A. A., Le Morvan, C., Drouet, M., Nalesso, V., Le Bert, M., et al. (2001). Localization of the 3' IgH locus elements that effect long-distance regulation of class switch recombination. *Immunity*, 15, 187–199. [https://doi.org/10.1016/S1074-7613\(01\)00181-9](https://doi.org/10.1016/S1074-7613(01)00181-9).
- Qian, J., Wang, Q., Dose, M., Pruett, N., Kieffer-Kwon, K.-R., Resch, W., et al. (2014). B cell super-enhancers and regulatory clusters recruit AID tumorigenic activity. *Cell*, 159, 1524–1537. <https://doi.org/10.1016/j.cell.2014.11.013>.
- Rouaud, P., Vincent-Fabert, C., Fiancette, R., Cogné, M., Pinaud, E., & Denizot, Y. (2012). Enhancers located in heavy chain regulatory region (hs3a, hs1.2, hs3b, and hs4) are dispensable for diversity of VDJ recombination. *The Journal of Biological Chemistry*, 287, 8356–8360. <https://doi.org/10.1074/jbc.M112.341024>.
- Rouaud, P., Vincent-Fabert, C., Saintamand, A., Fiancette, R., Marquet, M., Robert, I., et al. (2013). The IgH 3' regulatory region controls somatic hypermutation in germinal center B cells. *The Journal of Experimental Medicine*, 210, 1501–1507. <https://doi.org/10.1084/jem.20130072>.
- Rozen, S., Skaletsky, H., Marszalek, J. D., Minx, P. J., Cordum, H. S., Waterston, R. H., et al. (2003). Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature*, 423, 873–876. <https://doi.org/10.1038/nature01723>.
- Saintamand, A., Rouaud, P., Garot, A., Saad, F., Carrion, C., Oblet, C., et al. (2015). The IgH 3' regulatory region governs  $\mu$  chain transcription in mature B lymphocytes and the B cell fate. *Oncotarget*, 6, 4845–4852. <https://doi.org/10.18632/oncotarget.3010>.
- Saintamand, A., Rouaud, P., Saad, F., Rios, G., Cogné, M., & Denizot, Y. (2015). Elucidation of IgH 3' region regulatory role during class switch recombination via germline deletion. *Nature Communications*, 6, 7084. <https://doi.org/10.1038/ncomms8084>.
- Saintamand, A., Vincent-Fabert, C., Garot, A., Rouaud, P., Oruc, Z., Magnone, V., et al. (2016). Deciphering the importance of the palindromic architecture of the immunoglobulin heavy-chain 3' regulatory region. *Nature Communications*, 7, 10730. <https://doi.org/10.1038/ncomms10730>.
- Saintamand, A., Vincent-Fabert, C., Marquet, M., Ghazzau, N., Magnone, V., Pinaud, E., et al. (2017). E $\mu$  and 3'RR IgH enhancers show hierarchic unilateral dependence in mature B-cells. *Scientific Reports*, 7, 442. <https://doi.org/10.1038/s41598-017-00575-0>.
- Santos, J. M., Braikia, F.-Z., Oudinet, C., Dauba, A., & Khamlichi, A. A. (2019). Two modes of *cis*-activation of switch transcription by the IgH superenhancer. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 14708–14713. <https://doi.org/10.1073/pnas.1902250116>.
- Sepulveda, M., Garrett, F., Pricewhelan, A., & Birshstein, B. (2005). Comparative analysis of human and mouse 3' Regulatory regions identifies distinctive structural features. *Molecular Immunology*, 42, 605–615. <https://doi.org/10.1016/j.molimm.2004.09.006>.



- Skaletsky, H., Kuroda-Kawaguchi, T., Minx, P. J., Cordum, H. S., Hillier, L., Brown, L. G., et al. (2003). The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature*, *423*, 825–837. <https://doi.org/10.1038/nature01722>.
- Stavnezer, J., Björkman, A., Du, L., Cagigi, A., & Pan-Hammarström, Q. (2010). Mapping of switch recombination junctions, a tool for studying DNA repair pathways during immunoglobulin class switching. *Advances in Immunology*, *108*, 45–109. <https://doi.org/10.1016/B978-0-12-380995-7.00003-3>.
- Thomas-Claudepierre, A.-S., Schiavo, E., Heyer, V., Fournier, M., Page, A., Robert, I., et al. (2013). The cohesin complex regulates immunoglobulin class switch recombination. *The Journal of Experimental Medicine*, *210*, 2495–2502. <https://doi.org/10.1084/jem.20130166>.
- Vian, L., Pękowska, A., Rao, S. S. P., Kieffer-Kwon, K.-R., Jung, S., Baranello, L., et al. (2018). The energetics and physiological impact of cohesin extrusion. *Cell*, *173*, 1165–1178.e20. <https://doi.org/10.1016/j.cell.2018.03.072>.
- Vincent-Fabert, C., Fiancette, R., Pinaud, E., Truffinet, V., Cogne, N., Cogne, M., et al. (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. <https://doi.org/10.1182/blood-2010-01-264689>.
- Vincent-Fabert, C., Truffinet, V., Fiancette, R., Cogné, N., Cogné, M., & Denizot, Y. (2009). Ig synthesis and class switching do not require the presence of the hs4 enhancer in the 3' IgH regulatory region. *Journal of Immunology*, *182*, 6926–6932. <https://doi.org/10.4049/jimmunol.0900214>.
- Volpi, S. A., Verma-Gaur, J., Hassan, R., Ju, Z., Roa, S., Chatterjee, S., et al. (2012). Germine 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. *Journal of Immunology*, *188*, 2556–2566. <https://doi.org/10.4049/jimmunol.1102763>.
- Wuerffel, R., Wang, L., Grigera, F., Manis, J., Selsing, E., Perlot, T., et al. (2007). S-S synapsis during class switch recombination is promoted by distantly located transcriptional elements and activation-induced deaminase. *Immunity*, *27*, 711–722. <https://doi.org/10.1016/j.immuni.2007.09.007>.
- Yamane, A., Resch, W., Kuo, N., Kuchen, S., Li, Z., Sun, H., et al. (2011). Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nature Immunology*, *12*, 62–69. <https://doi.org/10.1038/ni.1964>.
- Yan, C. T., Boboila, C., Souza, E. K., Franco, S., Hickernell, T. R., Murphy, M., et al. (2007). IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature*, *449*, 478–482. <https://doi.org/10.1038/nature06020>.
- Zhang, X., Yoon, H. S., Chapdelaine-Williams, A. M., Kyritsis, N., & Alt, F. W. (2021). Physiological role of the 3'IgH CBEs super-anchor in antibody class switching. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(3). <https://doi.org/10.1073/pnas.2024392118>.
- Zhang, X., Zhang, Y., Ba, Z., Kyritsis, N., Casellas, R., & Alt, F. W. (2019). Fundamental roles of chromatin loop extrusion in antibody class switching. *Nature*, *575*, 385–389. <https://doi.org/10.1038/s41586-019-1723-0>.
- Zhang, Y., Zhang, X., Ba, Z., Liang, Z., Dring, E. W., Hu, H., et al. (2019). The fundamental role of chromatin loop extrusion in physiological V(D)J recombination. *Nature*, *573*, 600–604. <https://doi.org/10.1038/s41586-019-1547-y>.
- Zhou, J., Ashouian, N., Delepine, M., Matsuda, F., Chevillard, C., Riblet, R., et al. (2002). The origin of a developmentally regulated Igh replicon is located near the border of regulatory domains for Igh replication and expression. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 13693–13698. <https://doi.org/10.1073/pnas.212392399>.



**Figure 20 : Représentation de l'organisation du noyau en territoires chromosomiques et compartimentation A/B**  
Chaque territoire de couleur correspond à un territoire chromosomique différent. La compartimentation A/B est également représenté avec le compartiment A en bleu et le compartiment B en rose

## Partie IV. Organisation du noyau et architecture de la chromatine

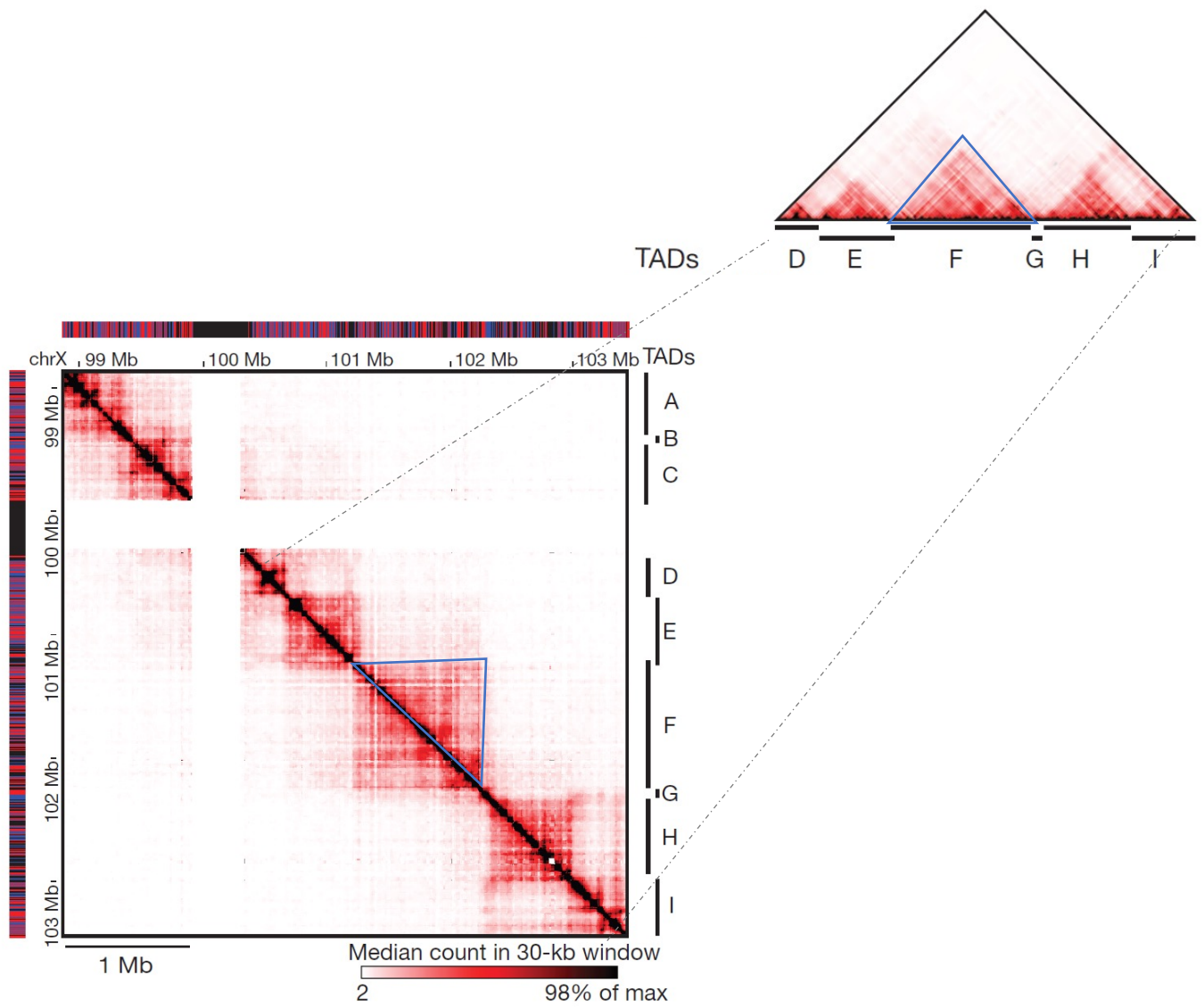
---

Le noyau isole l'information génétique du reste des organites de la cellule. Il arbore une organisation particulièrement précise nécessaire à la régulation de l'expression des gènes. Dans cette partie les différents niveaux d'organisation du noyau seront abordés dans un contexte général puis de façon appliquée au lymphocyte B.

### IV.1. Territoires chromosomiques et compartiments A et B

Dans le noyau interphasique, les chromosomes (donc les gènes) ne sont pas répartis de façon fortuite. Les chromosomes sont répartis dans des zones du noyau appelées territoire chromosomique (TC) (**Figure 20**) (Szczepińska et al., 2019). Le nombre de TC est propre à chaque type cellulaire et stade de différenciation (Bonev et al., 2017; Hnisz et al., 2013). Les TC constituent un aspect important de l'architecture du noyau. La distribution radiale des chromosomes au sein du noyau serait liée à leur densité génique. En effet, il a été démontré que, dans des lymphocytes humains, les chromosomes avec une forte densité de gènes ont une localisation plutôt centrale dans le noyau alors que les chromosomes avec une faible densité génique sont, quant à eux, positionnés à la périphérie nucléaire (Cremer et al., 2003, 2001; Croft et al., 1999). Il a été montré que l'organisation des TC chez les rongeurs dépend également de la densité génique des chromosomes (Mayer et al., 2005). De la même façon, il semblerait que les loci et gènes répliqués de façon précoce occupent une position centrale au sein du noyau alors que les gènes répliqués de façon tardive et peu exprimés se localisent, à l'inverse, plus en périphérie (Grasser et al., 2008; Takizawa et al., 2008).

Au-delà des TC, le génome est divisé en deux compartiments distincts arbitrairement nommés A et B (**Figure 20**). Le compartiment A est associé à une chromatine plus ouverte, transcriptionnellement active. Au sein du compartiment B, une fréquence de contact interne plus élevée traduit une compaction de la chromatine plus importante et donc transcriptionnellement moins active. De façon globale, le compartiment A est associé à l'euchromatine et le compartiment B à l'hétérochromatine. De manière analogue, le compartiment A est associé à une réplication précoce alors que le compartiment B se réplique de façon tardive (Ryba et al., 2010). Alors qu'ils peuvent varier selon les types cellulaires, les compartiments A et B sont stables entre les cellules de même type (Stevens et al., 2017). Les compartiments A et B présentent des sous-compartiments caractérisés par des marques chromatiniennes distinctes. Ainsi, les sous-compartiments A1 et A2 présentent des marques actives de la chromatine telles que H3K36me3, H3K79me3, H3K27ac et H3K4me1. Plus spécifiquement, le sous-compartiment A2 est défini par la présence de la marque inactive H3K9me3 et contient également des gènes plus longs que le sous-compartiment A1 (Rao et al., 2014). Le compartiment B est lui subdivisé en trois sous-compartiments : B1, B2 et B3. Le sous-compartiment B1 possède les caractéristiques de



**Figure 21 : Exemple de carte de contacts révélant la présence de TAD**

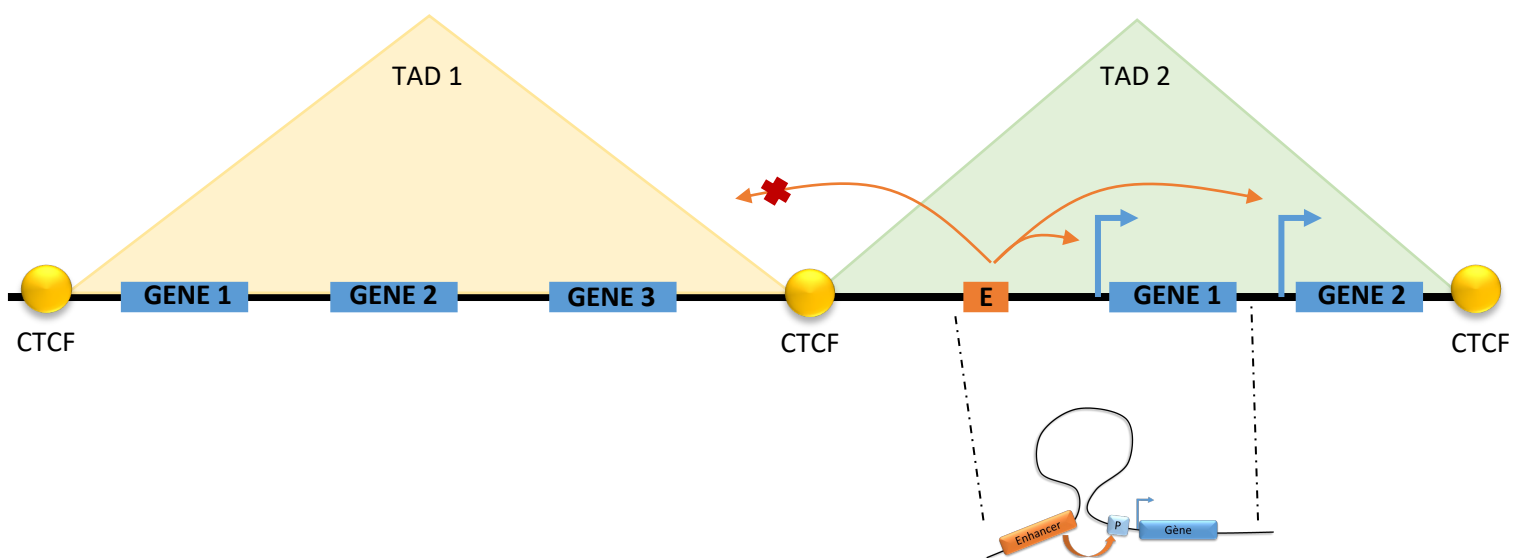
Cette figure est une carte de contacts réalisée à partir de données de Hi-C issu de Nora *et al*, 2012  
 Le triangle bleu met en évidence un des TAD

l'hétérochromatine facultative avec les marques H3K27me3 enrichie et H3K36me3 absente. Le sous-compartiment B2 est associé à l'hétérochromatine péri-centromérique alors que le sous-compartiment B3 coïncide avec l'hétérochromatine constitutive située à proximité de la lamine nucléaire (Rao et al., 2014). Les contacts entre gènes, à longue distance voir même interchromosomiques, sont favorisés au sein d'un même compartiment (A/A et B/B) alors que ces interactions physiques sont quasi inexistantes entre compartiments A et B (Lieberman-Aiden et al., 2009).

Les compartiments A et B constituent un niveau supérieur d'organisation nucléaire. Dans les paragraphes suivants, des niveaux supplémentaires d'organisation seront décrits.

#### **IV.2. TAD: "Topologically Associated Domains"**

Les compartiments A et B, précédemment décrits, contiennent des domaines qui arborent des hautes fréquences de contact interne. Ces domaines portent le nom de TAD (pour « *Topologically Associated Domains* »), ils ont été découverts suite à l'émergence de nouvelles techniques de biologie moléculaire associées au séquençage haut débit telles que le Hi-C (Dixon et al., 2012; Nora et al., 2012; Sexton et al., 2012). Les TAD sont des unités structurales tridimensionnelles dans lesquelles deux régions d'ADN distantes de plusieurs centaines de pb peuvent interagir entre elles. Les TAD sont des structures très conservées au cours du développement embryonnaire et entre différents types cellulaires. Dans les lignages dérivés de cellules souches embryonnaires, les TAD évoluent peu alors que la compartimentation A/B s'opère et arbore une certaine plasticité : certains gènes transitent du compartiment A vers le compartiment B et inversement (Dixon et al., 2015). Le génome murin est composé d'environ 2000 TAD ayant une taille variant d'une dizaine de kb à quelques Mb avec une taille médiane de 880 kb (Dixon et al., 2012). Chez la souris, l'ensemble des TAD couvre plus de 90% du génome alors que les 10% restant semblent ne pas être organisés mais plutôt distribués de façon aléatoire dans le noyau. Sur des cartes de contact issues de données de Hi-C, ces TAD sont visualisés sous la forme de triangles, traduisant des fréquences de contact élevées de la chromatine située à l'intérieur même de régions isolées du reste du génome (**Figure 21**) (Dixon et al., 2012; Rao et al., 2014). Les TADs sont délimités en 5' et en 3' par la présence de régions isolatrices (ou TAD border) qui arborent une concentration importante en sites de fixation pour la protéine CTCF (Dixon et al., 2012). La suppression de la protéine CTCF affecte essentiellement les contacts intra-TAD. La perte de ce facteur peut aussi induire une augmentation de la perméabilité des frontières, ce qui a pour conséquence une augmentation des contacts chromatiniens inter-TAD (Nora et al., 2017). La bonne conformation des TAD requiert donc la présence de la protéine CTCF (Hansen et al., 2017; Rowley and Corces, 2018; Zuin et al., 2014).



**Figure 22 : Schématisation des fonctions des TAD**

Deux TAD différents sont représentés. Le TAD 1 est tardivement répliqué et le TAD 2 est répliqué de façon précoce. L'effet de l'enhancer E est limité au TAD 2 où il contribue à l'expression des gènes présents dans ce TAD alors qu'il ne peut pas moduler l'expression des gènes du TAD 1.

Un exemple de boucle entre l'enhancer E et le promoteur d'un gène cible, permise par la facilitation des contacts au sein du TAD.

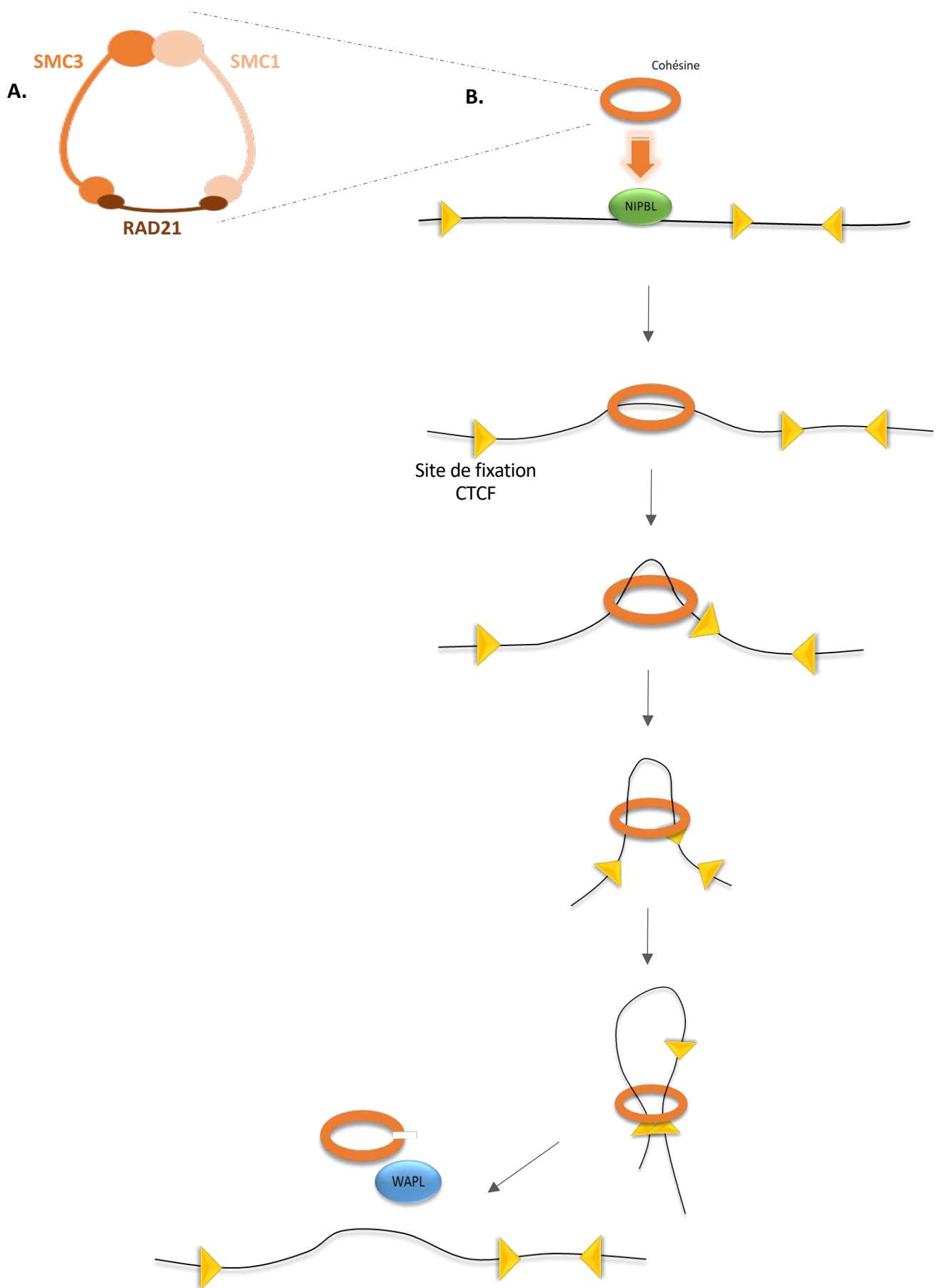
La fonction exacte des TAD n'est pas encore complètement décrite mais ces derniers conditionnent des niveaux de régulation à part entière dans certains processus biologiques, notamment la régulation de l'expression des gènes (**Figure 22**). En fonction du type cellulaire ou du stade de différenciation de la cellule, un TAD peut osciller entre le compartiment A et le compartiment B, permettant respectivement l'expression ou la répression des gènes qu'il contient. La facilitation des contacts entre des régions d'ADN, d'un même TAD, rend notamment possible les interactions entre différents éléments *cis*-régulateurs, notamment des promoteurs et des enhancers, ces derniers étant parfois situés à plusieurs kb des gènes qu'ils régulent (Kagey et al., 2010; Seitan et al., 2013) (**Figure 22**). De même les enhancers présents au sein d'un TAD verront leurs effets sur d'autres gènes considérablement limités voir abrogés si ces derniers sont localisés dans un TAD différent (Wendt et al., 2008). Il est aussi décrit que les TAD sont des unités stables de régulation temporelle de la réplication : les domaines de réplication correspondent généralement à un TAD (Marchal et al., 2019; Pope et al., 2014). De par leurs diverses fonctions, les TAD constituent l'unité structurelle fondamentale de l'organisation du génome (Dekker and Heard, 2015; Dixon et al., 2016). Ces derniers contribueraient à délimiter l'action de certains mécanismes. Pour exemple, dans les lymphocytes B, la susceptibilité de certaines régions du génome à l'hypermutation somatique pourrait être apporté par des éléments particuliers appelés *DIVAC* dont l'action serait cantonnée au sein de TAD (Senigl et al., 2019a).

L'étude des cartes de Hi-C révèle également la présence de boucles plus petites, appelées aussi sous-TAD (Rao et al., 2014). Ces sous-unités topologiques sont, quant à elle, moins bien conservées entre les cellules d'une même population et suggèrent donc une certaine spécificité tissulaire et fonctionnelle. Bien que leur existence soit admise, il est néanmoins possible d'imaginer que les sous-TAD détectés ne détiennent pas tous une fonction biologique réelle : leur présence sur les cartes de contact pourrait correspondre à un stade intermédiaire (TAD en cours de formation) dans une partie de la population cellulaire (Chang et al., 2020). La formation des TAD et des boucles dans les sous-TAD nécessite l'action d'un complexe multi-protéique : la cohésine (Sofueva et al., 2013; Zuin et al., 2014).

### **IV.3. Mécanisme d'extrusion de boucle**

#### **IV.3.1. Le modèle de l'extrusion de boucle**

Le complexe de la cohésine inclut deux sous-unités ATPases nommées SMC1 et SMC3 (SMC pour *Structural Maintenance of Chromosomes*). Ces deux sous-unités sont liées par une tierce protéine : Rad21. L'association de SMC1, SMC3 et Rad21 confère à la cohésine une forme en anneau (**Figure 23A**) (Gruber et al., 2003). D'autres protéines peuvent interagir avec la cohésine afin de réguler son association ou sa dissociation avec l'ADN. C'est le cas de NIPBL et WAPL, respectivement impliquées



**Figure 23 : Représentation schématique de la cohésine et du mécanisme d'extrusion de boucle**

A. La cohésine est un complexe tri-protéique composé de SMC1, SMC3 et RAD21

B. La cohésine est chargée sur l'ADN grâce à NIPBL et progresse sur l'ADN en formant une boucle qui passe dans son anneau jusqu'à la rencontre de deux sites CTCF (représentés par les triangles jaunes) convergents.

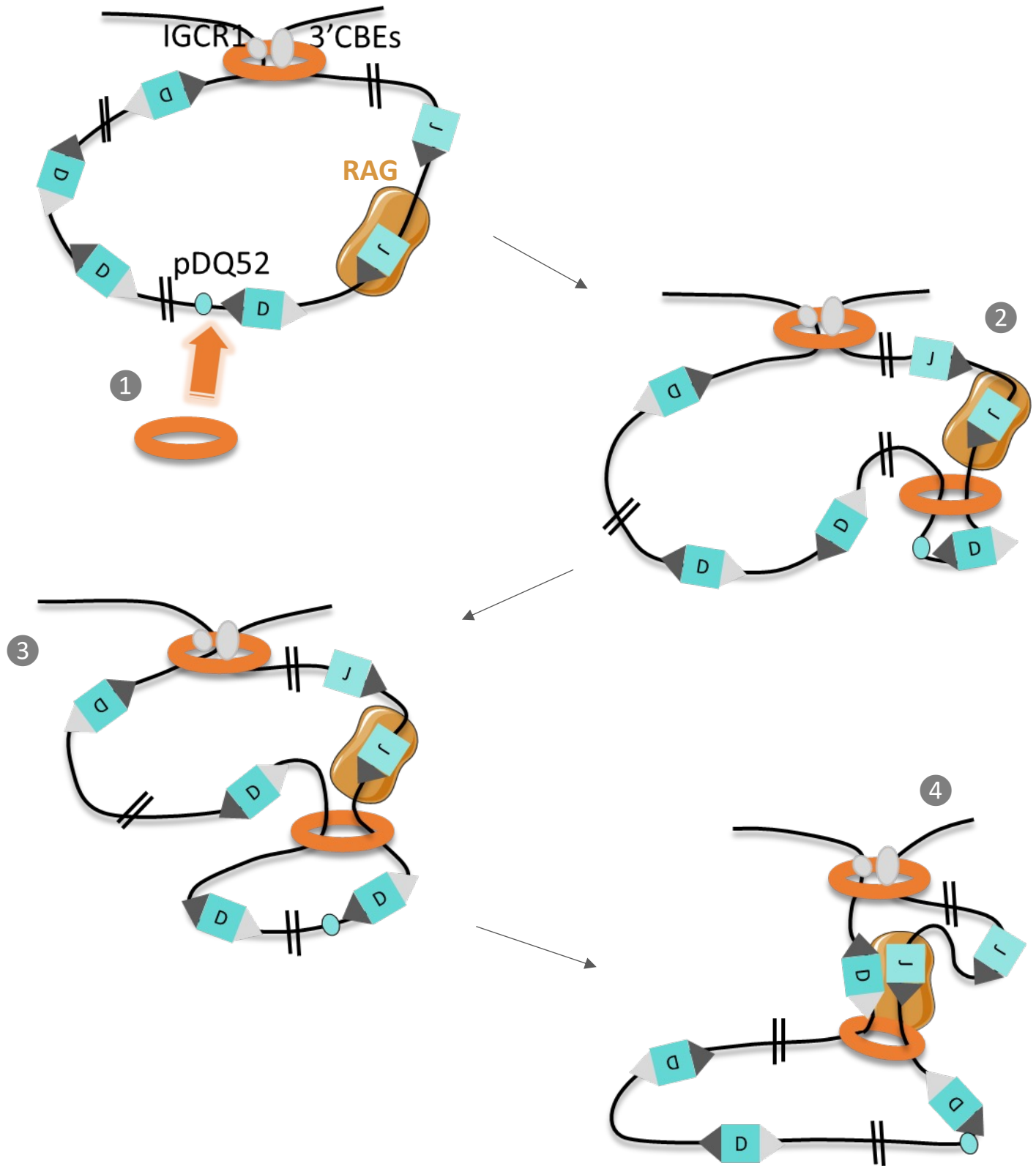


dans l'association et la dissociation du complexe de la cohésine à l'ADN. La présence de ces deux acteurs semble essentielle à l'extrusion de boucle et par conséquent à la formation des TAD. En effet, la délétion de NIPBL chez la souris entraîne la dissolution des TAD (Schwarzer et al., 2017) par manque d'association de la cohésine avec l'ADN alors que la déplétion de WAPL augmente le temps de résidence de la cohésine sur l'ADN (Gandhi et al., 2006; Kueng et al., 2006), favorisant la formation de boucle de plus grande taille (Gassler et al., 2017; Haarhuis et al., 2017; Wutz et al., 2017).

Le complexe tri-protéique de la cohésine s'associe à l'ADN grâce à NIPBL (Ciosk et al., 2000; Higashi et al., 2020; Shi et al., 2020). Grâce à sa forme annulaire, la cohésine encercle l'ADN et progresse sur celui-ci le plus souvent de façon bidirectionnelle, permettant le passage de la double hélice en son sein. Ceci mène à la formation d'une boucle de plus en plus grande. Elle progresse de cette façon jusqu'à la rencontre de deux sites CTCF convergents (**Figure 23B**) (Nanni et al., 2020; Rao et al., 2014; Vietri Rudan et al., 2015). L'inversion d'un site CTCF est suffisante pour induire la disparition d'une boucle ou l'apparition d'une boucle de plus grande taille (de Wit et al., 2015). Cependant, des sites CTCF non convergents peuvent tout de même faire office de faible point d'ancrage (Rowley and Corces, 2018). Ce mécanisme progressif explique la présence de boucles intra-TAD observées sur les cartes de contact. De même, la cohésine peut progresser de façon unidirectionnelle depuis un site CTCF convergent jusqu'à un autre, permettant la formation de boucles plus ou moins grandes. La formation de boucle ne dépendrait pas uniquement de la présence de sites CTCF : il est également possible que la transcription soit à l'origine d'un arrêt de la cohésine (Bonev et al., 2017; X. Zhang et al., 2019, 2019). Effectivement, il a été montré que la cohésine a tendance à s'accumuler de façon plus importante entre des gènes dont la transcription est convergente, c'est-à-dire à l'endroit où la transcription sens et antisens se rencontrent (Busslinger et al., 2017; Lengronne et al., 2004). De plus, de nombreuses frontières de TAD contiennent des gènes de ménage, ce qui étaye l'idée d'une corrélation entre la transcription active d'un gène et sa capacité à former une frontière de TAD indépendamment de CTCF (Dixon et al., 2012). Cependant, la transcription à elle seule pourrait ne pas être suffisante pour établir une frontière de TAD : la présence de facteurs de transcription contribuerait à perturber/stopper la progression de la cohésine sur l'ADN (Bonev et al., 2017).

#### **IV.3.2. L'extrusion de boucle chez le lymphocyte B**

Il a été montré par l'équipe de F. Alt que le mécanisme d'extrusion de boucle joue un rôle primordial chez le lymphocyte B. En effet, ce mécanisme participerait activement aux événements de recombinaison aussi bien lors de la phase précoce que lors de la phase tardive de l'ontogénie B que sont respectivement les recombinaisons V(D)J et CSR, détaillées dans la partie II.



**Figure 24 : Mécanisme d'extrusion de boucle pendant la recombinaison VDJ**

(1) La cohésine est chargée au niveau du promoteur pDQ52. (2) Elle progresse de façon bidirectionnelle jusqu'à la rencontre du complexe RAG. (3) Elle progresse ensuite unidirectionnellement sur l'ADN et permet ainsi le scan de l'ADN par RAG. (4) La cohésine s'arrête et permet de positionner correctement les deux segments (D et J<sub>H</sub>) à recombinaison.

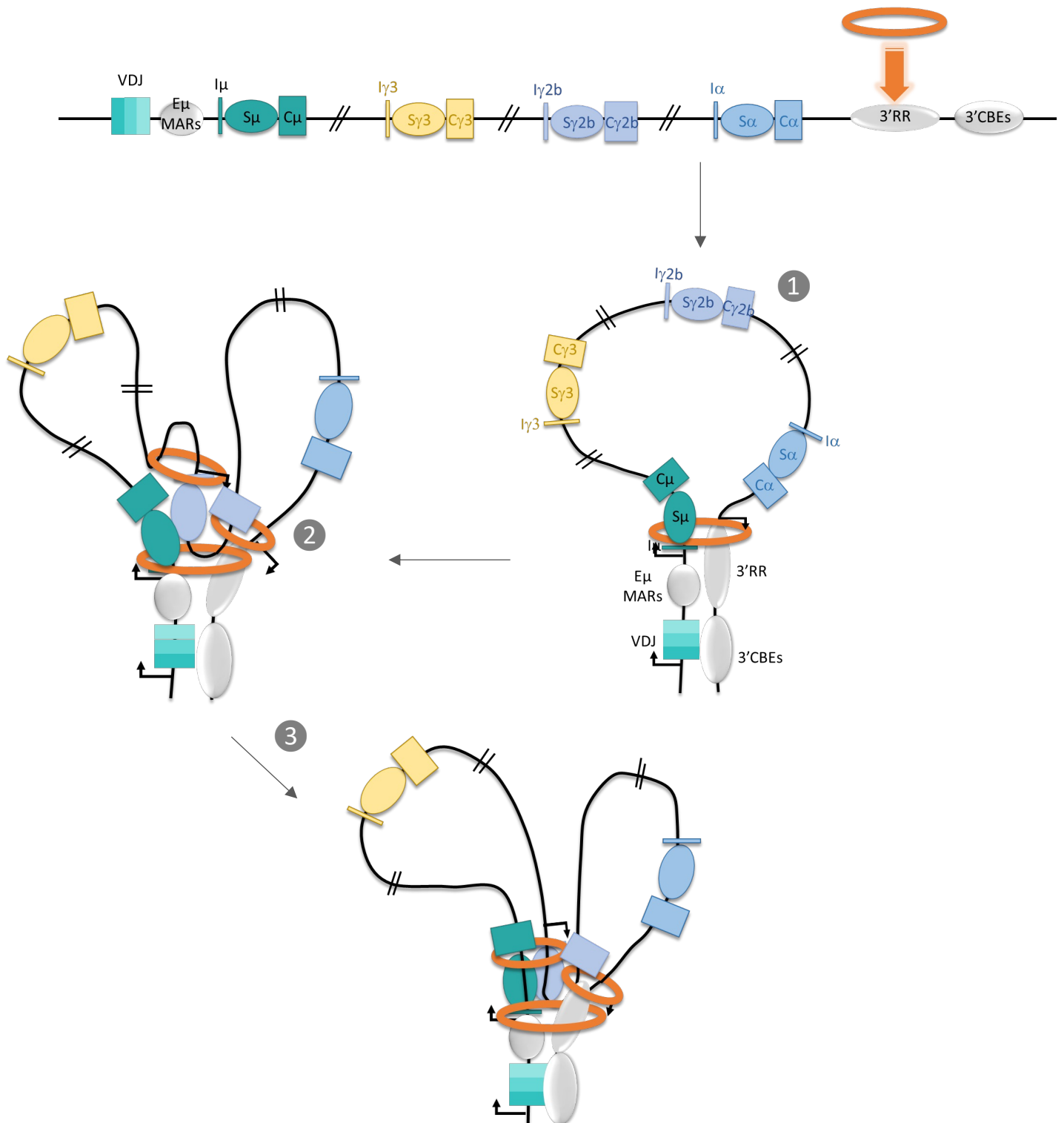
#### IV.3.2.1. L'extrusion de boucle pendant la recombinaison V(D)J

Les segments  $V_H$  s'étalent sur une région d'un peu plus de 2Mb. Cette région  $V_H$  contient de nombreux sites de fixation de CTCF dont l'orientation est convergente vis-à-vis des CBEs localisés en aval de la région 3' du locus. A partir du stade de cellule pro-B, des interactions entre ces sites convergents sont détectées ; traduisant alors une contraction du locus *IgH* médiée par le mécanisme d'extrusion de boucle (Ba et al., 2020; Dai et al., 2021). Pendant la recombinaison VDJ, d'autres boucles sont formées par ce même mécanisme. Ce mécanisme est initié au niveau du centre de recombinaison, au sein duquel une grande concentration de complexes RAG est observée (Y. Zhang et al., 2019). Par exemple, pendant le réarrangement  $D_H$  vers  $J_H$ , le complexe de la cohésine est chargé au niveau du promoteur *pDQ52* grâce à NIPBL (Figure 24, n°1) ; l'ADN est ensuite extrudé bidirectionnellement jusqu'à ce que la cohésine soit bloquée par le complexe RAG qui a reconnu une séquence RSS adjacente à un segment  $J_H$  (Figure 24, n°2). La suite de l'extrusion se fait alors de façon unidirectionnelle (Figure 24, n°3). Le déroulement de la région  $D_H$  à travers l'anneau que forme la cohésine, permet au complexe RAG de scanner les différents segments  $D_H$ . Enfin, l'extrusion est stoppée lorsque la progression de la cohésine est empêchée par le complexe RAG qui reconnaît la séquence RSS adjacente au segment  $D_H$  à recombiner (Figure 24, n°4). De cette façon, les segments à recombiner sont correctement orientés, le complexe RAG peut alors initier les CDBs.

D'autres boucles sont initiées lors de la recombinaison VDJ, il s'agit de boucles plus petites qui apparaissent notamment entre différents éléments *cis*-régulateurs dispersés au sein de la région VDJ. Ainsi, l'occurrence de boucles au sein de la portion  $V_H$  assure l'utilisation de l'ensemble des segments  $V_H$ , qu'ils soient distaux, intermédiaires ou proximaux (Medvedovic et al., 2013). En effet, la perte de ces boucles entraîne une modification du répertoire  $V_H$  avec une diminution de la contraction du locus *IgH* et par conséquent une diminution de l'utilisation des segments  $V_H$  distaux (Guo et al., 2011). Ce dernier effet est notamment apprécié en absence de PAX5, protéine clé du développement lymphocytaire B précoce. Dans les faits, la délétion de PAX5 entraîne une augmentation de l'expression de *Wapl*, gène codant la protéine du même nom ayant pour fonction le déchargement de la cohésine de l'ADN (Hill et al., 2020).

#### IV.3.2.2. L'extrusion de boucle pendant la CSR

Les segments codants les régions constantes se trouvent dans une portion d'environ 250kb encadrée par les régions *EμMARs* et *3'RR*, comme décrit précédemment dans la partie I. Bien que ces régions soient dépourvues de sites CTCF, le groupe de Kenter a décrit, il y a une quinzaine d'années, la formation de boucle entre les enhancers majeurs du locus. Ces boucles reliant *EμMARs* et la *3'RR* associent les régions donneuses et acceptrices de switch à la suite de l'activation des lymphocytes B (Wuerffel et al., 2007). Ce n'est que bien plus tard, en 2019, que la formation de ces synapses de switch



**Figure 25 : Mécanisme d'extrusion de boucle pendant la CSR**

(1) La cohésine est chargée au niveau de la 3'RR, qui et extrude l'ADN unidirectionnellement jusqu'à I $\mu$  (2), où la transcription sens et antisens, qui sont constitutives et donc présentes dans des B au repos, est convergente. Suite à une stimulation antigénique, la transcription est initiée au niveau du I cible, par exemple ici Ig2b. Une autre extrusion s'opère donc entre la région g2b et la 3'RR ainsi qu'entre g2b et  $\mu$  (3). La séquentialité de ces événements conduit à la formation de deux boucles mettant en contact les régions switch à recombinaison (4).

a été associée au mécanisme d'extrusion de boucle par la cohésine ; un mécanisme décrit comme fondamental à la CSR (X. Zhang et al., 2019). La boucle « basale » entre la 3'RR et la région *EμMARs* dans les cellules B au repos est médiée par la cohésine et forme le centre de CSR (CSRC pour *CSR Center*) (**Figure 25**, n°1) (X. Zhang et al., 2019). La formation de cette boucle, comme l'ensemble des boucles formées par le mécanisme d'extrusion, est dynamique dans le temps ; cette interaction est donc susceptible d'être défaite et reformée plusieurs fois dans les cellules B naïves. Lors de l'activation de la cellule B, la transcription initiée à partir des promoteurs germinaux des régions donneuses et acceptrices va permettre le recrutement de NIPBL et par conséquent le chargement de la cohésine, (**Figure 25**, n°2). Il s'en suit la formation de sous-boucles, au cœur du CSRC, formé par la boucle basale 3'RR-*EμMARs*, qui aura pour conséquence le rapprochement des régions switch donneuses nouvellement transcrites ; ce mécanisme exclut alors les régions switch non transcrites du CSRC (**Figure 25**, n°3) (X. Zhang et al., 2019). En complément du rapprochement physique des régions switch donneuse et acceptrice, le mécanisme d'extrusion de boucle est aussi impliqué dans la résolution des CDB (Arnould et al., 2021) qui se produisent lors de la CSR. Ainsi, le clivage des régions switch donneuse et acceptrice par AID s'opère dès lors que ces régions se trouvent maintenues par la cohésine. Ce mécanisme permettrait aux régions switch de rester à proximité l'une de l'autre et faciliterait le recrutement des acteurs de la voie NHEJ, limitant ainsi la survenue de translocation.

La mise en évidence du mécanisme d'extrusion de boucle dans la cellule B permet aussi d'expliquer les phénomènes de switch séquentiel (Shen et al., 2021) : par exemple les recombinaisons de *Sμ* vers *Sγ1* puis *Sμ/Sγ1* vers *Sε* ou encore de *Sμ* vers *Sγ3* puis *Sμ/Sγ3* vers *Sγ2b*, observées notamment lors de stimuli par l'IL-4 et le LPS, respectivement. En effet, il est facile d'imaginer que la cohésine progresse jusqu'à la première région acceptrice transcrite, puis, une fois la recombinaison effectuée, progresse jusqu'au second site activement transcrit, correspondant à la deuxième région acceptrice.

Un avantage incontestable de la formation de boucle par le mécanisme d'extrusion est que ce dernier assure l'occurrence d'évènements de délétion lors des recombinaisons de CSR, prévenant tout risque d'évènement non-productif d'inversion du locus *IgH*. La CSR délétionnelle s'accompagne de l'élimination de la portion génique qui était située entre les régions switch à recombinaison alors que la CSR inversionnelle conduit à l'inversion du fragment situé entre les deux régions à recombinaison sans qu'il ne soit pour autant éliminé ; ce dernier évènement conduit à la perte de l'expression d'une chaîne lourde fonctionnelle d'Ig.

Les éléments *cis*-régulateurs du locus *IgH*, notamment la 3'RR et la région *EμMARs*, seraient des sites de chargement de la cohésine et assureraient une certaine fonction de barrière en limitant la progression du mécanisme d'extrusion de boucle au-delà du CSRC (X. Zhang et al., 2019). Un tel mécanisme, assuré par les régions régulatrices, assurerait la survenue d'une CSR délétionnelle avec une jonction entre *Sμ* et une région switch acceptrice *Sx* située en aval. Cependant, l'implication de la

3'RR dans la formation de ces boucles est remise en question par l'équipe de Pavri qui propose que la formation de la boucle basale est plutôt médiée par les sites *CBE* localisés en 3' de la 3'RR (Costea et al., 2023). Néanmoins, nos travaux viennent en partie contredire cette hypothèse. Grâce à nos différents modèles mutant pour la 3'RR, nous montrons que les enhancers contenus dans cette région sont nécessaires à la formation des boucles pendant la CSR (**Article n°1** (Bruzeau et al., 2023)). En complément, le modèle présentant une délétion partielle des 3'CBEs (*hs567* KO) ne présente pas de défaut de CSR ni de formation de boucles (Volpi et al., 2012; Zhang et al., 2021). En revanche, la délétion totale de la région 3'CBEs entraîne bien un défaut de switch vers tous les isotypes exceptés IgG1. Ce dernier point est en contradiction avec l'étude de Costea et ses collègues qui font mention de l'implication de la région 3'CBEs dans la formation de boucles en condition de stimulation LPS+IL4 (qui entraîne le switch vers IgG1).

#### **IV.4. Organisation du noyau chez le lymphocyte B et rôle des enhancers**

L'organisation du noyau ainsi que la formation de boucles de chromatine représentent un niveau à part entière de régulation du développement du lymphocyte B et contribuent à l'orchestration ordonnée des événements de recombinaison précoces et tardifs. La revue ci-après, que je signe en première auteure et publiée dans *Frontiers in Immunology*, décrit l'implication des éléments *cis*-régulateurs *d'Ig* pour l'adressage des loci *d'Ig* au sein du noyau ainsi que pour la formation de boucles pendant les événements de recombinaison (Bruzeau et al., 2022).



# Contribution of Immunoglobulin Enhancers to B Cell Nuclear Organization

Charlotte Bruzeau, Jeanne Cook-Moreau, Eric Pinaud and Sandrine Le Noir\*

UMR CNRS 7276, INSERM 1262 and Université de Limoges: Contrôle de la Réponse Immune B et des Lymphoproliférations, 2 Rue du Pr. Descottes, Limoges, France

## OPEN ACCESS

### Edited by:

Paolo Casali,  
The University of Texas Health Science  
Center at San Antonio, United States

### Reviewed by:

Jayanta Chaudhuri,  
Memorial Sloan Kettering Cancer  
Center, United States  
Uttiya Basu,  
Columbia University, United States

### \*Correspondence:

Sandrine Le Noir  
sandrine.le-noir@unilim.fr

### Specialty section:

This article was submitted to  
B Cell Biology,  
a section of the journal  
Frontiers in Immunology

Received: 17 February 2022

Accepted: 26 May 2022

Published: 24 June 2022

### Citation:

Bruzeau C, Cook-Moreau J, Pinaud E  
and Le Noir S (2022) Contribution of  
Immunoglobulin Enhancers to B Cell  
Nuclear Organization.  
Front. Immunol. 13:877930.  
doi: 10.3389/fimmu.2022.877930

B cells undergo genetic rearrangements at immunoglobulin gene (*Ig*) loci during B cell maturation. First *V(D)J* recombination occurs during early B cell stages followed by class switch recombination (CSR) and somatic hypermutation (SHM) which occur during mature B cell stages. Given that RAG1/2 induces DNA double strand breaks (DSBs) during *V(D)J* recombination and AID (Activation-Induced Deaminase) leads to DNA modifications (mutations during SHM or DNA DSBs during CSR), it is mandatory that *IgH* rearrangements be tightly regulated to avoid any mutations or translocations within oncogenes. *Ig* loci contain various *cis*-regulatory elements that are involved in germline transcription, chromatin modifications or RAG/AID recruitment. *Ig cis*-regulatory elements are increasingly recognized as being involved in nuclear positioning, heterochromatin addressing and chromosome loop regulation. In this review, we examined multiple data showing the critical interest of studying *Ig* gene regulation at the whole nucleus scale. In this context, we highlighted the essential function of *Ig* gene regulatory elements that now have to be considered as nuclear organizers in B lymphocytes.

**Keywords:** immunoglobuline genes, enhancers, nuclear organization, B lymphocytes, chromatin loops

## INTRODUCTION

To produce highly specific antibodies, B cells undergo genetic modifications of their immunoglobulin (*Ig*) genes. Among these events, *V(D)J* recombination takes place in the bone marrow during the early steps of B cell development and occurs in an antigen-independent manner. Mature B cells migrate towards secondary lymphoid organs and continue their differentiation once stimulated by antigens. This process integrates secondary beneficial DNA remodeling events including class switch recombination (CSR) and somatic hypermutation (SHM) but can also induce B cell death through locus suicide recombination (LSR), a detrimental rearrangement that abrogates surface B cell receptor expression (1). These events, all mediated by the activation-induced deaminase (AID) enzyme, characterize the late antigen-dependent phase of developing cells (Figure 1).

Nuclear organization and chromosome distribution in B-lineage cells have to be considered as important parameters for the control of all these events. Chromosomes are not randomly scattered in the nucleus; their positions change in a dynamic fashion especially during the B cell lifespan. Indeed, distinct organizational levels co-exist in the genome on nucleosomal (genes and loci),

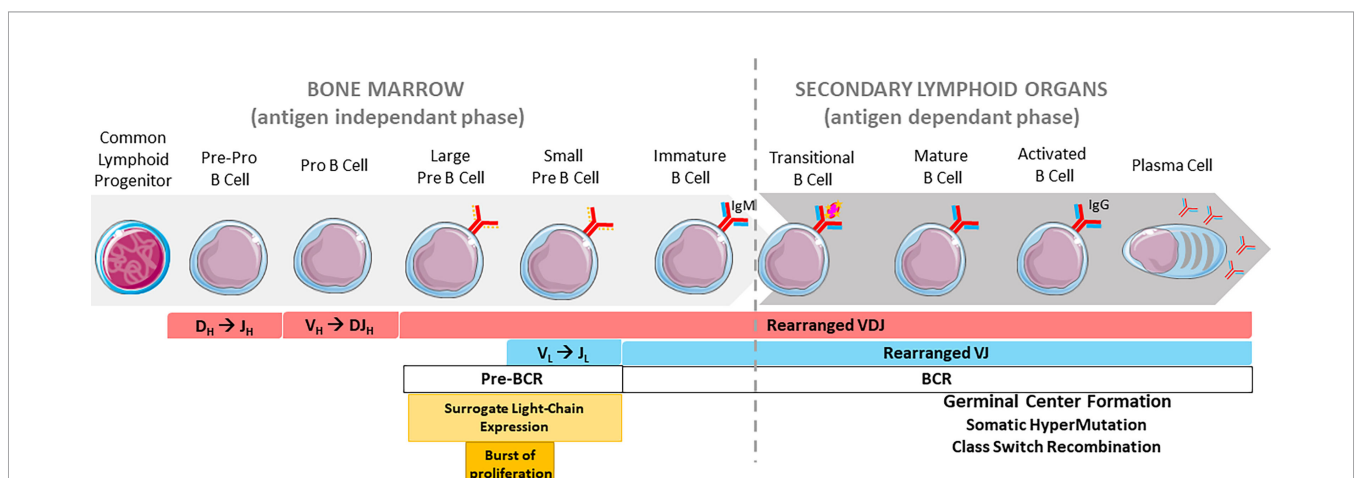
supranucleosomal (chromatin domains and compartments) and nuclear (chromosome territories) scales (2–4). In addition to histone mark patterns that reflect the nucleosomal context, the emergence of new molecular biology approaches led to the discovery of Topologically Associated Domains (TADs) and the dynamic loop extrusion model (5). Such methods, based on chromosome capture conformation (6) (3C, 4C and Hi-C), make it possible to evaluate additional levels of gene regulation at the supranucleosomal scale. This particular level of nuclear topology includes TAD structures, A and B chromatin compartments, DNA loops and interchromatin space. Gene transcription takes place in the A-euchromatin compartment whereas B-heterochromatin prevents it. In a simplified scheme, A and B compartments are respectively positioned at the center of nuclei and at the nuclear periphery (4). More recently, the Volk's group showed that heterochromatin and euchromatin, respectively defined as B and A compartments, are localized at the nucleus periphery, leaving the nucleus center devoid of nucleosomes (7). As a structural unit of genome organization, a TAD is a large chromosomal region in which the contact frequencies between genes or regulatory regions are higher than elsewhere in the genome. TADs themselves are subdivided into multiple sub-TAD structures (chromatin loops) that undergo dynamic cell-type specific connections. The mouse genome contains around 2000 TADs, each with an approximate mean size of 1 megabase (Mb) (8). Indeed, some dynamic processes drive chromatin regions into a free space termed the interchromatin compartment in order to permit gene segment interactions. Such contacts between gene portions occur either in active (A) or inactive (B) chromatin compartments and it is widely recognized that interactions take place within the same TAD. Among chromatin compartments, long-range homologous contacts (A–A or B–B) are largely favored over heterologous contacts (A–B) (9). Moreover, additional TAD interactions exist since chromosome portions are not only able to establish close contacts in *cis*, but also in *trans* with other chromosomes (10). Some

of these *trans* interactions have been documented in the case of olfactory receptor (11) and Th2 cytokines genes (12). At TAD extremities, TAD borders are enriched in CTCF (CCCTC-binding Factor) insulator protein, mediator complexes (MED1, MED12) as well as active histone marks (H3K4me3 and H3K36me3) (8). TAD borders display specific “insulating” features, preventing loci located on each side of this border to establish contacts (8).

By considering genome nuclear topology, these emerging models are particularly relevant for the tightly-regulated *Ig* gene loci. Most *Ig* gene regulation studies have so far been performed at the nucleosomal scale (epigenetic modifications and regulatory transcription of loci and gene segments). The increasing interest in understanding gene regulation at the whole nucleus scale prompted B cell scientists to revisit previous models at both supranucleosomal (DNA loops and TADs) and nuclear (chromosome territories and nuclear position) levels (13–22). To provide a clearer picture of how B cell development is tightly regulated by the nuclear location of *Ig* genes, including chromosome looping and loci positioning in the mouse, we will begin with an overview of *Ig* genes and their enhancers and then focus on the role of their main enhancers on 3D-nuclear organization. The relationship between *Ig* and respective enhancers will be discussed in this review.

## OVERVIEW OF IMMUNOGLOBULIN GENES AND THEIR REGULATORY ELEMENTS

In mice, *Ig* genes are encoded by three loci located on three distinct chromosomes. The immunoglobulin heavy chain (*IgH*) locus lies on chromosome 12. Immunoglobulin light (*IgL*) chain loci contain either kappa immunoglobulin light chain (*Igκ*) or lambda immunoglobulin light chain (*Igλ*) genes and are respectively located on chromosomes 6 and 16. In mice, at least 95% of B cells express *Igκ* light chains (13, 14).



**FIGURE 1** | Overview of B cell development. Chronological order of B cell stages in the bone marrow (left) and secondary lymphoid organs (right). The pre-BCR is represented by a continuous red line and a discontinuous orange line (surrogate light chain). The BCR is represented by continuous red and blue lines.



## Immunoglobulin Heavy (IgH) Chain Loci

The *IgH* locus spans approximately 3Mb, in its germline configuration, and contains various *cis*-regulatory regions (**Figure 2A**). From 5' to 3', the 5'*hs123ab* elements, are situated upstream from the first  $V_H$  segment (15). The intergenic  $V_H D_H$  region, located between the most distal  $D_H$  segment, *DFL16*, and the most proximal  $V_H$  segment,  $V_{H7183a.2,3}$ , contains six DNase I sensitive sites (*hs1* to 6) among which *hs4* and *hs5* are CTCF binding sites (17, 18). This set is also called Intergenic Control Region 1 (*IGCR1*) (16, 17). The promoter/enhancer *pDQ52*, is located just upstream from *DQ52* (18). The  $E\mu$ -MARs region, spanning about 1kb, is composed of a 220-base pair (bp) core enhancer element (*cE $\mu$* ) flanked by two matrix attachment regions (MARs) located between the last  $J_H$  exon and the  $S\mu$  region. Between the *C $\gamma$ 1* and *C $\gamma$ 2b* constant genes, two transcriptional enhancers *hRE1* and *hRE2* (19) form the  $\gamma$ 1E regulatory element (20). At the 3' extremity, the 3'*Regulatory Region* (3'*RR*), spanning more than 30 kb, is composed of four enhancers, called *hs3a*, *hs1.2*, *hs3b* and *hs4* (21). The central *hs1.2* enhancer is flanked by inverted repeated intervening sequences (*IRIS*) that form a 25kb-long quasi-palindrome (22). Downstream from this palindrome, *hs4* is the most distal element harboring an enhancer activity within the 3'*RR* (23). Although highly divergent in different species, *IRIS* sequences always stand as inverted copies on both sides of *hs1.2*, conserving a singular symmetry within the 3'*RR* (24). Downstream from the 3'*RR*, four *hs* elements (*hs5*, *hs6*, *hs7* and *hs8*) lie in a region containing ten CTCF Binding Elements (3'*CBEs*) (25, 26). This region acts as an insulator and delimits the 3' *IgH* TAD border (25, 26).

Our group and others developed numerous mouse models carrying deletions within the *IgH* regulatory elements that helped elucidate  $E\mu$  and 3'*RR* super-enhancer functions during early and late B cell development (20, 27–37).

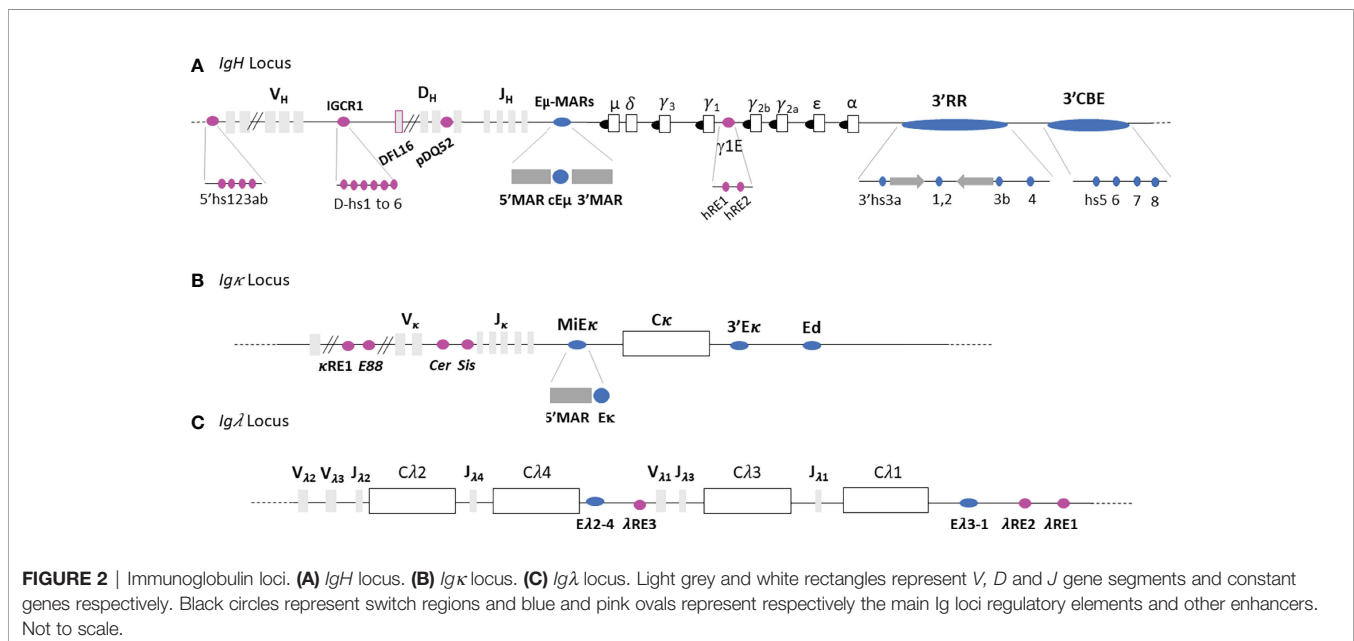
## Immunoglobulin Light Chain Loci: Ig $\kappa$ and Ig $\lambda$

The 3.2 Mb *Ig $\kappa$*  locus contains regulatory elements within the  $V\kappa$  segments (**Figure 2B**) which are partially homologous to those within the *IgH*:  $\kappa$ *RE1* (19) and *E88* (38). Six DNase I hypersensitive sites (*hs*), *hs1* to 6, are located in the intervening  $V\kappa$ - $J\kappa$  region (39, 40). *Hs1-2* forms *Cer* (Contracting element for recombination) and *hs3-6* comprises the *Sis* (Silencer in the intervening sequence) elements. A MAR-intronic  $E\kappa$  enhancer (*MiE $\kappa$* ) (41–43), located downstream from the last  $J\kappa$  segment, is composed of a 5' flanking matrix attachment region (MAR $\kappa$ ) and an intronic enhancer. The 3' $E\kappa$  enhancer region is situated downstream from the unique  $C\kappa$  gene (44). A third regulatory region, called *Ed* (due to its downstream location within the *Ig $\kappa$*  locus), is located distal to 3' $E\kappa$  (45).

In comparison to *IgH* and *Ig $\kappa$*  loci, the *Ig $\lambda$*  locus is smaller (200kb) and uniquely organized. It comprises four families which contain a pair of  $J_\lambda$  and  $C_\lambda$  segments with only three  $V_\lambda$  segments (46) (**Figure 2C**). The *Ig $\lambda$*  locus contains two main enhancers called  $E_{\lambda 2-4}$ , located between  $C_{\lambda 4}$  and  $V_{\lambda 1}$ , and  $E_{\lambda 3-1}$  located downstream from  $C_{\lambda 1}$  (47). Three supplementary elements featuring enhancer activity have been described:  $\lambda$ *RE3*,  $\lambda$ *RE2* and  $\lambda$ *RE1* (19).  $\lambda$ *RE3* lies between  $E_{\lambda 2-4}$  and  $V_{\lambda 1}$  while  $\lambda$ *RE1* and  $\lambda$ *RE2* are located close to the  $E_{\lambda 3-1}$  enhancer. Both *Ig $\lambda$*  enhancers,  $E_{\lambda 2-4}$  and  $E_{\lambda 3-1}$  are involved in transcription and *VJ* rearrangement regulation (47).  $\lambda$ *RE* elements, especially  $\lambda$ *RE1* and  $\lambda$ *RE3*, have been shown to potentiate the enhancer activity of  $E_{\lambda 3-1}$  and  $E_{\lambda 2-4}$ , in pro-B and plasma cells respectively (19).

## Immunoglobulin Joining Chain Loci: IgJ

Located on chromosome 5 in mice, the *IgJ* loci encode the joining peptide (J) chain which promotes active IgA (48) or IgM (49) secretion by ensuring efficient assembly of these Ig subtypes in plasma cells.



## NUCLEAR ORGANIZATION DRIVEN BY IG ENHANCERS AT EARLY B CELL STAGES

Early B cell development takes place in bone marrow where a lymphoid precursor progresses through different stages to the final immature stage. This progression is concomitant with  $V(D)J$  recombination. Once engaged in the B cell lineage after expression of B cell specific transcription factors (50), the common lymphoid precursor (CLP) differentiates into a pre-pro B and then pro-B cell that undergoes  $D_H-J_H$  rearrangement at the  $IgH$  locus (51). Once the  $DJ_H$  segment rearranges (52), the pro-B cell joins a  $V_H$  segment to the previously rearranged  $DJ_H$  segment and progresses to the large pre-B cell stage. At this stage, the  $IgH$  locus is completely rearranged and the cell expresses a pre-B Cell Receptor (pre-BCR) at its surface. The pre-BCR, indispensable for B cell development (53), is composed of a functional IgH chain linked to an invariant surrogate light chain, altogether associated with the  $Ig\alpha-Ig\beta$  transmembrane heterodimer signaling component. Pre-BCR signaling stops  $IgH$  rearrangement and triggers a burst of proliferation leading to the small pre-B cell stage and the occurrence of  $V_L-J_L$  rearrangements at  $IgL$  loci. Successful  $IgL$  chain rearrangement and production leads to membrane IgM expression and consequently a functional BCR (50), on the immature B cell. These cells then migrate towards secondary lymphoid organs and continue their maturation (Figure 1).

### IgH Sub-Nuclear Positioning and Chromatin Loops

$Ig$  loci positioning within the nucleus is dynamic throughout early B cell development and it has been clearly shown that

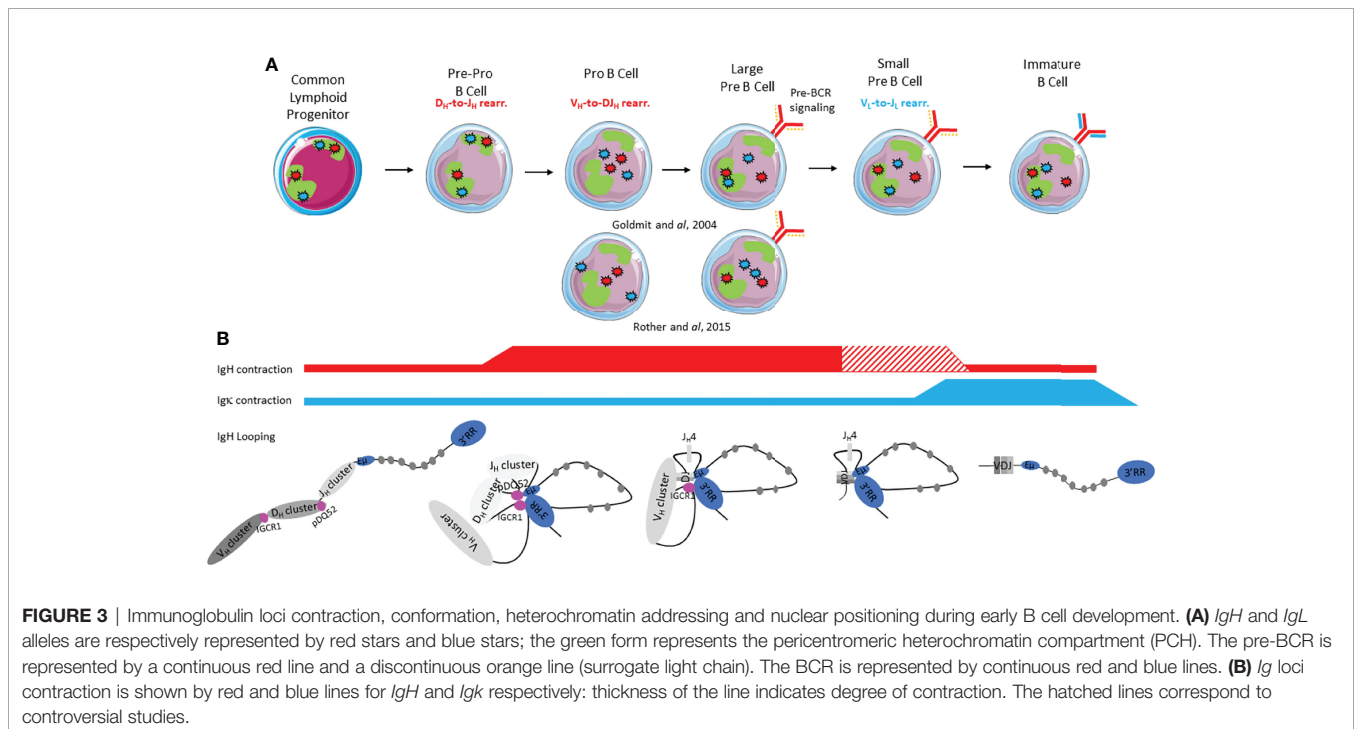
nuclear organization provides a critical level of regulation during  $V(D)J$  recombination, particularly for loci contraction/decontraction, nuclear positioning (center vs periphery) and heterochromatin addressing (Figure 3).

#### IgH Sub-Nuclear Positioning

Dynamic repositioning of the  $IgH$  locus throughout early stages is involved in  $IgH$  accessibility in order to correctly orchestrate  $V(D)J$  recombination (Figure 3A). In CLP and pre-pro-B cells,  $IgH$  loci are preferentially located at the nuclear periphery within a repressive compartment (54). While only few changes occur at immature stages (55), the pro-B cell stage undergoes, in contrast, global genome reorganization since around 20% of genes switch to A-B or B-A compartments (56). Indeed, both  $IgH$  alleles are relocated to the active A compartment to allow  $D_H-J_H$  recombination on both alleles. At the large pre-B cell stage, once  $VDJ_H$  recombination completed, pre-BCR signaling induces repositioning of one  $IgH$  allele into a repressive centromeric domain. Although pericentromeric recruitment has been associated with allelic exclusion (57), this nuclear location does not preclude transcriptional expression of the non-productive allele (58). Monoallelic  $IgH$  addressing to pericentromeric heterochromatin (PCH) remains until the immature B cell stage (57).

#### IgH Locus Contraction and Chromatin Loops

Beyond nuclear positioning,  $IgH$  allele contraction/decontraction, as well as TAD and loop reorganization represent an additional level of regulation, considered as critical for early B cell development (59). At the pre-pro B cell stage, both  $IgH$  alleles are decontracted (60) and at the pro-B cell stage both  $IgH$  loci



**FIGURE 3** | Immunoglobulin loci contraction, conformation, heterochromatin addressing and nuclear positioning during early B cell development. **(A)**  $IgH$  and  $IgL$  alleles are respectively represented by red stars and blue stars; the green form represents the pericentromeric heterochromatin compartment (PCH). The pre-BCR is represented by a continuous red line and a discontinuous orange line (surrogate light chain). The BCR is represented by continuous red and blue lines. **(B)**  $Ig$  loci contraction is shown by red and blue lines for  $IgH$  and  $IgL$  respectively: thickness of the line indicates degree of contraction. The hatched lines correspond to controversial studies.

contract to bring  $D_H$  et  $J_H$  clusters closer to promote  $D_H$  to  $J_H$  rearrangement (**Figure 3B**). 3C (Chromosome Capture Conformation) and 4C (Circular Chromosome Capture Conformation) experiments revealed that *IgH* locus compaction was mediated by large loops in which  $E\mu$  interacts with  $3'RR$ , *IGCR1* and  $V_H$  regions ( $5' V_H7183$  and  $5' V_H558$ ) (60). During the transition from pro-B to large pre-B cell stages, the entire  $V_H$  region is brought into juxtaposition with  $D_H$  elements due to extensive *IgH* locus contraction (60). This contraction is completed by looping within the  $V_H$  region to facilitate *VDJ* recombination (57, 61). This builds a rosette-like structure and allows equal usage between proximal, middle and distal  $V_H$  genes, essential for the generation of a diversified immunoglobulin repertoire (62). After productive *IgH* recombination and in response to pre-BCR signaling, locus contraction is reversed in small pre-B cells (60) or in large pre-B cells (57) depending on the studies. This decontracted state remains until the immature stage (57). Discrepancies between Rother's (60) and Roldan's (57) reports could be due to different probes used in 3D-FISH. Rother used three different probes encompassing distal  $V_H$ , proximal  $V_H$  and  $C_H$  genes whereas Roldan employed only two probes including one spanning *C $\gamma$ 1* and the other  $V_H558$  (distal). Ultimately, *IgH* looping plays a critical role in *VDJ* recombination by forming a "recombination center" containing *pDQ52*, all four  $J_H$  segments and *cE $\mu$*  (17). These loops permit RAG scanning between two recognition signal sequences (*RSS*) and therefore focal RAG targeting activity during  $D_H$  to  $J_H$  joining. This loop extrusion, is thought to ensure better sequence recognition by the RAG complex, avoiding off-targeting (63).

While *IgH* locus contraction and looping are important to permit equal usage of all  $V_H$  genes, nuclear positioning of *IgH* loci seems more determinant than locus contraction in the orchestration of ordered *VDJ* rearrangements (60).

## **Ig $\kappa$ Nuclear Positioning and Chromosome Loops**

In contrast to the *IgH* locus, the exact kinetics of *Ig $\kappa$*  loci contraction and relocation within the nucleus remain unclear but it is admitted that dynamic changes also occur at *Ig $\kappa$*  loci. In CLP and pre-pro B cells, both *Ig $\kappa$*  alleles are located in PCH. At the pro-B cell stage, *Ig $\kappa$*  loci relocate to a more central and active area within the nucleus (54, 64, 65). From large pre-B cells and until the immature stage, one *Ig $\kappa$*  allele becomes more closely associated with active chromatin and the other *Ig $\kappa$*  allele stays in the PCH (64) (**Figure 3A**). This chronology is still debated. Rother et al. described relocation of both *Ig $\kappa$*  loci only at pre-B cell stages (60). In small pre-B cells, contraction occurs at the *Ig $\kappa$*  locus, in preparation for *V $\kappa$ -J $\kappa$*  recombination and this contracted status persists until the immature B cell stage (39, 57) (**Figure 3B**). While some controversy still remains concerning contraction of *Ig $\kappa$*  locus in all early developmental stages, it is clear that redistribution of intra-loci interactions occurs at the small pre-B cell stage (60, 66). This redistribution, mediated by pre-BCR signaling, results in *Ig $\kappa$*  looping through ordered coordination between *MiE $\kappa$* , *3'E $\kappa$*  and *Sis* regulatory elements spread throughout the locus (67).

## **Role of Ig Enhancers in B Cell Nuclear Organization**

### **IgH V- $D_H$ Intergenic Region and IGCR1**

Beyond the observation that *IGCR1* interacts with the  $3'CBEs$  insulator region in pro-B cells (68, 69); recent studies suggest an insulation function for *IGCR1* itself. In an *IGCR1* deficient mouse model, interaction between the whole  $V-D_H$  intergenic region and  $3'CBEs$  was abolished while  $E\mu/3'RR$  interaction was not (69). In *IGCR1* deficient pro-B cells,  $E\mu/IGCR1$  interaction was decreased as expected but  $E\mu/V_H81$  and  $E\mu/3'RR$  interactions still occurred (70). Moreover, the recombination center, in which *DFL1.6/DQ52/E $\mu$ /hs4* interact in a *wt* context, was disturbed in *IGCR1* deficient pro-B cells. The absence of *IGCR1* allows an alternative bigger loop to occur between  $V_H81X/DQ52/E\mu/hs4$ . By including proximal  $V_H$  segments in the former loop, this new chromosome conformation alters *V(D)J* repertoire by decreasing distal  $V_H$  usage (70). According to the actual model, *IGCR1* seems to work as an insulator to delimit  $E\mu$  action/function. These findings are contradicted by Busslinger's group who carried out 4C-Seq experiments in *IGCR1* deficient pro-B cells where interaction patterns remain unchanged compared to *Rag $^{-/-}$*  pro-B cells (71). This discrepancy between studies could be explained by the restriction enzymes used in 4C experiments. When Busslinger's group used sequentially *HindIII* (6pb cutter) and *Sau3AI* enzymes, Sen's group used two 4pb cutter enzymes (*MseI* and *NlaIII*) that generate smaller fragments and probably offer higher resolution (**Table 1**).

### **IgH DQ52 Enhancer and E $\mu$ Region**

Many studies have contributed to the identification and unraveling of the role of *Ig* gene enhancers on 3D loci conformation during early B cell development proving that such regulatory elements are clearly involved in subnuclear positioning and chromosome looping (intra-TAD modification). To decipher the role of *pDQ52* and the  $E\mu$  enhancer, Guo and colleagues generated two mouse models (17):  $P^E^-$  (both *pDQ52/DQ52* and  $E\mu$  enhancers are deleted) and  $P^E^+$  (only promotor *pDQ52* and *DQ52* segments are deleted). Using 3D-FISH, authors demonstrated that the  $E\mu$  enhancer is required for *IgH* locus contraction in pro-B cells. Indeed, in  $P^E^-$ , but not in the  $P^E^+$  model, the large loops between  $V-D_H$  intergenic regions (*DFL16.1*, *IGCR1*),  $E\mu$  and the  $3'RR$  are drastically reduced underlining the important role of  $E\mu$  in loop formation. Moreover, ChIP-Seq experiments coupled with 3C experiments demonstrated that such large domain interactions are shaped by CTCF and cohesin proteins (68). In this way, CTCF-mediated *IgH* looping facilitates the generation of a diversified repertoire by juxtaposing distal  $V_H$  to  $D_H$  regions (71). It has also been shown that, in pro-B cells, *cE $\mu$*  deletion leads to *IgH* relocation in close proximity to the nuclear periphery (17). Taken together, these studies show that the  $E\mu$  region has a pivotal role during early *VDJ* recombination by regulating both long range interactions and *IgH* sub-nuclear positioning. While  $E\mu$  seems to facilitate *IgH* loops in developing B cells, the presence of this enhancer is not strictly necessary since another study showed, by 4C-Seq, that several long-range interactions remain detectable in the absence of  $E\mu$  (71) (**Table 1**).

**TABLE 1** | Contribution of Immunoglobulin enhancers to B Cell Nuclear Organization.

	Enhancers	Early development	Late development
IgH locus	<b>5'hs123ab</b>		
	<b>IGCR1</b>	*VDH-3'CBEs interactions (17, 69). *Maintain recombination center in pro B cells (DFL1.6; DQ52; Eμ; hs4) (17, 69). *Formation of alternative larger loop between VH81X, DQ52, Eμ and hs4 (70).	
	<b>pDQ52</b>		
	<b>Eμ-MAR</b>	*Loop formation between DFL16.1, IGCR1, Eμ and 3'RR. (17, 69). *Maintain <i>IgH</i> loci at nuclear periphery in pro B cells (17, 69).	
	<b>γ1E</b>		
	<b>3'RR</b>		*Maintain proximity of both IgH loci (Le Noir et al., 2016) (86). *hs3b and hs4 enhancers required for interaction between 3'RR and Eμ regions (74).
Igκ locus	<b>3'CBE</b>		*Total 3'CBE region involved in interactions between Eμ and targeted Switch regions (25).
	<b>κRE1</b>	*Interaction between Vκ/Jκ with MiEκ, 3'Eκ and 3'Ed (38).	
	<b>E88</b>		
	<b>hs1-2=Cer</b>	* <i>Igk</i> locus contraction (40).	
	<b>hs3-6=Sis</b>	* <i>IgH</i> and <i>Igk</i> addressing to PCH in pre-B cells (39).	
	<b>MiEκ</b>	* <i>Igk</i> locus contraction (60) Maintain <i>Igk</i> in an active compartment (72).	
	<b>3'Eκ</b>	*Maintain <i>Igk</i> in active compartment at pre B cell stage (65) * <i>IgH</i> decontraction, relocalization to PCH and <i>IgH-Igk</i> association (72).	*Maintain <i>IgH</i> , <i>Igk</i> and <i>IgJ</i> loci close to nuclear periphery in plasma cells (65).
	<b>3'Ed</b>		

In the absence of study depicting the role of Ig enhancers the boxes are empty. Bold correspond to the name of the enhancers.

## Igκ E88

The *E88* element located in the intervening *V-Jκ* region also participates in the generation of a diversified Ig repertoire (38). Endowed with enhancer activity starting at pre-B cell stages, *E88* regulates long-range *Igκ* chromatin interactions and participates in sub-TAD determination of the *Igκ* locus. *E88* deletion in pro-B cells disrupts interactions between *Vκ/Jκ* genes and three other known enhancers, especially *MiEκ*, but also to some extent *3'Eκ* and *3'Ed* (38). It is then reasonable to consider *E88* as a major hub of *Igk* locus interactions critical for regulation of *Igk* repertoire diversity.

## Igκ Cer and Sis Elements

While of critical interest to generate a diversified antibody repertoire of the *Igκ* chain (39, 40); *Cer* and *Sis* regulatory elements located within the *Igκ* locus are also involved in *Ig* loci PCH addressing. It is now established that *IgH* and *Igκ* loci are monoallelically repositioned into PCH in pre-B cells (72). In *Sis*<sup>-/-</sup> pre-B cells, both *IgH* and *Igk* alleles do not relocate to PCH. While *Igκ* locus contraction and looping still occur normally in the absence of this element, *Sis*-deficient B cells harbor a biased *Igκ* repertoire with increased proximal *Vκ* gene usage accompanied by decreased distal *Vκ* utilization. Taken

together, these results show that the *Sis* element is required for *Igκ* (*cis*-acting) and *IgH* (*trans*-acting) monoallelic positioning within the nucleus and somehow promotes a diversified antibody repertoire (39). *Cer* deletion leads to a strong increase in proximal *Vκ* usage with decreased distal *Vκ* usage. *Cer*<sup>-/-</sup> pre-B cells show normal epigenetic marks throughout the *Igκ* locus but the compaction level is clearly decreased suggesting that *Cer* regulates *Igκ* locus contraction (40). As a whole, *Sis* seems to be more involved in nuclear *Ig* loci positioning and *Cer* in *Igκ* locus conformation. The idea that nuclear positioning, rather than loci contraction, promotes *Ig* gene recombination (60) is thereby questioned by the reciprocal functions of *Sis* and *Cer* elements in the *Igκ* locus.

## Igκ MiEκ Element

*Igκ* locus contraction, mediated by *MiEκ*, seems to occur at the pro-B cell stage and remains until the pre-B cell stage. This observation indicates that locus contraction is lineage, but not stage specific (60). In pro-B cells, *MiEκ* deletion both provokes positioning of both *Igκ* alleles to PCH and reduces physical distance between *IgH* and *Igκ* loci evidence of *Ig* loci crosstalk at this stage. In pre-B cells, *MiEκ* deletion enforces both *Igk* and *IgH* allele positioning to PCH (76). The current model proposes



that *MiEκ* is more implicated in *Ig* loci sub-nuclear organization than in locus looping.

### 3' E $\kappa$ Region

Involvement of 3'E $\kappa$  in early rearrangement of the *Igκ* locus is clearly established. This enhancer region is necessary for *Igκ* germline transcription activation, which is a prerequisite for V $\kappa$ -J $\kappa$  recombination (77). In line with the previously observed transcription defect, deletion of the 3'E $\kappa$  enhancer induces premature repositioning of *Igκ* into PCH at the pre-B cell stage. Moreover, according to Park and colleagues, in pro-B cells, *Igκ* seems to be distant from the *IgH* locus although both *IgH* and *Igκ* appear to be located at the nuclear periphery (65). More strikingly, in 3'E $\kappa$  deficient mice, Skok and colleagues also documented multiple defects in *IgH* decontraction, relocalization to PCH and *IgH*-*Igκ* association (76). Altogether these results suggest that *IgH* decontraction is dependent on PCH repositioning and *Igκ*-*IgH* colocalization. This particular model highlights an unexpected *trans* – acting effect upon deletion of an *Ig* gene enhancer. This observation supposes strong crosstalk between *Ig* loci in developing B cells and supports the hypothesis that regulatory regions are involved in interactions between loci. In parallel, crosstalk between *Igκ* and *IgH* is a little more frequent in pre-B cells than in pro-B cells and becomes almost inexistent at immature stages (65).

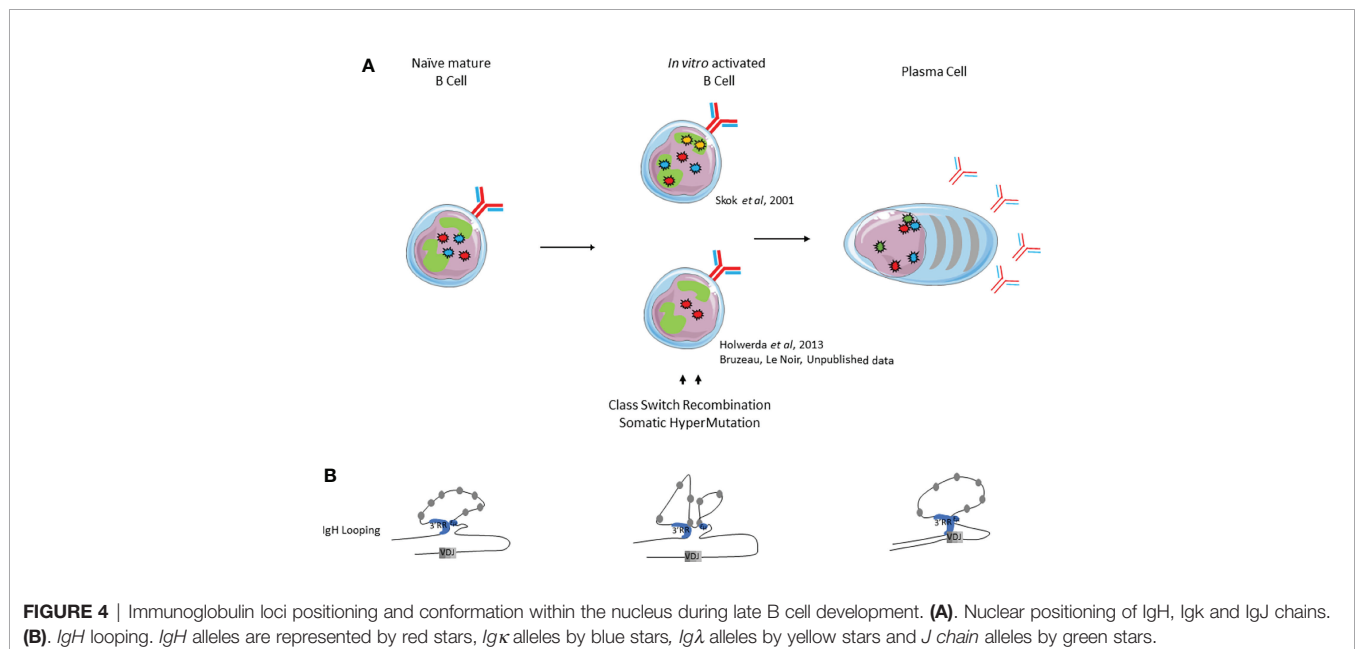
## NUCLEAR ORGANIZATION DRIVEN BY IG ENHANCERS IN LATE DEVELOPING B CELLS

In secondary lymphoid organs, mature B cells can encounter soluble or membrane antigens able to engage their BCR and

induce proper B cell activation for secondary remodeling events. Within the germinal center (GC), B cells undergo to two major genetic rearrangements initiated by AID (78): SHM and CSR at *IgH* loci whereas only SHM occurs at *IgL* loci. By inducing frequent point mutations into the variable regions of *Ig* loci, SHM is a driving force for antibody affinity maturation. CSR results in  $\mu$  heavy chain replacement by another subtype (IgG, IgA or IgE) after initiating DNA double strand breaks and recombination. In addition, locus suicide recombination (LSR) is another rearrangement which leads to deletion of all *IgH* constant genes and therefore induces B cell death by abrogating surface B cell receptor expression (1). In parallel to such secondary remodeling events, activated B cells differentiate into antibody secreting plasma cell or memory B cells (73).

### IgH Nuclear Positioning and Loops in Mature Naive and Activated B Cells

Mature B cell differentiation and its accompanying late genetic remodeling events are also characterized by changes in nuclear organization and chromatin loops. Pioneering studies used 3D-FISH to show that *IgH* alleles were not located in similar positions within nuclei of resting naive and activated B cells. At the resting mature B cell stage, both *IgH* alleles are localized in euchromatin whereas their respective nuclear position upon *in vitro* activation is still debated. According to Skok's lab, the *IgH* allele colocalized with PCH and replicated later, suggesting that this allele is the unproductive allele (72) (Figure 4). These elements imply that the non-productive *IgH* allele is “tagged and maintained as excluded” by nuclear location. Although when using the same 3D-FISH approach in stimulated cells, the De Latt group showed that both *IgH* loci are sitting in an active compartment (79) (Figure 4A). This observation is in agreement with another study in which *IgH* alleles relocalized to the nuclear



periphery in proliferating splenic B cells (80). By using mouse models carrying IgM of “a” and “b” allotypes (respectively from C57Bl6 and SV129 backgrounds), Holwerda and colleagues also showed, by 4C-Seq, that tardive replication in splenic B cells was lymphoid specific but independent of nuclear location and topology of *Ig* loci (79). The fact that allelic exclusion might not be driven by nuclear location is in agreement with previous studies showing *IgH* biallelic expression in mature B cells (58, 81). Evidence for loop formation was first provided by the Kenter group that described high frequencies of interaction between *E $\mu$*  and 3'RR in mature resting B cells. Loop conformation within *IgH* changes upon *in vitro* stimulation by lipopolysaccharide (LPS) with or without interleukin (IL-4), which respectively induce CSR mostly towards IgG3 and IgG1; by acquiring additional contact between previously interacting enhancers and acceptor switch region (*S $\gamma_3$*  or *S $\gamma_1$* ) involved in CSR (75) (Figure 4B). The advantage of this architectural scaffolding, promoting synapsis between *S* regions, has been mechanistically demonstrated to facilitate CSR by the loop extrusion mechanism (74, 82). Briefly, cohesin is loaded at the 3'RR end and the extrusion mechanism bring together 3'RR and *E $\mu$*  regions. An additional internal loop is further formed to juxtapose the transcriptionally active *S* regions. This particular conformation first allows AID recruitment at both *S $\mu$*  and *S* acceptor regions to induce DSB and second maintains *S* regions together for ligation by the NHEJ pathway [for review (83)].

### Ig $\kappa$ Nuclear Positioning in Mature Naive and Activated B Cells

Similar observations can be made for *IgL* loci (Figure 4). *Ig $\kappa$*  monoallelic positioning in PCH is often observed in activated splenic B cells and this PCH-localized allele replicates later compared to the other one (64, 84). The current model proposes that this monoallelic positioning drives both light chain allelic and isotypic exclusion. Effectively, both *Ig $\lambda$*  alleles are located in repressive areas in the nuclear periphery of *in vitro* stimulated murine B cells, which mostly express *Ig $\kappa$* . When *Ig $\lambda$*  expression is forced, one *Ig $\lambda$*  allele is recruited to a central permissive compartment within the nucleus and both *Ig $\kappa$*  alleles are repositioned to the nuclear periphery, thus inhibiting *Ig $\kappa$*  expression (72). Moreover, colocalization of *Ig $\kappa$*  with *IgH* loci is observed more frequently upon LPS stimulation of B cells (65).

### Ig Nuclear Positioning and Chromosome Looping in Plasma Cells

The intense antibody secreting function of plasma cells requires high levels of immunoglobulin gene transcription. For this purpose, nuclear organization can now be considered as one important level of regulation as described by Garrard and colleagues (65). Although located on three different chromosomes, *Ig* genes (*IgH*, *Ig $\kappa$*  and *Ig $\lambda$* ) in plasma cells often undergo physical clustering by forming pairs or triplets. Such clusters, often composed of functional alleles, preferentially localize within the same transcription factory near the nuclear periphery (Figure 4A). In this study, ChIP-3C-seq experiments performed with anti-RNA PolII antibodies indicate that physical

interactions within transcription factories are mediated by *Ig* gene enhancers: *cE $\mu$*  and 3'RR for *IgH*, *Ei $\kappa$*  and 3'E $\kappa$  for *Ig $\kappa$*  loci. Nuclear location also facilitates transport of *Ig* transcripts from the nucleus towards the endoplasmic reticulum (65). Moreover, 3C experiments performed by Birshstein's group underscored the importance of *IgH* chromatin conformation in a plasma cell line. In this study, authors showed that physical interactions between *V $H$*  genes and the 3'RR were involved in efficient *IgH* transcription, heavy chain expression and ensuing antibody synthesis (85). All of these elements demonstrate that nuclear organization contributes favorably to massive antibody synthesis (65).

### Role of Ig Enhancers on B Cell Nuclear Organization and Loops (Table 1)

While globally less documented, the implication of *Ig* gene enhancers in B cell nuclear organization (*Ig* positioning and loop formation) has been proposed in the context of the first-described *IgH* locus loop bringing the 3'RR and *E $\mu$*  regions into close contact before CSR. Upon LPS-stimulation, this configuration acquires additional contact between the acceptor *S* region involved in CSR and previous interacting *IgH* enhancers (75). In plasma cells, *Ig* loci “coordination” seems to be mediated by the 3'E $\kappa$  regulatory element since its deletion leads to a decrease in cohabitation of all *Ig* loci with decreases in *IgH-Ig $\kappa$* , *IgH-IgJ* and *Ig $\kappa$ -IgJ* communication. Likewise, nuclear localization of these loci seems to be modified with the relocation of *IgH*, *Ig $\kappa$*  and *IgJ* alleles further from the nuclear periphery in 3'E $\kappa$ <sup>-/-</sup> plasma cells compared to *wt* plasma cells. Together, these results suggest that interactions between *Ig* genes (including interaction between *IgH* and *IgJ*) are mediated by *Ig $\kappa$* , especially by 3'E $\kappa$  enhancer. Moreover, relocation and mis-cohabitation of *Ig* loci correlate with decreased transcription of each *Ig* gene, as observed in 3'E $\kappa$ <sup>-/-</sup> plasma cells in comparison with *wt* cells (65).

*IgH* positioning within the nucleus might also be supported by the 3'RR. Complete 3'RR deletion leads to an increase in distance between both *IgH* loci in activated B cells (86). The pioneering study by Kenter's group showed that *IgH* looping requires an intact 3'RR. This statement is supported by the model devoid of its two last enhancers, *hs3b* and *hs4*, that led to a decrease in interactions between 3'RR and *E $\mu$*  regions in resting and LPS ± IL4 stimulated splenic B cells (75). In contrast, partial deletion of 3'CBEs (*hs5* to 7) does not impair loop formation at resting and activated B cell stages (26). However, Alt's group reported that total 3'CBEs deletion of led to significant decreases in interactions between *E $\mu$*  and targeted switch regions in activated B cells (25). Deletion of the *cE $\mu$*  enhancer seems to have no impact on interactions between 3'RR and *E $\mu$*  regions since the contact frequencies in *cE $\mu$*  KO B cells are comparable to *wt* B cells (75). Changes in *IgH* locus conformation mediated by interactions between promoters and enhancers, could impair CSR by limiting the activity of promoters located upstream from each constant gene, and therefore, their ability to initiate prerequisite germline transcription (87).

## DISCUSSION

Nuclear positioning and chromosome looping have clearly been shown to display a functional role in early developing B cells and in B cells in general. Beyond some discrepancies, studies overall showed that positioning and conformation of *Ig* loci play a major role in ordered H and L chains rearrangements. Furthermore, at such early stages, nuclear repositioning of *Ig* loci to PCH is moreover integral to allelic exclusion mechanism while chromosome looping optimizes antibody repertoire constitution. Regulatory regions of *Ig* gene loci largely participate in positioning and looping. At both *IgH* and *Igk*, *E $\mu$*  and 3'RR as well as *Cer* elements are essential for the proper loop formation in their respective loci. In contrast, *Igk* *Sis* element is supposed to ensure crosstalk between *IgH* and *Igk* loci since influencing *cis*- as well as *trans*-positioning. Similarly, 3'*E $\kappa$*  also promotes, at early stages, temporal association between *IgH* and *Igk*.

In contrast, in mature B cells, the role of *Ig* loci position within the nucleus remains more elusive since studies unveil various differences. Nevertheless, *IgH* chromosome looping has been distinctly shown to play a role in CSR, especially by bringing donor *S $\mu$*  and acceptor switch regions in close proximity to optimize switch recombination events. During SHM, primary transcription of the *V* exon, potentially in both sense and antisense direction, is an important prerequisite for AID-induced mutations. In KO mouse models in which the 3'RRpalindromic structure is disrupted, the observed SHM defects correlate to a decrease in *V<sub>H</sub>* primary transcription (29, 35, 36). As transcription is an essential step to initiate loop extrusion prior CSR, establishing a link between SHM and loop extrusion would be of critical interest. In addition, the function of ncRNA in chromosome topology during CSR (88) and SHM (89) has been elucidated. In plasma cells, only few studies reported that *Ig* loci are colocalized to transcription factories,

surely to improve antibody production. Similarly, the role of *Ig* loci conformation throughout terminal B cell development is also of major interest.

*IgH* locus positioning within the nucleus during CSR seems to be important for legitimate maintenance. While *IgH* nuclear location in close proximity to an oncogene is a contributing factor for translocation (90); its distal 3'RR enhancers *hs3b* and *hs4* have been shown to be only involved in oncogene activation but not in translocation (91). Given that AID activity initiates DNA modifications, it is mandatory that *IgH* rearrangements be tightly regulated to avoid any mutations or translocations within oncogenes (so called "AID off targets"). Several AID off target genes (*ie: Pax5*, *Il4ra* and *Inf8*) lie in close proximity to *IgH* during CSR and altogether are located in a chromosomal territory containing a high AID concentration (90). As a consequence, the non-random positioning and conformation of *Ig* loci could be widely considered as critical to maintain B cell genome integrity. In this way, their respective regulatory elements could play the role of B cell genome guardians thereby avoiding illegitimate events that contribute to lymphomagenesis.

## AUTHOR CONTRIBUTIONS

CB wrote the manuscript and prepared the figures. SN wrote the manuscript. EP provided critical feedback. All authors contributed to the article and approved the submitted version.

## FUNDING

CB was supported by a Ph.D. fellowship of the french Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation. ANR-21-CE15-0001-01.

## REFERENCES

- Laffleur B, Denis-Lagache N, Cook-Moreau J, Tinguely A, Delpy L, et al. AID-Driven Deletion Causes Immunoglobulin Heavy Chain Locus Suicide Recombination in B Cells. *Science* (2012) 336:931–4. doi: 10.1126/science.1218692
- Dekker J, Rippe K, Dekker M, Kleckner N. Capturing Chromosome Conformation. *Science* (2002) 295:1306–11. doi: 10.1126/science.1067799
- Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. Looping and Interaction Between Hypersensitive Sites in the Active Beta-Globin Locus. *Mol Cell* (2002) 10:1453–65. doi: 10.1016/S1097-2765(02)00781-5
- Ea V, Baudement M-O, Lesne A, Forné T. Contribution of Topological Domains and Loop Formation to 3D Chromatin Organization. *Genes* (2015) 6:734–50. doi: 10.3390/genes6030734
- Wutz G, Várnai C, Nagasaka K, Cisneros DA, Stocsits RR, Tang W, et al. Topologically Associating Domains and Chromatin Loops Depend on Cohesin and are Regulated by CTCF, WAPL, and PDS5 Proteins. *EMBO J* (2017) 36:3573–99. doi: 10.15252/embj.201798004
- Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, et al. Spatial Partitioning of the Regulatory Landscape of the X-Inactivation Center. *Nature* (2012) 485:381–5. doi: 10.1038/nature11049
- Amiad-Pavlov D, Lorber D, Bajpai G, Reuveny A, Roncato F, Alon R, et al. Live Imaging of Chromatin Distribution Reveals Novel Principles of Nuclear Architecture and Chromatin Compartmentalization. *Sci Adv* (2021) 7:eabf6251. doi: 10.1126/sciadv.abf6251
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, et al, et al. Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions. *Nature* (2012) 485:376–80. doi: 10.1038/nature11082
- Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* (2009) 326:289–93. doi: 10.1126/science.1181369
- Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, et al. Nuclear Organization of Active and Inactive Chromatin Domains Uncovered by Chromosome Conformation Capture–on–Chip (4C). *Nat Genet* (2006) 38:1348–54. doi: 10.1038/ng1896
- Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel , et al. Interchromosomal Interactions and Olfactory Receptor Choice. *Cell* (2006) 126:403–13. doi: 10.1016/j.cell.2006.06.035
- Spilianakis CG, Lalioti MD, Town T, Lee GR, Flavell RA. Interchromosomal Associations Between Alternatively Expressed Loci. *Nature* (2005) 435:637–45. doi: 10.1038/nature03574
- Arakawa H, Shimizu T, Takeda S. Re-Evaluation of the Probabilities for Productive Rearrangements on the K and X Loci. *Int Immunol* (1996) 9(8):91–9. doi: 10.1093/intimm/8.1.91
- Engel H, Rolink A, Weiss S. B Cells are Programmed to Activate B and  $\emptyset$  for Rearrangement at Consecutive Developmental Stages. *Eur J Immunol* (1999) 10(29):2167–76. doi: 10.1002/(SICI)1521-4141(199907)29:07<2167::AID-IMMU2167>3.0.CO;2-H

15. Pawlitzky I, Angeles CV, Siegel AM, Stanton ML, Riblet R, Brodeur PH. Identification of a Candidate Regulatory Element Within the 5' Flanking Region of the Mouse Igh Locus Defined by Pro-B Cell-Specific Hypersensitivity Associated With Binding of PU.1, Pax5, and E2A. *J Immunol Baltim Md 1950* (2006) 176:6839–51. doi: 10.4049/jimmunol.176.11.6839
16. Featherstone K, Wood AL, Bowen AJ, Corcoran AE. The Mouse Immunoglobulin Heavy Chain V-D Intergenic Sequence Contains Insulators That May Regulate Ordered V(D)J Recombination. *J Biol Chem* (2010) 285:9327–38. doi: 10.1074/jbc.M109.098251
17. Guo C, Gerasimova T, Hao H, Ivanova I, Chakraborty T, et al. Two Forms of Loops Generate the Chromatin Conformation of the Immunoglobulin Heavy-Chain Gene Locus. *Cell* (2011) 147:332–43. doi: 10.1016/j.cell.2011.08.049
18. Alessandrini A, Desiderio SV. Coordination of Immunoglobulin DJH Transcription and D-To-JH Rearrangement by Promoter-Enhancer Approximation. *Mol Cell Biol* (1991) 11:2096–107. doi: 10.1128/mcb.11.4.2096-2107.1991
19. Predeus AV, Gopalakrishnan S, Huang Y, Tang J, Feeny AJ, Oltz EM, et al. Targeted Chromatin Profiling Reveals Novel Enhancers in Ig H and Ig L Chain Loci. *J Immunol Baltim Md 1950* (2014) 192:1064–70. doi: 10.4049/jimmunol.1302800
20. Amoretti-Villa R, Rogier M, Robert I, Heyer V, Reina-San-Martin B. A Novel Regulatory Region Controls IgH Locus Transcription and Switch Recombination to a Subset of Isotypes. *Cell Mol Immunol* (2019) 16:887–9. doi: 10.1038/s41423-019-0267-4
21. Lieberson R, Giannini SL, Birshtein BK, Eckhardt LA. An Enhancer at the 3' End of the Mouse Immunoglobulin Heavy Chain Locus. *Nucleic Acids Res* (1991) 19:933–7. doi: 10.1093/nar/19.4.933
22. Chauveau C, Cogné M. Palindromic Structure of the IgH 3' locus Control Region. *Nat Genet* (1996) 14:15–6. doi: 10.1038/ng0996-15
23. Madisen L, Groudine M. Identification of a Locus Control Region in the Immunoglobulin Heavy-Chain Locus That Deregulates C-Myc Expression in Plasmacytoma and Burkitt's Lymphoma Cells. *Genes Dev* (1994) 8:2212–26. doi: 10.1101/gad.8.18.2212
24. Sepulveda M, Garrett F, Pricewhelan A, Birshtein B. Comparative Analysis of Human and Mouse 3' Regulatory Regions Identifies Distinctive Structural Features. *Mol Immunol* (2005) 42:605–15. doi: 10.1016/j.molimm.2004.09.006
25. Zhang X, Yoon HS, Chapdelaine-Williams AM, Kyritsis N, Alt FW. Physiological Role of the 3'IgH CBEs Super-Anchor in Antibody Class Switching. *Proc Natl Acad Sci* (2021) 118:e2024392118. doi: 10.1073/pnas.2024392118
26. Volpi SA, Verma-Gaur J, Hassan R, Ju Z, Roa S, Chatterjee S, et al. 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. *J Immunol* (2012) 188:2556–66. doi: 10.4049/jimmunol.1102763
27. Vincent-Fabert C, Fiancette R, Pinaud E, Denizot Y. Genomic Deletion of the Whole IgH 3J Regulatory Region (Hs3a, Hs1,2, Hs3b, and Hs4) Dramatically Affects Class Switch Recombination and Ig Secretion to All Isotypes. *Blood* (2010) 116:5. doi: 10.1182/blood-2010-01-264689
28. Perlot T, Alt FW, Bassing CH, Suh H, Pinaud E. Elucidation of IgH Intronic Enhancer Functions via Germ-Line Deletion. *Proc Natl Acad Sci* (2005) 102:14362–7. doi: 10.1073/pnas.0507090102
29. Garot A, Marquet M, Saintamand A, Bender S, Le Noir S, Rouaud P, et al. Sequential Activation and Distinct Functions for Distal and Proximal Modules Within the IgH 3' Regulatory Region. *Proc Natl Acad Sci* (2016) 113:1618–23. doi: 10.1073/pnas.1514090113
30. Vincent-Fabert C, Truffinet V, Fiancette R, Cogné N, Cogné M, Denizot Y, et al. Ig Synthesis and Class Switching Do Not Require the Presence of the Hs4 Enhancer in the 3' IgH Regulatory Region. *J Immunol* (2009) 182:6926–32. doi: 10.4049/jimmunol.0900214
31. Bébin A-G, Carrion C, Marquet M, Cogné N, Lecardeur S, Cogné M, et al. *In Vivo* Redundant Function of the 3' IgH Regulatory Element HS3b in the Mouse. *J Immunol* (2010) 184:3710–7. doi: 10.4049/jimmunol.0901978
32. Pinaud E, Khamlichi AA, Le Morvan C, Drouet M, Nalesso V, Le Bert M, et al. Localization of the 3' IgH Locus Elements That Effect Long-Distance Regulation of Class Switch Recombination. *Immunity* (2001) 15:187–99. doi: 10.1016/S1074-7613(01)00181-9
33. Manis JP, van derStoep N, Tian M, Ferrini R, Davidson L, Bottaro A, et al. Class Switching in B Cells Lacking 3' Immunoglobulin Heavy Chain Enhancers. *J Exp Med* (1998) 188:1421–31. doi: 10.1084/jem.188.8.1421
34. Cogné M, Lansford R, Bottaro A, Zhang J, Gorman J, Young , et al. A Class Switch Control Region at the 3' End of the Immunoglobulin Heavy Chain Locus. *Cell* (1994) 77:737–47. doi: 10.1016/0092-8674(94)90057-4
35. Le Noir S, Boyer F, Lecardeur S, Brousse M, Oruc Z, Cook-Moreau J, et al. Functional Anatomy of the Immunoglobulin Heavy Chain 3' Super-Enhancer Needs Not Only Core Enhancer Elements But Also Their Unique DNA Context. *Nucleic Acids Res* (2017) 45:5829–37. doi: 10.1093/nar/gkx203
36. Saintamand A, Vincent-Fabert C, Garot A, Rouaud P, Oruc Z, Magnone V, et al. Deciphering the Importance of the Palindromic Architecture of the Immunoglobulin Heavy-Chain 3' Regulatory Region. *Nat Commun* (2016) 7:10730. doi: 10.1038/ncomms10730
37. Bruzeau C, Moreau J, Le Noir S, Pinaud E. Panorama of Stepwise Involvement of the IgH 3' Regulatory Region in Murine B Cells. *Adv Immunol* (2021) 149:95–114. doi: 10.1016/bs.ai.2021.03.004
38. Barajas-Mora EM, Kleiman E, Xu J, Carrico NC, Lu H, Oltz EM, et al. A B-Cell-Specific Enhancer Orchestrates Nuclear Architecture to Generate a Diverse Antigen Receptor Repertoire. *Mol Cell* (2019) 73:48–60.e5. doi: 10.1016/j.molcel.2018.10.013
39. Xiang Y, Zhou X, Hewitt SL, Skok JA, Garrard WT. A Multifunctional Element in the Mouse *Igk* Locus That Specifies Repertoire and Ig Loci Subnuclear Location. *J Immunol* (2011) 186:5356–66. doi: 10.4049/jimmunol.1003794
40. Xiang Y, Park S-K, Garrard WT.  $\text{V}\kappa$  Gene Repertoire and Locus Contraction Are Specified by Critical DNase I Hypersensitive Sites Within the  $\text{V}\kappa$ - $\text{J}\kappa$  Intervening Region. *J Immunol* (2013) 190:1819–26. doi: 10.4049/jimmunol.1203127
41. Cockerill PN, Garrard WT. Chromosomal Loop Anchorage of the Kappa Immunoglobulin Gene Occurs Next to the Enhancer in a Region Containing Topoisomerase II Sites. *Cell* (1986) 44:273–82. doi: 10.1016/0092-8674(86)90761-0
42. Emorine L, Kuehl M, Weir L, Leder P, Max EE. A Conserved Sequence in the Immunoglobulin  $\text{J}\kappa$ - $\text{C}\kappa$  Intron: Possible Enhancer Element. *Nature* (1983) 304:447–9. doi: 10.1038/304447a0
43. Bergman Y, Rice D, Grosschedl R, Baltimore D. Two Regulatory Elements for Immunoglobulin Kappa Light Chain Gene Expression. *Proc Natl Acad Sci U S A* (1984) 81:7041–5. doi: 10.1073/pnas.81.22.7041
44. Meyer KB, Neuberger MS. The Immunoglobulin Kappa Locus Contains a Second, Stronger B-Cell-Specific Enhancer Which is Located Downstream of the Constant Region. *EMBO J* (1989) 8:1959–64. doi: 10.1002/j.1460-2075.1989.tb03601.x
45. Liu Z-M, George-Raizen JB, Li S, Meyers KC, Chang MY, Garrard WT, et al. Chromatin Structural Analyses of the Mouse *Igk* Gene Locus Reveal New Hypersensitive Sites Specifying a Transcriptional Silencer and Enhancer. *J Biol Chem* (2002) 277:32640–9. doi: 10.1074/jbc.M204065200
46. Carson S, Wu GE. A Linkage Map of the Mouse Immunoglobulin Lambda Light Chain Locus. *Immunogenetics* (1989) 29:173–9. doi: 10.1007/BF00373642
47. Hagman J, Rudin CM, Haasch D, Chaplin D, Storb U. A Novel Enhancer in the Immunoglobulin Lambda Locus is Duplicated and Functionally Independent of NF Kappa B. *Genes Dev* (1990) 4:978–92. doi: 10.1101/gad.4.6.978
48. Hendrickson BA, Conner DA, Ladd DJ, Kendall D, Casanova JE, Corthesy B, et al. Altered Hepatic Transport of Immunoglobulin A in Mice Lacking the J Chain. *J Exp Med* (1995) 182:1905–11. doi: 10.1084/jem.182.6.1905
49. Erlandsson L, Andersson K, Sigvardsson M, Lycke N, Leanderson T. Mice With an Inactivated Joining Chain Locus Have Perturbed IgM Secretion. *Eur J Immunol* (1998) 28:2355–65. doi: 10.1002/(SICI)1521-4141(199808)28:08<2355::AID-IMMU2355>3.0.CO;2-L
50. Hardy RR, Kincaid PW, Dorshkind K. The Protean Nature of Cells in the B Lymphocyte Lineage. *Immunity* (2007) 26:703–14. doi: 10.1016/j.immuni.2007.05.013
51. Nutt SL, Heavey B, Rolink AG, Busslinger M. Commitment to the B-Lymphoid Lineage Depends on the Transcription Factor Pax5. *Nature* (1999) 401:7. doi: 10.1038/44076
52. Chakraborty T, Chowdhury D, Keyes A, Jani A, Subrahmanyam R, Ivanova , et al. Repeat Organization and Epigenetic Regulation of the DH-C $\mu$  Domain



- of the Immunoglobulin Heavy-Chain Gene Locus. *Mol Cell* (2007) 27:842–50. doi: 10.1016/j.molcel.2007.07.010
53. Mårtensson I-L, Rolink A, Melchers F, Mundt C, Licence S, Shimizu T, et al. The Pre-B Cell Receptor and its Role in Proliferation and Ig Heavy Chain Allelic Exclusion. *Semin Immunol* (2002) 14:335–42. doi: 10.1016/S1044-5323(02)00666-0
  54. Kosak ST. Subnuclear Compartmentalization of Immunoglobulin Loci During Lymphocyte Development. *Science* (2002) 296:158–62. doi: 10.1126/science.1068768
  55. Johanson TM, Lun ATL, Coughlan HD, Tan T, Smyth GK, Nutt SL, et al. Transcription-Factor-Mediated Supervision of Global Genome Architecture Maintains B Cell Identity. *Nat Immunol* (2018) 19:1257–64. doi: 10.1038/s41590-018-0234-8
  56. Lin YC, Benner C, Mansson R, Heinz S, Miyazaki K, Miyazaki M, et al. Global Changes in the Nuclear Positioning of Genes and Intra- and Interdomain Genomic Interactions That Orchestrate B Cell Fate. *Nat Immunol* (2012) 13:1196–204. doi: 10.1038/ni.2432
  57. Roldán E, Fuxa M, Chong W, Martinez D, Novatchkova M, Busslinger M, et al. Locus ‘Decontraction’ and Centromeric Recruitment Contribute to Allelic Exclusion of the Immunoglobulin Heavy-Chain Gene. *Nat Immunol* (2005) 6:31–41. doi: 10.1038/ni1150
  58. Daly J, Licence S, Nanou A, Morgan G, Mårtensson I-L. Transcription of Productive and Nonproductive VDJ-Recombined Alleles After IgH Allelic Exclusion. *EMBO J* (2007) 26:4273–82. doi: 10.1038/sj.emboj.7601846
  59. Kumari G, Sen R. Chromatin Interactions in the Control of Immunoglobulin Heavy Chain Gene Assembly. *Adv Immunol* (2015) 128:41–92. doi: 10.1016/bs.ai.2015.08.001
  60. Rother MB, Palstra R-J, Jhunjhunwala S, van Kester KAM, van Ijcken WFJ, Hendriks RW, et al. Nuclear Positioning Rather Than Contraction Controls Ordered Rearrangements of Immunoglobulin Loci. *Nucleic Acids Res* (2016) 44:175–86. doi: 10.1093/nar/gkv928
  61. Sayegh C. Visualization of Looping Involving the Immunoglobulin Heavy-Chain Locus in Developing B Cells. *Genes Dev* (2005) 19:322–7. doi: 10.1101/gad.1254305
  62. Jhunjhunwala S, van Zelm MC, Peak MM, Cutchin S, Riblet R, van Dongen JJM, et al. The 3d Structure of the Immunoglobulin Heavy-Chain Locus: Implications for Long-Range Genomic Interactions. *Cell* (2008) 133:265–79. doi: 10.1016/j.cell.2008.03.024
  63. Zhang Y, Zhang X, Ba Z, Liang Z, Dring EW, Hu H, et al. The Fundamental Role of Chromatin Loop Extrusion in Physiological V(D)J Recombination. *Nature* (2019) 573:600–4. doi: 10.1038/s41586-019-1547-y
  64. Goldmit M, Ji Y, Skok J, Roldán E, Jung S, Cedar H, et al. Epigenetic Ontogeny of the Igk Locus During B Cell Development. *Nat Immunol* (2005) 6:198–203. doi: 10.1038/ni1154
  65. Park S-K, Xiang Y, Feng X, Garrard WT. Pronounced Cohabitation of Active Immunoglobulin Genes From Three Different Chromosomes in Transcription Factories During Maximal Antibody Synthesis. *Genes Dev* (2014) 28:1159–64. doi: 10.1101/gad.237479.114
  66. Fitzsimmons SP, Bernstein RM, Max EE, Skok JA, Shapiro MA. Dynamic Changes in Accessibility, Nuclear Positioning, Recombination, and Transcription at the Ig Kappa Locus. *J Immunol Baltim Md* 1950 (2007) 179:5264–73. doi: 10.4049/jimmunol.179.8.5264
  67. Stadhouders R, de Bruijn MJW, Rother MB, Yuvaraj S, Ribeiro de Almeida C, Kolovos P, et al. Pre-B Cell Receptor Signaling Induces Immunoglobulin  $\kappa$  Locus Accessibility by Functional Redistribution of Enhancer-Mediated Chromatin Interactions. *PLoS Biol* (2014) 12:e1001791.
  68. Degner SC, Verma-Gaur J, Wong TP, Bossen C, Iverson GM, Torkamani A, et al. CCCTC-Binding Factor (CTCF) and Cohesin Influence the Genomic Architecture of the Igh Locus and Antisense Transcription in Pro-B Cells. *Proc Natl Acad Sci* (2011) 108:9566–71. doi: 10.1073/pnas.1019391108
  69. Guo C, Chunguang Yoon HS, Franklin A, Jain S, Ebert A, Cheng H-L, et al. CTCF-Binding Elements Mediate Control of V(D)J Recombination. *Nature* (2011) 477:424–30. doi: 10.1038/nature10495
  70. Qiu X, Kumari G, Gerasimova T, Du H, Labaran L, Singh A, et al. Sequential Enhancer Sequestration Dysregulates Recombination Center Formation at the IgH Locus. *Mol Cell* (2018) 70:21–33.e6. doi: 10.1016/j.molcel.2018.02.020
  71. Medvedovic J, Ebert A, Tagoh H, Tamir IM, Schwickert TA, Novatchkova M, et al. Flexible Long-Range Loops in the VH Gene Region of the Igh Locus Facilitate the Generation of a Diverse Antibody Repertoire. *Immunity* (2013) 39:229–44. doi: 10.1016/j.immuni.2013.08.011
  72. Skok JA, Brown KE, Azuara V, Caparros M-L, Baxter J, Takacs K, et al. Nonequivalent Nuclear Location of Immunoglobulin Alleles in B Lymphocytes. *Nat Immunol* (2001) 2:848–54. doi: 10.1038/ni0901-848
  73. Shaffer AL, Lin KI, Kuo TC, Yu X, Hurt EM, Rosenwald A, et al. Blimp-1 Orchestrates Plasma Cell Differentiation by Extinguishing the Mature B Cell Gene Expression Program. *Immunity* (2002) 17:51–62. doi: 10.1016/S1074-7613(02)00335-7
  74. Zhang X, Zhang Y, Ba Z, Kyritsis N, Casellas R, Alt FW, et al. Fundamental Roles of Chromatin Loop Extrusion in Antibody Class Switching. *Nature* (2019) 575:385–9. doi: 10.1038/s41586-019-1723-0
  75. Wuerffel R, Wang L, Grigera F, Manis J, Selsing E, Perlot T, et al. S-S Synapsis During Class Switch Recombination Is Promoted by Distantly Located Transcriptional Elements and Activation-Induced Deaminase. *Immunity* (2007) 27:711–22. doi: 10.1016/j.immuni.2007.09.007
  76. Hewitt SL, Farmer D, Marszalek K, Cadera E, Liang H-E, Xu Y, et al. Association Between the Igk and Igh Immunoglobulin Loci Mediated by the 3’ Igk Enhancer Induces ‘Decontraction’ of the Igh Locus in Pre-B Cells. *Nat Immunol* (2008) 9:396–404. doi: 10.1038/ni1567
  77. Inlay M, Alt FW, Baltimore D, Xu Y. Essential Roles of the  $\kappa$  Light Chain Intronic Enhancer and 3’ Enhancer in  $\kappa$  Rearrangement and Demethylation. *Nat Immunol* (2002) 3:463–8. doi: 10.1038/ni790
  78. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T, et al. Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme. *Cell* (2000) 102:553–63. doi: 10.1016/S0092-8674(00)00078-7
  79. Holwerda SJB, van deWerken HJG, Ribeiro deAlmeida C, Bergen IM, de Bruijn MJW, Verstegen MJAM, et al. Allelic Exclusion of the Immunoglobulin Heavy Chain Locus is Independent of its Nuclear Localization in Mature B Cells. *Nucleic Acids Res* (2013) 41:6905–16. doi: 10.1093/nar/gkt491
  80. Yang Q, Riblet R, Schildkraut CL. Sites That Direct Nuclear Compartmentalization Are Near the 5’ End of the Mouse Immunoglobulin Heavy-Chain Locus. *Mol Cell Biol* (2005) 25:10. doi: 10.1128/MCB.25.14.6021-6030.2005
  81. Tinguely A, Chemin G, Péron S, Sirac C, Reynaud S, Cogné M, et al. Cross Talk Between Immunoglobulin Heavy-Chain Transcription and RNA Surveillance During B Cell Development. *Mol Cell Biol* (2012) 32:107–17. doi: 10.1128/MCB.06138-11
  82. Shen HM, Wuerffel R, Cantillo JF, Priyadarshi S, Lei X, Liang J, et al. Loop Extrusion Promotes an Alternate Pathway for Isotype Switching. *Cell Rep* (2021) 37:110059. doi: 10.1016/j.celrep.2021.110059
  83. Zhang Y, Zhang X, Dai H-Q, Hu H, Alt FW. The Role of Chromatin Loop Extrusion in Antibody Diversification. *Nat Rev Immunol* (2022). doi: 10.1038/s41577-022-00679-3
  84. Mostoslavsky R, Singh N, Tenzen T, Goldmit M, Gabay C, Elizur S, et al. Asynchronous Replication and Allelic Exclusion in the Immune System. *Nature* (2001) 414:5. doi: 10.1038/35102606
  85. Ju Z, Volpi SA, Hassan R, Martinez N, Giannini SL, Gold T, et al. Evidence for Physical Interaction Between the Immunoglobulin Heavy Chain Variable Region and the 3’ Regulatory Region. *J Biol Chem* (2007) 282:35169–78. doi: 10.1074/jbc.M705719200
  86. Le Noir S, Laffleur B, Carrion C, Garot A, Lecardeur S, Pinaud E, et al. The IgH Locus 3’ Cis-Regulatory Super-Enhancer Co-opts AID for Allelic Transvection. *Oncotarget* (2017) 8:12929–40. doi: 10.18632/oncotarget.14585
  87. Feldman S, Achour I, Wuerffel R, Kumar S, Gerasimova T, Sen R, et al. Constraints Contributed by Chromatin Looping Limit Recombination Targeting During Ig Class Switch Recombination. *J Immunol Baltim Md* 1950 (2015) 194:2380–9. doi: 10.4049/jimmunol.1401170
  88. Rothschild G, Zhang W, Lim J, Giri PK, Laffleur B, Chen Y, et al. Noncoding RNA Transcription Alters Chromosomal Topology to Promote Isotype-Specific Class Switch Recombination. *Sci Immunol* (2020) 5:eay5864. doi: 10.1126/sciimmunol.aay5864
  89. Laffleur B, Lim J, Zhang W, Chen Y, Pefanis E, Bizarro J, et al. Noncoding RNA Processing by DIS3 Regulates Chromosomal Architecture and Somatic

- Hypermethylation in B Cells. *Nat Genet* (2021) 53:230–42. doi: 10.1038/s41588-020-00772-0
90. Rocha PP, Micsinai M, Kim JR, Hewitt SL, Souza PP, Trimarchi T, et al. Close Proximity to Igh Is a Contributing Factor to AID-Mediated Translocations. *Mol Cell* (2012) 47:873–85. doi: 10.1016/j.molcel.2012.06.036
91. Gostissa M, Yan CT, Bianco JM, Cogné M, Pinaud E, Alt FW, et al. Long-Range Oncogenic Activation of IgH/c-Myc Translocations by the IgH 3' Regulatory Region. *Nature* (2009) 462:803–7. doi: 10.1038/nature08633

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Bruzeau, Cook-Moreau, Pinaud and Le Noir. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

## *Objectifs*

---

## Objectifs

---

Au cours de son développement, la cellule B subit trois grands remaniements géniques majeurs, la recombinaison V(D)J, la CSR et la SHM, qui ensemble, constituent une caractéristique unique de cette cellule. La régulation précise de ces évènements est primordiale afin d'éviter la survenue de tout évènement illégitime qui pourrait initier la lymphomagenèse. Cette régulation s'opère notamment *via* l'initiation de la transcription de façon ordonnée au niveau des régions cibles de ces mécanismes. Ainsi, ce sont les segments V, D et J qui sont transcrits lors de la recombinaison V(D)J ainsi que lors de la SHM alors que les régions switch se verront transcrites pour permettre la CSR. Cette régulation transcriptionnelle est en partie permise au locus *IgH* par la présence de deux régions *cis*-régulatrices : la région *E $\mu$ MARs* et la *3'RR*.

Le laboratoire s'est particulièrement illustré dans l'étude du rôle de ces deux régions à l'aide de différents modèles murins. Ainsi, la région *E $\mu$ MARs* est particulièrement impliquée dans le contrôle de la recombinaison VDJ par le biais de l'enhancer *cE $\mu$*  alors que les régions *MARs* sont, quant à elles, impliquées dans le processus de SHM (Marquet et al., 2014; Martin et al., 2023). A l'image de la région *E $\mu$ MARs*, l'étude de différents modèles mutants de la *3'RR* ont permis de mettre en avant la fonction de deux modules distincts : les enhanceurs seuls sont suffisants pour la CSR alors que la structure quasi-palindromique caractéristique de cette région est nécessaire pour la SHM (Garot et al., 2016; Le Noir et al., 2017; Rouaud et al., 2013; Saintamand et al., 2016).

Au-delà du contrôle transcriptionnel de la SHM et de la CSR, d'autres niveaux de régulation, tels que le niveau supranucléosomal (boucles d'ADN et compartiments A/B) et le niveau nucléaire (territoires chromosomiques), interviennent au cours de la maturation des cellules B. En effet, les loci *IgH* ainsi que les loci *IgL* se repositionnent de façon dynamique au sein du noyau en fonction des stades de développement de la cellule B (Holwerda et al., 2013; Kosak et al., 2002; Park et al., 2014; Skok et al., 2001). La recombinaison VDJ et la CSR sont aussi régulés au niveau supranucléosomal avec la formation de boucles de chromatine. Pendant la recombinaison VDJ, ces boucles assurent un répertoire d'Ig varié avec l'utilisation de l'ensemble des segments V, qu'ils soient distaux, intermédiaires ou proximaux (Guo et al., 2011; Medvedovic et al., 2013; Y. Zhang et al., 2019). D'autre part, ces boucles de chromatine sont nécessaires à la CSR en permettant le rapprochement de la région switch donneuse *S $\mu$*  avec la région switch acceptrice *S $\alpha$* , formant ainsi une synapse où les acteurs des voies de réparation des CDB sont efficacement recrutés (Wuerffel et al., 2007; X. Zhang et al., 2019; Zhang et al., 2021).

Des données du laboratoire suggèrent que la *3'RR* serait également impliquée dans l'adressage du locus *IgH* au sein du noyau. Un premier axe de mon travail de thèse consistait donc à étudier l'implication des deux régions *cis*-régulatrices majeures (*E $\mu$ MARs* et *3'RR*) du locus *IgH* dans l'organisation du noyau du lymphocyte B.

A l'aide de modèles murins portant la délétion complète de la 3'RR ou seulement de sa structure palindromique, ainsi qu'à l'aide d'un large éventail de techniques ; permettant d'étudier l'adressage du locus *IgH* dans le noyau, son accessibilité ainsi que la formation de boucles, nous avons pu mettre en évidence que les core enhancers de la 3'RR sont nécessaires et suffisants à pré-organiser le noyau afin de permettre une recombinaison de CSR efficace. Ce travail, que je signe en première auteure a été publié dans BioRxiv (Bruzeau et al., 2023) et sera soumis prochainement dans *Nucleic Acid Research* (**Résultats, Partie I.1**).

L'étude de la formation de boucle dans les modèles mutants pour la région *E $\mu$ MARs* a mis en évidence une augmentation des fréquences de contact au sein du locus *IgH*, traduisant un locus plus bouclé et suggérant un temps de résidence de la cohésine augmentée. Nous suspectons alors que la 3'RR constitue le point d'entrée majeur de la cohésine alors que la région *E $\mu$ MARs* serait un site de recrutement de WAPL, dont la fonction consiste au déchargement de la cohésine. Ce travail nécessite d'être approfondi et sera alors présenté sous forme de données préliminaires (**Résultats, Partie I.2**).

Enfin, le deuxième axe de mon travail de recherche s'est articulé autour des conséquences de la délétion des régions *cis*-régulatrices du locus *IgH* sur l'intégrité génomique de la cellule B. Nous avons donc étudié le phénomène d'hypermutation somatique pouvant intervenir au niveau de cibles dites « off-target » à partir de cellules B du GC triées de nos différents modèles murins. De façon surprenante, l'ensemble des gènes off-target ne se comporte pas de la même manière selon les mutants de la 3'RR (délétion partielle ou totale). Nous avons montré que l'oncogène *Bcl6*, cible illégitime connue d'AID, est systématiquement plus muté dès qu'une portion de la 3'RR est enlevée. Cette augmentation ne semble pas corrélée avec une quelconque augmentation de sa transcription. En revanche, elle semble liée au positionnement nucléaire de *Bcl6* puisque nous observons un rapprochement de *Bcl6* et d'*IgH* par FISH-3D (**Résultats, Partie II.1**). Nous avons réalisé la même étude pour l'oncogène *cMyc* et nos résultats révèlent une régulation toute autre, mettant en avant toute la complexité du maintien de l'intégrité du génome de la cellule B. (**Résultats, Partie II.2**). Enfin, nous avons étendu cette étude à deux autres gènes AID « off-target » : *Pim1* et *Cd83*. Les résultats de cette étude sont présentés sous forme de données préliminaires (**Résultats, Partie II.3**).

## *Résultats*

---

## Résultats

---

### Partie I. Etude de l'implication des éléments *cis*-régulateurs (3'RR et EμMARs) du locus des chaînes lourdes d'immunoglobuline dans l'organisation du noyau du lymphocyte B.

#### I.I. Article n°1 : Core enhancers of the 3'RR optimize *IgH* nuclear position and loop conformation for oriented CSR.

Charlotte Bruzeau<sup>1</sup>, Justine Pollet<sup>1</sup>, Morgane Thomas<sup>1,2</sup>, Zhaoqing Ba<sup>3</sup>, David Roulois<sup>4</sup>, Eric Pinaud<sup>1,\*</sup>, Sandrine Le Noir<sup>1,\*</sup>

<sup>1</sup>UMR CNRS7276, Inserm1262, Université de Limoges: Contrôle de la Réponse Immune B et des Lymphoproliférations, 2, rue du Dr. Marcland, 87025 Limoges, France, Team 2, B-NATION: B cell Nuclear Architecture, Immunoglobulin genes and Oncogenes.

<sup>2</sup>Current Address : Laboratoire Suivi des Thérapies Innovantes, Institut de Génétique Humaine, UMR 9002 CNRS-UM, Montpellier, France.

<sup>3</sup>National Institute of Biological Science, Beijing, 102206, China.

<sup>4</sup>UMR 1236, Univ Rennes, INSERM, Etablissement Français du Sang Bretagne, Rennes, France.

\* To whom correspondence should be addressed sandrine.le-noir@unilim.fr and eric.pinaud@unilim.fr

Il est admis que l'organisation nucléaire de la cellule B (position des gènes dans le noyau) et la configuration tridimensionnelle du locus *IgH* (boucles de chromatine) participent au bon déroulement des remaniements géniques survenant au cours du développement lymphocytaire B. Néanmoins, la régulation de ces deux niveaux d'organisation du génome n'est pas totalement décrite dans le contexte de la maturation du lymphocyte B. Plusieurs éléments pointent le rôle des éléments *cis*-régulateurs du locus *IgH* (EμMAR et 3'RR) dans la régulation de l'organisation du noyau et la formation de boucle de chromatines (Wuerffel et al., 2007; X. Zhang et al., 2019). Deux modèles murins, l'un portant la délétion totale de la région régulatrice en 3' (Δ3'RR, (Vincent-Fabert et al., 2010b)) et l'autre arborant uniquement les quatre core enhancers sans la structure quasi-palindromique de la 3'RR (c3'RR, (Le Noir et al., 2017)), en comparaison à des souris sauvages (*wt*), ont permis de décortiquer le rôle de la 3'RR et de sa structure particulière dans ces deux niveaux de régulation chez le lymphocyte B mature.

L'étude de l'accessibilité globale du génome par ATAC-Seq (Assay for Transposase Accessible Chromatin) n'a pas révélé de changement majeur de compartimentation A/B entre les différents modèles murins testés. En revanche, alors que la suppression de la 3'RR diminue globalement l'accessibilité du locus *IgH*, les core enhancers seuls (modèle c3'RR) suffisent à le maintenir ouvert. Ce constat est corrélé à un adressage plus important d'*IgH* à l'hétérochromatine en contexte Δ3'RR, exploré par hybridation fluorescente *in situ* en trois dimensions (DNA FISH-3D). Ces résultats mettent en évidence que les core enhancers de la 3'RR sont suffisants pour maintenir les loci *IgH* dans une configuration accessible.

L'exploration de la position relative des loci *IgH* (distance entre loci et distance entre loci et centre du noyau) depuis le stade B transitionnel jusqu'au stade B activé a révélé une dynamique de position des

loci. En contexte *wt*, lors de la transition du stade B transitionnel vers le stade B mature au repos, les allèles *IgH* se rapprochent (à la fois l'un de l'autre et du centre du noyau) puis ils procèdent à des mouvements inverses (en s'éloignant l'un de l'autre et du centre du noyau) lors de l'activation. En revanche, en contexte  $\Delta 3'RR$ , une perte totale du mouvement des loci *IgH* (à la fois l'un vis-à-vis de l'autre et par rapport au centre du noyau) est observée. La présence des seuls core enhancers permet de restaurer une dynamique normale des loci.

L'étude de la configuration tridimensionnelle du locus *IgH* que nous avons réalisée par 3C-HTGTS (Chromosome Conformation Capture High Throughput Genome-wide Translocation Sequencing) révèle le rôle majeur des core enhancers dans la formation de la « grande boucle » entre la région *E $\mu$ MARs* et la *3'RR* au stade B au repos. En effet, cette grande boucle se forme dans les cellules B du modèle *c3'RR* alors qu'elle est complètement absente chez le mutant  $\Delta 3'RR$ . De même, les boucles internes, nécessaires à la CSR, qui impliquent les régions *S $\gamma$ 2b* et *S $\gamma$ 3* (ciblées après activation par le LPS), sont également dépendantes des core enhancers. En effet, les cellules B *c3'RR* arborent les mêmes fréquences de contact entre chacune des régions précédemment nommées que les cellules B *wt*. Au contraire, une absence totale de boucle est observée dans le modèle  $\Delta 3'RR$ .

Enfin, nous nous sommes intéressés à la structure des jonctions de CSR par la technique de LAM-HTGTS (Linear Amplification Mediated High Throughput Genome-wide Translocation Sequencing). En contexte sauvage et dans le modèle *c3'RR*, l'étude des jonctions de CSR indique que la recombinaison s'effectue par un mécanisme délétionnel, ce qui est classiquement observé. De façon attendue, très peu de jonctions de CSR sont détectées dans le modèle  $\Delta 3'RR$  ; toutefois la totalité de ces jonctions présente une orientation inverse, synonyme d'un mécanisme de CSR inversionnelle. L'ensemble de ces données permettent de faire un lien direct entre boucles de chromatine du locus *IgH* et recombinaison de CSR orientée et productive.

Nous proposons alors que les core enhancers de la *3'RR* permettent à eux seuls (sans intervention de la structure palindromique) d'optimiser la position du locus *IgH* dans le noyau ainsi que la formation des boucles de chromatine au sein du locus *IgH*, et ce afin de permettre une recombinaison de CSR optimale (car orientée) lors de l'activation antigénique.



# Core enhancers of the 3'RR optimize *IgH* nuclear position and loop conformation for oriented class switch recombination

Charlotte Bruzeau<sup>1</sup>, Justine Pollet<sup>1</sup>, Morgane Thomas<sup>1,2</sup>, Zhaoqing Ba<sup>3</sup>, David Roulois<sup>4</sup>, Eric Pinaud<sup>1,\*</sup>,  
Sandrine Le Noir<sup>1,\*</sup>

<sup>1</sup>UMR CNRS7276, Inserm1262, Université de Limoges: Contrôle de la Réponse Immune B et des Lymphoproliférations, 2, rue du Dr. Marcland, 87025 Limoges, France, Team 2, B-NATION: B cell Nuclear Architecture, Immunoglobulin genes and Oncogenes

<sup>2</sup>Current Adress : Laboratoire Suivi des Thérapies Innovantes, Institut de Génétique Humaine, UMR 9002 CNRS-UM ; Montpellier, France

<sup>3</sup>National Institute of Biological Sciences, Beijing, 102206, China.

<sup>4</sup>UMR 1236, Univ Rennes, INSERM, Etablissement Français du Sang Bretagne, Rennes, France.

\* To whom correspondence should be addressed [sandrine.le-noir@unilim.fr](mailto:sandrine.le-noir@unilim.fr) and [eric.pinaud@unilim.fr](mailto:eric.pinaud@unilim.fr)

## **Abstract**

Class switch recombination is an essential process which enabling B cells to adapt immunoglobulin subtypes to antigens. Transcription plays a crucial role in regulating CSR in which the *IgH 3'Regulatory Region (3'RR)* was identified as a key player. The *3'RR* stands at the 3' end of *IgH* locus and is composed of four core enhancers surrounded by inverted repeated sequences, forming a quasi-palindrome. In addition to transcriptional control, nuclear organization appears to be an important level in CSR regulation. Furthermore, the chromatin loops at *IgH* locus facilitate an efficient CSR recombination by bringing the donor and acceptor switch regions closer together. However, the precise control mechanisms governing both of these processes remain partially understood. Here, using the reference DNA 3D-FISH technique combined with various high throughput approaches, we showed that *3'RR* core enhancers are necessary and sufficient to preorganize resting B cell nuclei to facilitate a deletional CSR mechanism at activated stage. We demonstrated that the *3'RR* core enhancers regulate *IgH* locus addressing in the nuclei, control *IgH* locus accessibility and orchestrate *IgH* loops formation. Our findings pinpointed an additional regulation level of mechanisms underlying B cell diversification.

## Introduction

Late stages of B cell development are characterized by immunoglobulin genes remodeling to produce highly specific antibodies. Somatic hypermutation (SHM) and class switch recombination (CSR) are the two main Immunoglobulin Heavy Chain (*IgH*) locus modifications that both require transcription of target sequences and cytidine deamination by the activation-induced cytidine deaminase (AID) enzyme (1). It has been extensively documented in literature that these events are under the transcriptional control of the super-enhancer located at *IgH* locus 3' end: the *IgH* 3' Regulatory Region (3'RR) (2). In mice, this region is composed of four core enhancers, (*hs3a*, *hs1.2*, *hs3b* and *hs4*), which are surrounded by inverted repeated sequence (IRIS) centered on *hs1.2* enhancer and delimited by *hs3a* and *hs3b* (3) (Figure 1A). *IgH* locus forms a topologically associated domain (TAD) as demonstrated by chromosome conformation capture (3C)-based studies (4). The 3' end of the locus is characterized by an insulator region composed of 10 CTCF (CCCTC-binding factor) binding site named 3' CTCF binding elements (3'CBE) which delimited the 3' *IgH* TAD border (5). Several mouse models carrying partial or total deletions of 3'RR region allow to decipher the exact role of both core enhancers and quasi-palindromic structure on SHM, CSR, transcription and immunoglobulin production during B cell development (3). The 3'RR entire deletion leads to a 2-fold decrease of *IgH* germline transcription (GLT), and consequently, totally abolishes SHM and CSR events (6). Interestingly 3'RR core enhancers (*c3'RR* mouse model) are mandatory for full activity of the 3'RR super-enhancer, and the surrounding sequences do not constitute useless junk DNA, but contribute to 3'RR activity: while CSR is supported by enhancers alone, SHM requires both 3'RR enhancers and its palindromic architecture (7, 8).

Beyond transcriptional control of SHM and CSR, additional levels of regulation occurred during late B cell maturation. Most of Ig gene regulation studies have been so far performed at nucleosomal scale (epigenetic modifications and regulatory transcription of loci and gene segments). The increasing interest for the understanding of gene regulation at the whole nuclei scale raises the necessity to revisit previous models at both supranucleosomal (DNA loops and A/B compartments) and nuclear (chromosome territories and gene positioning) levels (9). To date, numerous studies showed that nuclear positioning of *IgH* loci is dynamic and regulated during early and late B cell development and *IgH* remodeling regulation, especially CSR recombination, also occurs at supranucleosomal level (10). CSR mechanism involves DNA double strand breaks (DSB) at *S $\mu$*  donor and *S $\alpha$*  acceptor regions (11). Mammalian switch (S) regions, located upstream of each constant gene except C $\delta$ , consist of 1 to 10 kb-long highly repetitive G-rich DNA sequences containing clusters of RGYW AID deamination hotspots (12). AID-induced DSB at S regions are followed by the deletion of DNA segment located between both DNA DSB and finally replacing *C $\mu$*  constant gene by one of the six *C $\alpha$*  constant gene (*IgG3*, *IgG1*, *IgG2a*, *IgG2b*, *IgA*, *IgE*). The recombination event between DNA ends is essentially ensured by the non-homologous end joining (NHEJ) complex to achieve productive deletional CSR (13). Mature resting B cells ensure an *IgH* locus configuration "prepared for CSR" by bringing in close contact the 3'RR and *E $\mu$* -regions, thereby forming a 50-200Kb "big chromatin loop" that contains all S regions. Upon stimulation, this configuration acquires additional contact between the acceptor S region involved in CSR and the previously interacting *IgH* enhancers (14–16). In this new layout, S donor and S acceptor regions are aligned to favor productive deletional CSR. DSB response factors inhibition impacts the ratio

of deletional vs inversional CSR pathway and promotes long end resection in S regions highlighting the important role of DNA repair pathway for efficient deletional CSR (17–19). The current model proposes that such loops are formed by loop extrusion, a mechanism consisting in loading cohesin complexes at transcribed enhancers or promoters (20–22) followed by chromatin extrusion until cohesin meets convergently oriented CTCF sites (5). In the specific case of *IgH* locus, cohesin is loaded at enhancers (either 3'RR or E $\mu$  regions) to first initiate the “big chromatin loop” between 3'RR and I $\mu$ /S $\mu$  regions at the resting stage. Upon B cell stimulation, new internal loops are shaped to bring together transcribed S regions in order to form the CSR center (CSRC) (15, 16), a structure that favors deletional CSR recombination (17).

The role of 3'RR and its components, core enhancers and palindromic structure, remains to be determined in the context of CSR center formation. Here, by taking advantage of two mouse models devoid of either the entire 3'RR ( $\Delta$ 3'RR (6)) or only its palindromic structure (c3'RR model (7)), we evaluated by multiple approaches the role of 3'RR core enhancers in B cell nuclear organization and *IgH* locus architecture. Interestingly, we showed that the four 3'RR core enhancers are necessary and sufficient to pre-organize resting B cell nucleus to allow deletional CSR recombination at activated stage.

## Materials & Methods

**Mice.** 8- to 10- weeks old *wt*,  $\Delta 3'RR$ ,  $c3'RR$  or  $\Delta hs5-7$  (kindly gifted by Barbara Birshstein) mice were bred and maintained in SOPF conditions at 21–23°C with a 12-h light/dark cycle. Procedures were reviewed and approved by the Ministère de l'Education Nationale de l'Enseignement Supérieur et Recherche autorisation APAFIS#16151-2018071716292105v3.

**Cell Culture.** Splenocytes were collected, non-B cells were removed with mouse B cell kit isolation from stem cell (REF#19854) to collect resting B cells. *In vitro* stimulated B cells were obtained after three or four days of culture at  $1.10^6$  cells/mL in RPMI 1640 medium supplemented with 10% FCS, 2mM Glutamine (Eurobio), 1% Eagle's Non-essential Amino Acids (Eurobio), 50U/ml of penicillin-streptomycin (Gibco), 1mM sodium pyruvate (Eurobio), 129  $\mu$ M 2- $\beta$ mercaptoethanol (Sigma-Aldrich). Cells were stimulated by adding 1 $\mu$ g/mL lipopolysaccharide (LPS EB ultrapure Invivogen).

**Immunization.** Mice were immunized by intraperitoneal injection of 200  $\mu$ L of Sheep Red Blood Cells (SRBC), challenged at day 10 and sacrificed at day 18. Splenic B cells were sorted by using CD43 negative selection (Miltenyi #130-049-801).

**Flow cytometry and Cell Sorting.** To sort transitional B cells, B splenocytes were labeled with anti-B220-APC (RA3-6B2 clone, BD 553092) and anti-CD93-BV421 (AA4.1 clone, BD 561990) conjugated antibodies. Transitional (B220+/CD93+) B cells were sorted on ARIA III FACS (BD Biosciences). Plasmablasts were labelled with anti-B220-BV421 (RA3-6B2 clone, BD 562922) and anti-CD138-APC (281-2 clone, BD 558626).

**DNA 3D-FISH.** DNA 3D FISH was performed as previously described (23). Briefly, resting and *in vitro* stimulated B cells for three days were dropped onto poly-L-lysine slides and fixed with 4% paraformaldehyde for 10min at room temperature (RT). After PBS washes, cells were permeabilized with pepsin 0.02%/HCl 0.1M for 10min at 37°C. Then cells were washed with PBS and post-fixed with 1% paraformaldehyde for 5min at RT. DNA and probes were denatured in 70% formamide/2X SSC at 72°C and at 95°C for 5min respectively. DNA and probes were hybridized at 37°C overnight. The *IgH* and  $\gamma$ -satellite probes (which highlight heterochromatin) were respectively labelled with dCTP-biotin or dUTP-digoxigenin (Invitrogen REF#19518018 or #11573152910) and dUTP-Alexa Fluor 488 (Invitrogen REF#C11397). Slides were washed in 1X SSC at 72°C for 5min and then incubated with streptavidin-Alexa Fluor 594 (Molecular Probes (life) REF#S32356) (1/300<sup>e</sup> in 4X SSC) during 1h at RT and mounted with vectashield containing DAPI (Vector Labs). Images were acquired with an epifluorescence microscope (NIKON). Optical sections separated by 0.2 $\mu$ m were captured and stack were deconvoluted and analyzed using Huygens and Volocity softwares, respectively. Separation of alleles was measured from the center of each signal. Volumetric pixel size was 0.065 $\mu$ m in xy and 0.2 $\mu$ m in z-direction.

**ATAC-Seq.** 10,000 *in vitro* stimulated cells for three days were centrifuged at 500g at 4°C during 10min. Then cells were incubated during 3min in lysis buffer (0,1% Tween-20, 0,1% Nonidet P40, 0,01% Digitonin, 10mM Tris-HCl pH7.4, 10mM NaCl, 3mM MgCl<sub>2</sub>) on ice and resuspended in 10 $\mu$ l of transposition mix (1X TD Buffer from Illumina, Transposase from Illumina, 0,01% digitonin, 0,1% Tween-20). Tagmentation was performed on a thermomixer at 1000rpm 37°C during 30min. Transposed DNA

was cleaned-up with minElute Qiagen kit (#28006). Transposed DNA was amplified by PCR using primers Ad1 (5'-AATGATACGGCGACCACCGAGATCTACACTCGTTCCGGCAGCGTCAGATGTG-3') and Ad2.X (5'-CAAGCAGAAGACGGCATAACGAGAT**BARECODE**GTCTCGTGGGCTCGGAGATGT-3'). Libraries were sequenced 2\*100bp on NovaSeq 6000.

**ATAC-qPCR.** DNA accessibility was quantified in resting B cells by qPCR using Bioline SYBR Hi-ROX kit (BIO-92020). Accessibility of C $\mu$  region was quantified using C $\mu$ -Forward (5'-CTTCCCAAATGTCTTCCCC-3') and C $\mu$ -Reverse (5'-TGCGAGGTGGCTAGGTAAGT-3') as well as C $\gamma$ 1 using C $\gamma$ 1-Forward (5'-GCCCAAATACTCCATGGTC-3') and C $\gamma$ 1-Reverse (5'-CAACGTTGCAGGTGACGGT-3'). C $\mu$  and C $\gamma$ 1 accessibility were normalized to Lig1, using Lig1-Forward (5'-CTCTTCTCCCCGACTGTAC3') and Lig1-Reverse (5'-GAGGCTGCTGGGAGTTGTAG-3'), gene identified as invariant between our sample as described in the "analysis of ATAC-Seq section".

**3C-HTGTS.** 3C-HTGTS was performed as previously described (5). Briefly, 10 million of resting or three days stimulated B cells were crosslinked with 2% formaldehyde 10% FCS PBS for 10min at RT under rotation. Crosslinking was stopped by adding glycine at 0.1M. Then, cells were lysed in 50mM Tris, 150mM NaCl, 5mM EDTA, 0.5%NP-40,1% TX-100 supplemented with protease inhibitor (ROCHE #11873580001). Nuclei were resuspended in 0,3% SDS for 1h at 37°C at 900rpm and then neutralized with Triton TX-100 for 1h. DNA restriction was performed using CviQ1 (Thermo Fisher ER0211) in B buffer (Thermofisher #BB5) overnight at 37 °C, before heat inactivation for 25 min at 65°C. Overnight ligation was performed at 16°C at 300rpm. Next, DNA was treated by proteinase K and RNase and cleaned by phenol/chloroform. After 3C step, the LAM-HTGTS protocol was performed (24). Briefly, 3C DNA was sonicated using the Bioruptor (Diagenode; two pulses at low intensity for 20s), and 10  $\mu$ g was used for the LAM-HTGTS step. A 3'*CBE* (5'biotin-CACTGTCCAGACAGCAAACC-3'), *I $\mu$ /S $\mu$*  (5'biotin-GCAGACCTGGGAATGTATGGT-3') or 3'*RR* (5'biotin-GGACTGCTCTGTGCAACAAC-3') bait was used for primer elongation. These single-stranded DNA fragments were incubated with streptavidin beads (Dynabeads C1 streptavidin beads; Invitrogen) overnight at RT and washed with BW buffer (1M NaCl, 5mM Tris-HCl pH7.4, 0,5mM EDTA pH 8.0). A universal I7 adaptor (5'-GCGACTATAGGGCACGCGTGGNNNNN-3'NH<sub>2</sub> and 5'-P CCACGCGTGCCCTATAGTCGC-3'NH<sub>2</sub>) was ligated before the nested PCR performed with 3'*CBE*-bait nested primer (5'-ACACTCTTCCCTACACGACGCTCTTCCGATCT**BARECODE**ACCGGCATGTTTCATCAACAC-3') or *I $\mu$ /S $\mu$* -bait nested primer (5'-ACACTCTTCCCTACACGACGCTCTTCCGATCT**BARECODE**ACACAAAGACTCTGGACCTC-3') or 3'*RR*-bait nested primer (5'-ACACTCTTCCCTACACGACGCTCTTCCGATCT**BARECODE**CAAGCTGGGGTCAGAGCATG-3') and universal I7 reverse (5'-CTCGGCATTCTGCTGAACCGCTCTTCCGATCTGACTATAGGGCACGCGTGG-3'). After the Tagged PCR with I7 and I5 Illumina primer, PCR products were cleaned using PCR clean-up kit (Macherey-Nagel REF#740609) and validated after migration on BioAnalyser (Agilent). 3C-HTGTS libraries were sequenced 300bp paired-end MiSeq V3 with 20% PhiX.

**LAM-HTGTS.** Genomic DNA (gDNA) was extracted from 10 million of 4 days stimulated B cells or CD43<sup>-</sup> B cells sorted from spleen of SRBC-immunized mice. LAM-HTGTS was performed as previously described (24) with a *Sμ* bait. Data analysis of MiSeq sequencing reads was performed as previously described (25). All sequence alignments were done with the mouse mm10 genome. Analysis of junction structure has been performed with CSReport tool (26).

**Analysis of ATAC-seq.** Nextera sequences were trimmed by TrimGalore (27) with only pairs of reads with their two reads of at least 55bp each were reported in the resulting fastqs and then processed by the nf-core/atacseq pipeline (28) using mm10 for reference genome. Differentially opened chromatin regions from comparisons of different conditions were extracted using DESeq2 (29) from the *DESeqDataSet* previously produce by the pipeline (28).

The normalized counts extracted from the *DESeqDataSet* were used to calculate the coefficient of variation (CV) of each consensus peaks. Peaks with coverage greater than or equal to the mean were ranked by increasing CV to find a suitable region for ATAC-qPCR normalization controls among the samples compared. The third least variable consensus peak hitting Lig1 gene from positions 13278266 to 13279324 of chromosome 7 on mm10 were selected to design primers for ATAC-qPCR normalization controls. Chromatin A/B compartments were inferred from the consensus peaks of nf-core/atacseq analysis (28) at 3000kb resolution by an extended version (github trichelab/compartmap repository) of the original R package Compartmap (30).

**Analysis of 3C-HTGTS.** Sequencing reads were aligned to the mm10 genome and processed as previously described (33). Each 3C-HTGTS library plotted for comparison was normalized by randomly selecting a number of junctions equal to the total number of junctions present in the smallest library in the comparison set.

**Statistical analysis.** Mann Whitney two-tailed tests were used for statistical analysis using GraphPad Prism software (\*p<0.05, \*\*p<0.01, \*\*\*p<0.001, \*\*\*\*p<0.0001).

## Results

### **3'RR core enhancers maintain *IgH* loci in an active compartment in mature resting B cells.**

To study the role of 3'RR core enhancers vs 3'RR palindromic structure, we compared two mouse models carrying total or partial deletion of the 3'RR to *wt* mice. We took advantage of *c3'RR* model (7) carrying only the four core enhancers and  $\Delta$ 3'RR model (6) carrying total deletion of the 3'RR region (Figure 1A). We compare, in these models, transitional splenic B cells (T) collected by flow cytometry (Supplementary Figure S1A), mature resting B cells (R) and *in vitro*-stimulated B cells for 3 days with LPS (S). To avoid any impact of cell subtypes representation in bulk culture at day 3, we verified the percentage of CD138 positive cell after LPS stimulation in the three models. As expected, the percentage of newly generated plasmablasts *in vitro* in the presence of LPS was comparable (Supplementary Figure S1B).

By using ATAC-Seq (Assay for Transposase-Accessible Chromatin) technic, we determined, with a high throughput approach, chromatin accessibility across the genome in resting and *in vitro* stimulated B cells from *wt*, *c3'RR* and  $\Delta$ 3'RR mice. First of all, a principal component analysis was performed to verify reproducibility between samples (Supplementary Figure S2). Then, we first analyzed all differentially opened (DO) chromatin regions. As expected a large number of DO regions was identified upon B cell stimulation (Supplementary Table S1A). In contrast, when we compared mutants to *wt* in resting and in *in vitro* stimulated B cells, only few differentially open chromatin regions were identified (less than 80 regions) (Supplementary Tables S1B and S2). Moreover, when we confronted the percentage of A and B compartments, no drastic change was observed between mutants nor in resting B cell neither in *in vitro* stimulated B cells (Figure 1B). Then, we focused exclusively on the *IgH* locus, we observed a decrease of its accessibility in resting B cells from  $\Delta$ 3'RR mutants in comparison to *wt* (Figure 1C, Top). To confirm this trend, we quantified ATAC-seq normalized read count in three distinct regions across the *IgH* locus (R1: around  $E\mu/S\mu$ , R2: constant regions and R3: 3'CBE; whose coordinates are available in Supplementary Table S3) and we observed a global decrease of *IgH* accessibility throughout the 3' part of the locus (Figure 1C, Middle). In parallel, we aimed to confirm this result by quantifying DNA accessibility by qPCR. For this purpose, we first identified an open region with low variability between samples to normalize qPCR. As expected, we noticed a slight decrease of *IgH* accessibility in both constant regions tested in  $\Delta$ 3'RR resting B cells (Figure 1C, Bottom). Surprisingly, when we compared *c3'RR* deficient and *wt* mice, we showed that the *c3'RR* resting B cells harbour the same profile as the *wt*, demonstrating that the 3'RR core enhancers are sufficient to maintain *IgH* loci in an open chromatin conformation (Figure 1C). In LPS-stimulated B cells, *IgH* locus is also globally less accessible in  $\Delta$ 3'RR model (Figure 1D). A slight decrease of *IgH* accessibility was shown in the *c3'RR* model across the locus.

Subsequently, we assessed addressing of *IgH* loci to pericentromeric heterochromatin (PCH) within B cells nuclei by a DNA 3D-FISH (Fluorescent *In Situ* Hybridization) approach using an *IgH* fluorescent probe (RP23-109B20) encompassing a region lying from  $D_H$  cluster to  $C\gamma 2a$  segment (Figure 2A) and a probe (major  $\gamma$ -satellite) specific to PCH as the "inactive zone". In a wild-type context, in all three cell



types tested, approximately 20% of cells have at least one *IgH* allele in the PCH, with a minor proportion of cells having both *IgH* alleles in PCH (Figure 2B, left). In  $\Delta 3'RR$  mice, around 15% of *IgH* alleles are localized in PCH at transitional stage as in *wt* mice. In contrast, in resting B cells, the proportion of B cells harboring at least one *IgH* allele colocalized to PCH reaches 35% and includes a higher proportion of cells for which both alleles localized to PCH (Figure 2B, middle). This configuration is maintained in *in vitro* stimulated B cells. Surprisingly, when only  $3'RR$  core enhancers, not the surrounding DNA packaging, are reintroduced ( $c3'RR$  model), practically the same profile as in *wt* was observed (Figure 2B, right). Overall, DNA 3D-FISH and ATAC-seq experiments showed a global decrease of *IgH* loci accessibility in resting B cells and in *in vitro* stimulated B cells from  $\Delta 3'RR$ . In contrast,  $c3'RR$  B cells are in an open chromatin configuration almost as in the *wt* context. All together these results show that in *wt* resting B cells, *IgH* loci were widely in an open chromatin compartment and that PCH addressing is governed only by  $3'RR$  core enhancers, and not by palindromic structure. These findings point out a crucial role of the  $3'RR$  core enhancers in regulating the nuclear organization of mature resting B cells.

### **$3'RR$ core enhancers are sufficient to regulate *IgH* loci position dynamics in nuclei of mature B cells**

Using DNA 3D-FISH, we also investigated the position of *IgH* loci within nuclei of transitional, mature resting and *in vitro* stimulated B cells from the three mouse models. We measured the *IgH* inter allelic distance as well as the distance between *IgH* loci and nuclear center at each stage of differentiation cited above. In the *wt* context, we observed that during the transition from transitional to mature resting B cell stage, *IgH* loci move closer together ( $p < 0.0001$ ), and then, move away from each other during *in vitro* activation ( $p = 0.0122$ ) (Figure 2C, Top). In *wt* mice, the same dynamic is observed when we measured the distances between *IgH* loci and the nuclear center: during transition from transitional to resting stage, *IgH* loci move closer to the center of nucleus ( $p < 0.0001$ ), while they move away from it upon *in vitro* stimulation ( $p < 0.0001$ ) (Figure 2C, Bottom). In contrast, B cells deficient for the entire  $3'RR$  ( $\Delta 3'RR$  model) lose the dynamics observed in *wt* context. In fact, the *IgH* interallelic distance remains constant (Figure 2D, Top). Similarly, distance between *IgH* loci and nucleus center does not differ through transitional to resting and stimulated stages (Figure 2D, Bottom). Interestingly, the reintroduction of  $3'RR$  core enhancers ( $c3'RR$  model) alone rescued normal dynamics: *IgH* loci get closer to each other during transitional to resting stage evolution ( $p < 0.0001$ ) and move away from each other upon *in vitro* stimulation ( $p = 0.0004$ ) (Figure 2E, Top). Comparably, in  $c3'RR$  resting B cells, *IgH* loci are closer to the nucleus center than in transitional B cells ( $p < 0.0001$ ) and then, *IgH* loci move away from nucleus center during activation ( $p < 0.0001$ ) (Figure 2E Bottom). These results pinpoint a dynamic relocalization of *IgH* loci during late B cell maturation orchestrated by the  $c3'RR$  enhancers alone and highlight the crucial role carried by the  $3'RR$  core enhancers in the nuclear positioning of *IgH* loci. They allow both the dynamic relative positioning of *IgH* loci to each other, and to the center of nucleus, especially at resting B cell stage. So, it is tempting to speculate that the  $3'RR$  core enhancers pre-organize resting B cell nucleus in order to properly position *IgH* loci for remodeling events occurring upon activation.

### **3'RR core enhancers allow *IgH* loop conformation in mature resting and stimulated B cells**

As a supplemental level of regulation, chromatin loop formation appears as a master regulator of gene remodeling, such as CSR, in B cells. The regulation of loop formation in the context of CSR is not fully described. Our mouse models allowed to assess the respective involvement of components of the 3'RR in *IgH* locus looping formation. To achieve this point, we performed, in resting and LPS-stimulated B cells from *wt* and mutants, high resolution 3C-HTGTS (15) with two complementary viewpoints: one located within the *I $\mu$ /S $\mu$*  region (named *I $\mu$ /S $\mu$*  bait) (Figure 3A) (17) and another one located between *hs5* and *hs6* enhancers within the 3' *IgH* TAD border region conserved between the three models (named 3'*CBE* bait) (Supplementary Figure S4A) for quantifying chromatin loop from bait to prey (Supplementary Table S4A for prey coordinates).

First, 3C-HTGTS performed in *wt* resting B cells with *I $\mu$ /S $\mu$*  bait showed long-range contacts between *I $\mu$ /S $\mu$*  and the 3'RR region (Figure 3B bedgraphs and middle bargraph, Supplementary Figure S3 left and Supplementary Table S4B for individual mice); these data reflect the first level of *IgH* locus folding ("big loop") and are consistent with both pioneer studies based on chromosome capture (14) and the more recent studies built on high throughput techniques (15). In contrast and as expected,  $\Delta$ 3'RR resting B cells were not able to perform any chromatin loop between *I $\mu$ /S $\mu$*  and the 3'RR region as a consequence of the total deletion of the 3'RR region (Figure 3B bedgraphs and middle bargraph and Supplementary Figure S3 left). In contrast, local interactions around *S $\mu$*  are maintained in this model showing that the 3'RR is not implicated in intra *I $\mu$ /S $\mu$*  loops (Figure 3B bedgraphs and right bargraph and Supplementary Figure S3 left). In this model, the decrease of long-range (3'RR) interactions from *I $\mu$*  was not compensated by any new interaction inside the locus (Figure 3C) nor outside the locus at the *eS* region, a phenomenon previously described in the mouse model devoid of the 3'*CBE* region in which the TAD border is disrupted (5) (Figures 3B bedgraphs and left bargraph). Strikingly, in resting *c3'RR* B cells, devoid of the palindromic structure, chromatin loops between *I $\mu$ /S $\mu$*  and 3'RR regions still occurred notably thanks to the *hs1.2* enhancer (Figure 3B bedgraphs and middle bargraph and Supplementary Figure S3 left). This finding was confirmed, albeit to lesser extent, by using the 3'*CBE* bait located downstream from 3' super-enhancer (Supplementary figure S4A). Similar chromatin loops pattern is observed in *wt* and *c3'RR* model (Supplementary figure S4B), harboring both high interaction frequencies between the *I $\mu$ /S $\mu$*  and the 3'RR regions. In contrast, in  $\Delta$ 3'RR, 3C-HTGTS did not detect any chromatin loop between *I $\mu$ /S $\mu$*  and the 3' super-enhancer (Supplementary figure S4B). Altogether, these results highlight, for the first time, a major role of 3' regulatory enhancers in *IgH* locus folding ("big loop", involving *I $\mu$ /S $\mu$*  and the 3' super-enhancer) of resting B cells. While the 3'RR core enhancers are necessary and sufficient for the *IgH* "big loop" formation, the 3'RR DNA packaging (supporting the quasi-palindromic organization) is not essential for such structure.

Then, we analyzed, in *in vitro* stimulated *wt* B cells with LPS for 3 days, the occurrence of chromatin loops by using the *I $\mu$ /S $\mu$*  viewpoint and globally we found the same pattern as in resting B cells with a more pronounced difference between genotypes. In *wt*, DNA loops detected by 3C-HTGTS revealed the "big loop" structure and the occurrence of new interactions between *I $\mu$ /S $\mu$*  and the targeted *S $\gamma$ 3* and *S $\gamma$ 2b* regions that normally undergoes germline transcription in LPS stimulation (Figure 3D, Supplementary Figure S3 right and Supplementary Table S4C for individual mice). This data reflects

the physiological *IgH* "internal loops" formed with targeted switch regions in the CSRC (15). In  $\Delta 3'RR$  model, the difference is even greater at this stage than in resting B cells, the *IgH* big loop is still not detectable, and a drastic decrease of contact frequencies between  $I\mu/S\mu$  and  $S\gamma 3$  or  $S\gamma 2b$  regions is observed (Figure 3D bedgraphs and bargraphs and supplementary figure S3 Right). As in resting B cells, the reduction in long-range interactions upon  $3'RR$  deletion (Figure 3D) was not compensated by any random interactions within *IgH* locus (Figures 3D bedgraphs and 3E) and the local interactions within  $S\mu$  are maintained (Figure 3D bedgraphs and right bargraphs and supplementary figure S3 Right). Interestingly, by using  $I\mu/S\mu$  bait in LPS-stimulated B cells from  $c3'RR$  model, the *IgH* big loop is still detected mainly through the hs1.2 enhancer, but slightly decreased because of the lack of palindromic structure (Figure 3D bedgraphs and right bargraph). On the other hand, contact frequencies between  $I\mu/S\mu$  and  $S\gamma 3$  or  $S\gamma 2b$  regions occurred in  $c3'RR$  model almost at the same level than in *wt* mice (Figure 3D bedgraphs and bargraphs), in accordance with a slight decrease of CSR toward this isotypes (7). 3C-HTGTS performed with the second viewpoint ( $3'CBE$  bait) confirmed all these findings. Indeed, high levels of interaction between the bait and  $I\mu/S\mu$  region were detected in *wt* and  $c3'RR$  mice as well as internal loops involving the  $S\gamma 3$  or  $S\gamma 2b$  regions (Supplementary figure S4C). And, as expected, there is still a defect of interactions between  $I\mu/S\mu$  and targeted switch regions in stimulated  $\Delta 3'RR$  B cells (Supplementary Figure S4C).

These data support several hypotheses regarding the establishment of *IgH* loops during B cell activation. Firstly, it seems quite clear that the  $3'RR$  core enhancers alone, without the packaging of surrounding DNA, are sufficient for the formation of *IgH* chromatin loops, the "big loop" and the internal CSR loops, in resting and in stimulated B cells. Second, the big loop formation at resting stage is a prerequisite for the formation of the internal loops and the CSRC formation in activated B cell. The internal loop defect identified by 3C-HTGTS, reflect the failure to form the CSRC which is in accordance with the strong CSR defect observed in the  $\Delta 3'RR$  model (34).

To fully conclude on the role of  $3'RR$  core enhancers as *IgH* chromatin organizers, we performed 3C-HTGTS on a mouse model carrying the partial deletion of  $3'CBE$  region (deletion from hs5 to hs7) (35) with a bait conserved upon deletion of the entire  $3'RR$  localized downstream *hs4* (bait  $3'RR$ ) (Supplementary Figures S5A and B). In this model, we observe the formation of the *IgH* big loop in resting B cells as well the internal CSR loops in stimulated B cells. All together these results showed that the  $3'CBE$  region is not involved in *IgH* chromatin loops formation (Supplementary Figures S5C and D).

### **$3'RR$ core enhancers are required for proper deletional CSR mechanism**

Productive CSR events are the result of deletional recombination events that occur by joining the correctly juxtaposed  $S\mu$  donor and  $Sx$  acceptor regions in such a way as to achieve deletion of intermediate DNA sequences. In some cases, the recombination event breaks and religates the large intermediate DNA segment between the two S regions in an inverted manner, resulting in an inversion of the *IgH* C gene, hence called inversional CSR, thereby prohibiting functional expression of heavy chain. Deletional or inversional recombinations within the  $S\mu$  region also occurred, reflecting the AID

targeting and leading to unproductive CSR. To further investigate the involvement of 3'RR core enhancers in CSR mechanism, we questioned the type of CSR recombination involved in our mutants. For that, we classified the junctions in two groups: successful CSR corresponding to a productive deletional CSR recombination and unsuccessful CSR including all recombination that occurred in the locus without producing productive CSR (ie: inversional CSR, intra S $\mu$  recombination) (Figure 4A). After *in vitro* stimulation with LPS for four days, we first compared the occurrence of CSR junctions from S $\mu$  to S $\gamma$ 3 and S $\gamma$ 2b by LAM-HTGTS using a bait localized within S $\mu$  (Figure 4A) as previously described (19, 23) and the resulting sequences have been analyzed through the dedicated pipeline. Coordinates used to analyze successful and unsuccessful CSR events are indicated in Table S5. As expected, in the *wt* context, almost half of the junctions were typical to CSR deletional mechanism (Figure 4B). The same trend is observed in the c3'RR model, in which a part of the junctions was in favor of productive deletional CSR (Figure 4B). Despite the drastic decrease of CSR mechanism characterizing the 3'RR-deficient model (6), very few CSR junctions could thus be detected in LPS-stimulated B cells. In this model, the large majority of junctions identified were characteristic of inversional CSR mechanism or occurred within S $\mu$ . To confirm these observations in *in vivo* context, we performed LAM-HTGTS in splenic B cells sorted from SRBC-immunized mice and found the same successful/unsuccessful CSR bias in the  $\Delta$ 3'RR mutant (Figure 4C). We then analyzed structural features of recombination junctions in both groups (successful and unsuccessful CSR) in *in vitro* LPS stimulated B cells and in splenic B cell sorted from SRBC-immunized mice. We found that the large majority of junctions are blunt or either with short deletion or microhomology in the both junction groups, in all models, a typical hallmark of the NHEJ repair pathway (Figure 4D and E). Increase of inversional CSR and long end resection within S $\mu$  were reported in case of specific DNA repair actors inhibition/deletion (17–19), however our findings suggest unsuccessful CSR identified in our  $\Delta$ 3'RR mutants is related to the impairment of chromatin loops and not to DNA repair actors recruitment since most of the junctions were blunt. These results suggested that the chromatin loops at the *IgH* locus allow productive CSR recombination by preventing inversional CSR.

## Conclusion

Our study clearly demonstrates that 3'RR core enhancers are necessary and sufficient to preorganize the resting B cell nucleus. Firstly, we show that they are mandatory for maintaining *IgH* loci in a transcriptionally active compartment. Secondly, we demonstrate that 3'RR core enhancers keep *IgH* loci close to each other and close to the nucleus center in an active compartment at resting B cell stage. Finally, we prove that these enhancers are essential for forming the *IgH* big loop between I $\mu$ /S $\mu$  and the 3'RR regions prior the internal loops involving the S regions at activated B cells stage. Altogether our results highlight that 3'RR core enhancers, beyond their regulatory function on *IgH* transcription (36), regulate *IgH* chromatin loops formation to prepare *IgH* locus to productive deletional CSR. We propose a simplistic model in which 3'RR core enhancers shape, by themselves, resting B cells at nuclear and supranucleosomal levels (See graphical abstract).

As previously described, in *wt* resting B cells, a small proportion of cells harbors an *IgH* locus localized in pericentromeric heterochromatin (37, 38). However, in *in vitro* stimulated B cells, their relative

positioning is still debate. According to Skok's lab, the unproductive *IgH* allele colocalized with PCH and replicated later (37), implying that the non-productive *IgH* allele is excluded only by its nuclear location. In contrast, we and the De Latt group showed, in *in vitro* activated B cells, that both *IgH* loci remain in euchromatin (38). This finding is in accordance with *IgH* biallelic expression observed in mature B cells (39, 40) and the occurrence of SHM and CSR on both *IgH* alleles (41). In this study, we found that *IgH* locus from  $\Delta 3'RR$  B cells is less accessible, which is consistent with previous ChIP-qPCR experiments that demonstrated a decrease of active chromatin marks (using anti-H3K9Ac antibody) in both constant and *S $\mu$*  regions in  $\Delta 3'RR$  stimulated B cells (34). This observation perfectly correlates with the BCR low phenotype found in this model in mature B cell subpopulations and enforces the role of *3'RR* in late stages of B cell differentiation (42). Moreover, *IgH* locus accessibility is in agreement with *IgH* constant region transcription; in *c3'RR* model the level of constant regions transcription is similar to the *wt* context whereas in  $\Delta 3'RR$  mice the transcription is largely decreased but not abolished (7). The reduced accessibility of the *IgH* locus in the  $\Delta 3'RR$  model may be explained by binding of chromatin-regulating factors in the *3'RR* region. It is known that the *3'RR* super-enhancer recruits Brg1, a protein forming part of the SWI/SNF chromatin remodeling complex, known to preferentially bind the enhancers with the aim to open the chromatin and activate the transcription. This factor was initially described to bind the *3'RR* in immature B cells (43, 44). More recently, Brg1 was found to bind the *3'RR* in mature B cells especially during the GC formation (45). ZMYND8, a protein described as a "chromatin reader", is also known to bind the *3'RR*. The conditional deletion of ZMYND8 in B cell results in a phenotype similar to the *3'RR* deletion (reduced CSR and SHM) suggesting the involvement of this factor in regulation of *IgH* locus at supranucleosomal scale (46).

On another scale, our data show that *3'RR* core enhancers are required for the establishment of chromatin loops at the *IgH* locus during CSR. However the Pavri Group recently published that the *IgH* big loop was mostly mediated by the *3'CBE* region and not by the *3'RR* region (47). This study was carried out only under LPS+IL4 stimulation conditions, inducing a switch to IgG1, an isotype known to be slightly differently regulated from other isotypes, since the switch to IgG1 persists, to a lesser extent, in the  $\Delta 3'RR$  model. Moreover, in the  $\Delta$ *hs567* mice (partial *3'CBE* deletion), it was shown, by 3C experiments, that the *IgH* big loops occurred as in a *wt* context in line with a normal IgG1 CSR recombination (35). In addition, in a model of *3'CBE* total deletion, a decrease of CSR is observed to all isotypes, except IgG1, suggesting a particular link between the *3'CBE* region and the CSR toward IgG1. To enforce the idea that IgG1 is a singular isotype, the Pavri Group showed that in LPS activation condition, a classical interaction between *I $\mu$ /S $\mu$*  and the *3'RR* region which is in line with our study.

$\Delta 3'RR$  and *c3'RR* models showed a drastic decrease in SHM (23) suggesting that the mechanism governing the nucleus organization and three-dimensional topology of *IgH* locus are different between GC and in *in vitro*-activated B cells. So "big loop", putting the *E $\mu$ MARs* and *3'RR* regions in contact, seems to not be involved in regulating the SHM process. This suggests a role of the palindromic structure in the formation of another loop, bringing the *3'RR* and *E $\mu$ MARs* regions into contact with the rearranged *pV<sub>H</sub>*.

The role of NHEJ DNA repair was important for maintaining the ratio of deletional vs inversionsal CSR pathway and long end resection in S regions (17–19). The  $\Delta 3'RR$  harbor the same bias of successful vs

unsuccessful CSR than in DSB response factors inhibition models. Moreover in these models the junctions present more microhomology that is not the case in the  $\Delta 3'RR$  mutant. This suggests that the 3'RR is not involved in the recruitment of NHEJ factors.

Described in 2019 by F. Alt's team, chromatin loop formation at *IgH* locus is mediated by the loop extrusion mechanism (15). It involves loading of the cohesin complex onto DNA by the NIPBL protein, extrusion of DNA into the ring of cohesin complex until encountering two converging CTCF binding sites, followed by dissociation of the cohesin complex from DNA by the WAPL protein. In the specific case of *IgH* locus, cohesin is loaded at enhancers but in the absence of the entire 3'RR, no loop was found throughout the *IgH* locus. It is tempting to speculate that the cohesin complex loading onto the *IgH* locus requires the 3'RR core enhancers.

**Data Availability.** Raw data from ATAC-seq, 3C-HTGTS and LAM-HTGTS have been deposited in the European Nucleotide Archive database under access number PRJEB52320.

### **Acknowledgments**

We thank BISCEm and the animal core facility team for help with mouse work on both practical and regulatory aspects. We are grateful to Emeline Lhuillier and the Genotoul Plateau GeT-Santé facility for technical assistance with ATAC-Seq sequencing. We thank Vonick Sibut for bio-informatic analysis advices on ATAC-Seq. We thank Adam Yongxin Ye and Xuefei Zhang for precious advices on 3C-HTGTS experiments (technical part and bio-informatic analysis), Mehdi Alizadeh and Emilie Guérin for advices on sequencing step. We are grateful to Dr Brice Laffleur for fruitful scientific discussion and critical reading of the manuscript. We also thank Pedro P Rocha for scientific and technical discussions. We thank Christelle Oblet for her invaluable help in creating the graphical abstracts. CB was supported by PhD fellowships from the French Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation. MT was supported by PhD fellowships from the French Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation and the Fonds Européen de Développement Régional (FEDER). JP was supported by Institut CARNOT CALYM. This work was supported by Agence Nationale de la recherche (ANR-21-CE15-0001-0), La Ligue Contre le Cancer (comités 87, 23 to EP and SLN) and Fondation ARC pour la recherche sur le cancer (PhD continuation fellowship to CB and MT).

### **Author contribution**

CB, MT and SLN performed experiments. EP and SLN conceived and supervised the study. JP performed the bio-informatic analysis. DR and ZB were very helpful respectively for ATAC-seq and 3C-HTGTS experiments. CB, EP and SLN wrote the manuscript.

**Competing financial interest:** The authors declare no competing financial interests.

## **REFERENCES**

1. Muramatsu,M., Kinoshita,K., Fagarasan,S., Yamada,S., Shinkai,Y. and Honjo,T. (2000) Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*, **102**, 553–563.
2. Pinaud,E., Marquet,M., Fiancette,R., Péron,S., Vincent-Fabert,C., Denizot,Y. and Cogné,M. (2011) The IgH locus 3' regulatory region: pulling the strings from behind. *Adv. Immunol.*, **110**, 27–70.
3. Bruzeau,C., Moreau,J., Le Noir,S. and Pinaud,E. (2021) Panorama of stepwise involvement of the IgH 3' regulatory region in murine B cells. *Adv. Immunol.*, **149**, 95–114.
4. Extremely Long-Range Chromatin Loops Link Topological Domains to Facilitate a Diverse Antibody Repertoire - PubMed.
5. Zhang,X., Yoon,H.S., Chapdelaine-Williams,A.M., Kyritsis,N. and Alt,F.W. (2021) Physiological role of the 3'IgH CBEs super-anchor in antibody class switching. *Proc. Natl. Acad. Sci. U. S. A.*, **118**, e2024392118.

6. Vincent-Fabert,C., Fiancette,R., Pinaud,E., Truffinet,V., Cogné,N., Cogné,M. and Denizot,Y. (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.
7. Le Noir,S., Boyer,F., Lecardeur,S., Brousse,M., Oruc,Z., Cook-Moreau,J., Denizot,Y. and Cogné,M. (2017) Functional anatomy of the immunoglobulin heavy chain 3' super-enhancer needs not only core enhancer elements but also their unique DNA context. *Nucleic Acids Res.*, **45**, 5829–5837.
8. Saintamand,A., Vincent-Fabert,C., Garot,A., Rouaud,P., Oruc,Z., Magnone,V., Cogné,M. and Denizot,Y. (2016) Deciphering the importance of the palindromic architecture of the immunoglobulin heavy-chain 3' regulatory region. *Nat. Commun.*, **7**, 10730.
9. Ea,V., Baudement,M.-O., Lesne,A. and Forné,T. (2015) Contribution of Topological Domains and Loop Formation to 3D Chromatin Organization. *Genes*, **6**, 734–750.
10. Bruzeau,C., Cook-Moreau,J., Pinaud,E. and Le Noir,S. (2022) Contribution of Immunoglobulin Enhancers to B Cell Nuclear Organization. *Front. Immunol.*, **13**, 877930.
11. Jung,S., Rajewsky,K. and Radbruch,A. (1993) Shutdown of class switch recombination by deletion of a switch region control element. *Science*, **259**, 984–987.
12. Stavnezer,J. (1996) Immunoglobulin class switching. *Curr. Opin. Immunol.*, **8**, 199–205.
13. Iwasato,T., Shimizu,A., Honjo,T. and Yamagishi,H. (1990) Circular DNA is excised by immunoglobulin class switch recombination. *Cell*, **62**, 143–149.
14. Wuerffel,R., Wang,L., Grigera,F., Manis,J., Selsing,E., Perlot,T., Alt,F.W., Cogne,M., Pinaud,E. and Kenter,A.L. (2007) S-S Synapsis during Class Switch Recombination Is Promoted by Distantly Located Transcriptional Elements and Activation-Induced Deaminase. *Immunity*, **27**, 711–722.
15. Zhang,X., Zhang,Y., Ba,Z., Kyritsis,N., Casellas,R. and Alt,F.W. (2019) Fundamental roles of chromatin loop extrusion in antibody class switching. *Nature*, 10.1038/s41586-019-1723-0.
16. Zhang,Y., Zhang,X., Dai,H.-Q., Hu,H. and Alt,F.W. (2022) The role of chromatin loop extrusion in antibody diversification. *Nat. Rev. Immunol.*, 10.1038/s41577-022-00679-3.
17. Dong,J., Panchakshari,R.A., Zhang,T., Zhang,Y., Hu,J., Volpi,S.A., Meyers,R.M., Ho,Y.-J., Du,Z., Robbiani,D.F., *et al.* (2015) Orientation-specific joining of AID-initiated DNA breaks promotes antibody class switching. *Nature*, **525**, 134–139.
18. Panchakshari,R.A., Zhang,X., Kumar,V., Du,Z., Wei,P.-C., Kao,J., Dong,J. and Alt,F.W. (2018) DNA double-strand break response factors influence end-joining features of IgH class switch and general translocation junctions. *Proc. Natl. Acad. Sci.*, **115**, 762–767.
19. Vincendeau,E., Wei,W., Zhang,X., Planchais,C., Yu,W., Lenden-Hasse,H., Cokelaer,T., Pipoli da Fonseca,J., Mouquet,H., Adams,D.J., *et al.* (2022) SHLD1 is dispensable for 53BP1-dependent V(D)J recombination but critical for productive class switch recombination. *Nat. Commun.*, **13**, 1–16.



20. Zuin, J., Franke, V., van IJcken, W.F.J., van der Sloot, A., Krantz, I.D., van der Reijden, M.I.J.A., Nakato, R., Lenhard, B. and Wendt, K.S. (2014) A Cohesin-Independent Role for NIPBL at Promoters Provides Insights in CdLS. *PLoS Genet.*, **10**, e1004153.
21. Dorsett, D. and Merckenschlager, M. (2013) Cohesin at active genes: a unifying theme for cohesin and gene expression from model organisms to humans. *Curr. Opin. Cell Biol.*, **25**, 327–333.
22. Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., *et al.* (2010) Mediator and Cohesin Connect Gene Expression and Chromatin Architecture. *Nature*, **467**, 430–435.
23. Le Noir, S., Laffleur, B., Carrion, C., Garot, A., Lecardeur, S., Pinaud, E., Denizot, Y., Skok, J. and Cogné, M. (2017) The IgH locus 3' cis-regulatory super-enhancer co-opts AID for allelic transvection. *Oncotarget*, **8**, 12929–12940.
24. Hu, J., Meyers, R.M., Dong, J., Panchakshari, R.A., Alt, F.W. and Frock, R.L. (2016) Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat. Protoc.*, **11**, 853–871.
25. Frock, R.L., Hu, J., Meyers, R.M., Ho, Y.-J., Kii, E. and Alt, F.W. (2015) Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. *Nat. Biotechnol.*, **33**, 179–186.
26. Boyer, F., Boutouil, H., Dalloul, I., Dalloul, Z., Cook-Moreau, J., Aldigier, J.-C., Carrion, C., Herve, B., Scaon, E., Cogné, M., *et al.* (2017) CSReport: A New Computational Tool Designed for Automatic Analysis of Class Switch Recombination Junctions Sequenced by High-Throughput Sequencing. *J. Immunol. Baltim. Md 1950*, **198**, 4148–4155.
27. Krueger, F., James, F., Ewels, P., Afyounian, E., Weinstein, M., Schuster-Boeckler, B., Hulselmans, G. and sclamons (2019) FelixKrueger/TrimGalore: v0.6.5 Zenodo.
28. Patel, H., Espinosa-Carrasco, J., Ewels, P., bot, nf-core, Langer, B., Syme, R., Peltzer, A., Behrens, D., Gabernet, G., Garcia, M.U., *et al.* (2022) nf-core/atacseq: nf-core/atacseq v2.0 - Iron Iguana Zenodo.
29. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
30. Fortin, J.-P. and Hansen, K.D. (2015) Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data. *Genome Biol.*, **16**, 180.
31. Lee, C.T., Cavalcante, R.G., Lee, C., Qin, T., Patil, S., Wang, S., Tsai, Z.T.Y., Boyle, A.P. and Sartor, M.A. (2020) Poly-Enrich: count-based methods for gene set enrichment testing with genomic regions. *NAR Genomics Bioinforma.*, **2**, lqaa006.
32. Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P. and Mesirov, J.P. (2011) Molecular signatures database (MSigDB) 3.0. *Bioinformatics*, **27**, 1739–1740.
33. Hu, J., Meyers, R.M., Dong, J., Panchakshari, R.A., Alt, F.W. and Frock, R.L. (2016) Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat. Protoc.*, **11**, 853–871.

34. Saintamand,A., Rouaud,P., Saad,F., Rios,G., Cogné,M. and Denizot,Y. (2015) Elucidation of IgH 3' region regulatory role during class switch recombination via germline deletion. *Nat. Commun.*, **6**, 1–11.
35. Volpi,S.A., Verma-Gaur,J., Hassan,R., Ju,Z., Roa,S., Chatterjee,S., Werling,U., Hou,H., Will,B., Steidl,U., *et al.* (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. *J. Immunol. Baltim. Md 1950*, **188**, 2556–2566.
36. Le Noir,S., Bonaud,A., Hervé,B., Baylet,A., Boyer,F., Lecardeur,S., Oruc,Z., Sirac,C. and Cogné,M. (2021) IgH 3' regulatory region increases ectopic class switch recombination. *PLoS Genet.*, **17**, e1009288.
37. Skok,J.A., Brown,K.E., Azuara,V., Caparros,M.L., Baxter,J., Takacs,K., Dillon,N., Gray,D., Perry,R.P., Merckenschlager,M., *et al.* (2001) Nonequivalent nuclear location of immunoglobulin alleles in B lymphocytes. *Nat. Immunol.*, **2**, 848–854.
38. Holwerda,S.J.B., van de Werken,H.J.G., Ribeiro de Almeida,C., Bergen,I.M., de Bruijn,M.J.W., Verstegen,M.J.A.M., Simonis,M., Splinter,E., Wijchers,P.J., Hendriks,R.W., *et al.* (2013) Allelic exclusion of the immunoglobulin heavy chain locus is independent of its nuclear localization in mature B cells. *Nucleic Acids Res.*, **41**, 6905–6916.
39. Daly,J., Licence,S., Nanou,A., Morgan,G. and Mårtensson,I.-L. (2007) Transcription of productive and nonproductive VDJ-recombined alleles after IgH allelic exclusion. *EMBO J.*, **26**, 4273–4282.
40. Tinguely,A., Chemin,G., Péron,S., Sirac,C., Reynaud,S., Cogné,M. and Delpy,L. (2012) Cross talk between immunoglobulin heavy-chain transcription and RNA surveillance during B cell development. *Mol. Cell. Biol.*, **32**, 107–117.
41. Chaudhuri,J., Basu,U., Zarrin,A., Yan,C., Franco,S., Perlot,T., Vuong,B., Wang,J., Phan,R.T., Datta,A., *et al.* (2007) Evolution of the immunoglobulin heavy chain class switch recombination mechanism. *Adv. Immunol.*, **94**, 157–214.
42. Garot,A., Marquet,M., Saintamand,A., Bender,S., Le Noir,S., Rouaud,P., Carrion,C., Oruc,Z., Bébin,A.-G., Moreau,J., *et al.* (2016) Sequential activation and distinct functions for distal and proximal modules within the IgH 3' regulatory region. *Proc. Natl. Acad. Sci. U. S. A.*, **113**, 1618–1623.
43. Holley,D.W., Groh,B.S., Wozniak,G., Donohoe,D.R., Sun,W., Godfrey,V. and Bultman,S.J. (2014) The BRG1 chromatin remodeler regulates widespread changes in gene expression and cell proliferation during B cell activation. *J. Cell. Physiol.*, **229**, 44–52.
44. Osipovich,O.A., Subrahmanyam,R., Pierce,S., Sen,R. and Oltz,E.M. (2009) Cutting edge: SWI/SNF mediates antisense Igh transcription and locus-wide accessibility in B cell precursors. *J. Immunol. Baltim. Md 1950*, **183**, 1509–1513.
45. Schmiedel,D., Hezroni,H., Hamburg,A. and Shulman,Z. (2021) Brg1 Supports B Cell Proliferation and Germinal Center Formation Through Enhancer Activation. *Front. Immunol.*, **12**, 705848.
46. Delgado-Benito,V., Rosen,D.B., Wang,Q., Gazumyan,A., Pai,J.A., Oliveira,T.Y., Sundaravinayagam,D., Zhang,W., Andreani,M., Keller,L., *et al.* (2018) The Chromatin Reader

ZMYND8 Regulates Igh Enhancers to Promote Immunoglobulin Class Switch Recombination. *Mol. Cell*, **72**, 636-649.e8.

47. Costea,J., Schoeberl,U.E., Malzl,D., von der Linde,M., Fitz,J., Gupta,A., Makharova,M., Goloborodko,A. and Pavri,R. (2023) A de novo transcription-dependent TAD boundary underpins critical multiway interactions during antibody class switch recombination. *Mol. Cell*, 10.1016/j.molcel.2023.01.014.

## Legend to figures

### Figure 1. 3'RR core enhancers maintained *IgH* loci in an active compartment in mature resting B cells.

**A.** Murine *IgH* locus scheme (not to scale). *IgH* constant genes are represented by squares, arrows represent the palindromic structure. Deletion of  $\Delta 3'RR$  (6, 34) and  $c3'RR$  (7) mice used in this study are schemed. **B.** A and B compartment proportions in resting and stimulated B cells from *wt* and mutants **C.** *IgH* locus accessibility by ATAC-Seq in splenic resting B cells from *wt* (n=2 independent mice),  $3'RR$  (n=2 independent mice) and  $c3'RR$  mice (n=2 independent mice). Bedgraphs are normalized to reads per million and visualized with IGV. *IgH* loci was split into three regions and the normalized reads coverage were counted in each region as indicated in each bargraph. Relative DNA accessibility by qPCR in resting B cells. Normalization was done to *Lig1* which is not differentially opened between models. q-PCR probes are indicated with black lines. **D.** As in B for *in vitro* LPS stimulated B cell from *wt* (n=2 independent mice),  $\Delta 3'RR$  (n=1 mouse) and  $c3'RR$  (n=2 independent mice).

### Figure 2. 3'RR core enhancers are sufficient to maintained *IgH* loci close together and close to the center in mature resting B cells

**A.** Localization of RP23-109B20 FISH probe is represented in red. **B.** Top: Bars graphs represent the percentage of *IgH* loci into pericentromeric heterochromatin (PCH, stained by  $\gamma$ -satellite major probe) in transitional (T) (*wt* = 116 nuclei analyzed from 3 independent mice) ( $\Delta 3'RR$  = 75 nuclei analyzed from 3 independent mice) ( $c3'RR$  = 108 nuclei analyzed from 2 independent mice), in resting (R) (*wt* = 232 nuclei analyzed from 3 independent mice) ( $\Delta 3'RR$  = 136 nuclei analyzed from independent mice) ( $c3'RR$  = 199 nuclei analyzed from 3 independent mice) and in stimulated (S) (*wt* = 239 nuclei analyzed from 4 independent mice) ( $\Delta 3'RR$  = 249 nuclei analyzed from 4 independent mice) ( $c3'RR$  = 740 nuclei analyzed from 3 independent mice) B cells by DNA-3D-FISH. Light and dark green respectively corresponds to the percentage of cells harboring one or two *IgH* allele in PCH. Bottom: Representative nuclei of resting B cells from each model (*IgH* loci in red / PCH in green, scale bars: 5 $\mu$ m). **C.** Violin plots represent raw *IgH-IgH* interallelic distances in  $\mu$ m from transitional to stimulated B cell stages in each model. **D.** As in (C) for *IgH*-nucleus center distances.

### Figure 3. 3'RR core enhancers are mandatory for *IgH* big loop and internal loops formation at resting and stimulated B cell stages.

**A.** Representation of murine *IgH* locus, the yellow arrow represents the 3C-HTGTS bait. **B.** Top: Representative bedgraphs from splenic resting B cell sample for *wt* (n=5 independent mice),  $\Delta 3'RR$  (n=3 independent mice) and  $c3'RR$  (n=2 independent mice) models, visualized with IGV. Bedgraphs were normalized to smallest sample. Bottom: Bargraphs of normalized reads count. **C.** Quantification of normalized reads count in the entire *IgH* locus in splenic resting B cells from *wt* and mutant mice. **D.** Same as in (B) in *wt* (n=4 independent mice),  $\Delta 3'RR$  (n=3 independent mice) and  $c3'RR$  (n=3 independent mice) *in vitro* LPS-stimulated B cells. **E.** Same as in (C). Scale is adapted to see interactions of interest.

### Figure 4. Deletional CSR in stimulated B cells requires 3'RR core enhancers.

**A.** Schematic representation of successful and unsuccessful CSR events. Successful CSR corresponds to classical deletional CSR between  $S\mu$  and targeted switch regions with excision of the intermediate DNA segment. Unsuccessful CSR events corresponds to inversional CSR, deletional intra- $S\mu$  and inversional intra- $S\mu$ . Location of the  $S\mu$  LAM-HTGTS bait is indicated in yellow. **B.** Bargraphs represent the proportion of successful CSR (in white) and unsuccessful CSR (in dark grey) determined by LAM-HTGTS in LPS stimulated B cells from *wt* (n=5 independent mice),  $\Delta 3'RR$  (n=3 independent mice) and  $c3'RR$  (n=3 independent mice) models. **C.** Same as in B for splenic B cells sorted from SRBC immunized mice (*wt* : n=2 independent mice,  $\Delta 3'RR$  : n = 3 independent mice,

c3'RR : n=1 mouse). **D.** Structure of successful (Right, white bargraphs) and unsuccessful (Left, dark grey bargraphs) CSR junctions from LPS stimulated splenic B cells analyzed with CSReport (26). Only mice with recombination events are shown. **E.** Same as in D for splenic B cells sorted from SRBC immunized mice.

#### **Graphical abstract.**

Position of *IgH* loci through B cell development (from transitional to stimulated stages) is represented by the red spots. In *wt* and *c3'RR* context, *IgH* loci get closer from each other and from nucleus center during evolution from transitional to mature resting stage and they relocate more at nuclear periphery, away one from each other, upon *in vitro* stimulation. In  $\Delta 3'RR$  model, this dynamic is lost and, moreover, *IgH* loci are more localized to pericentromeric heterochromatin (represented by green area) since the mature resting B cell stage and remain in after *in vitro* stimulation.

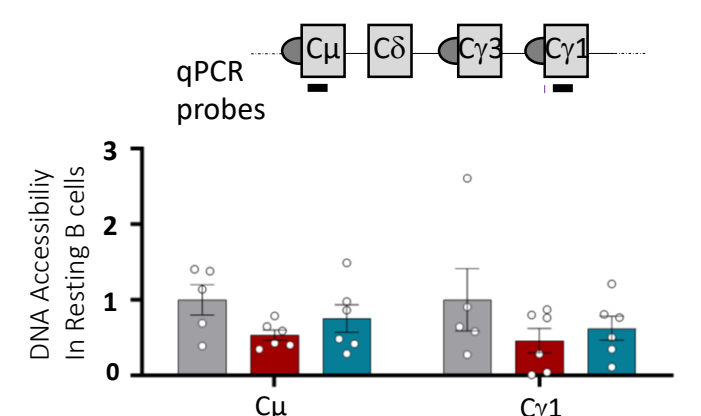
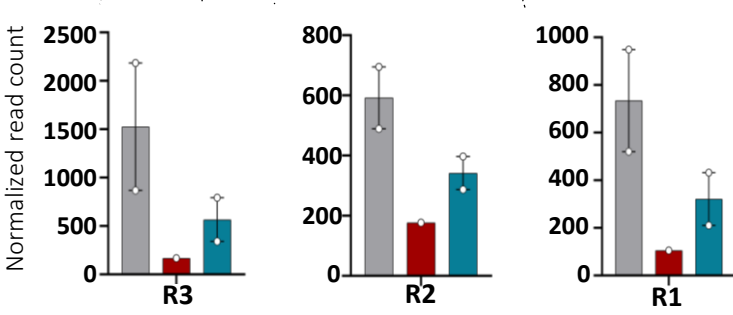
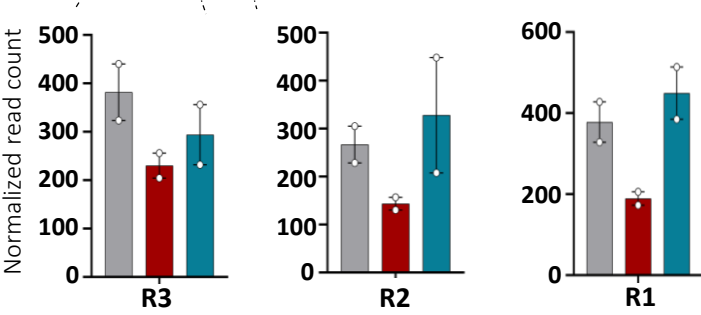
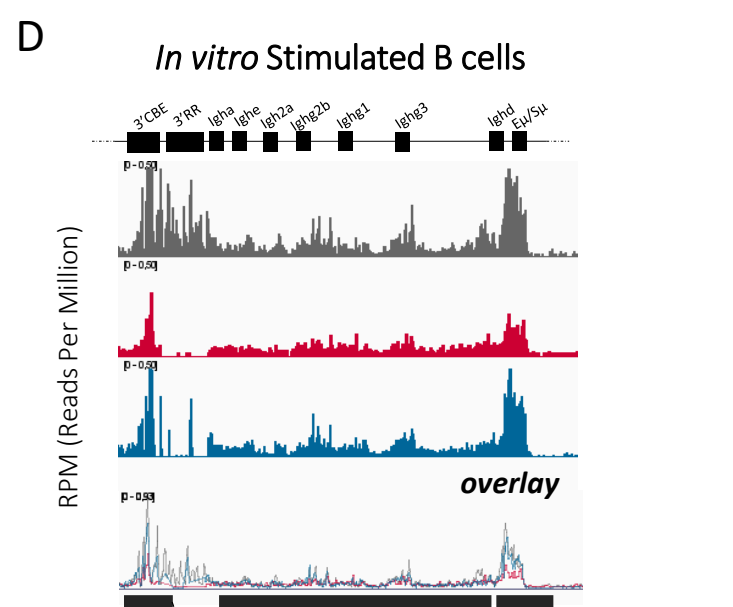
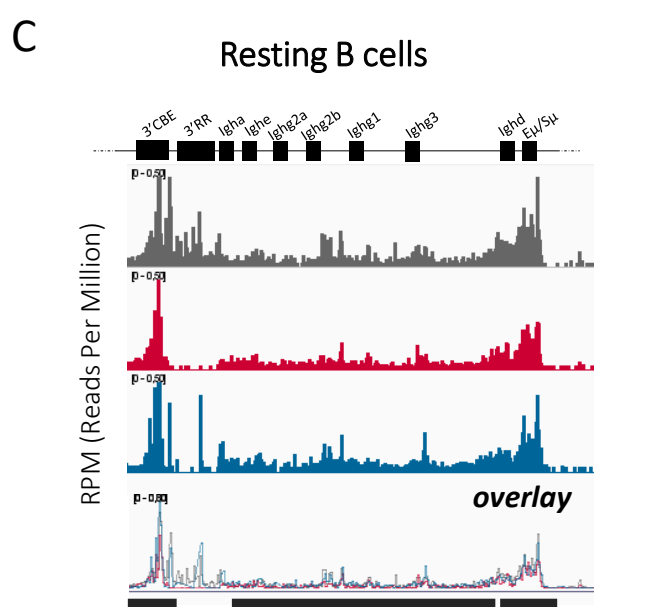
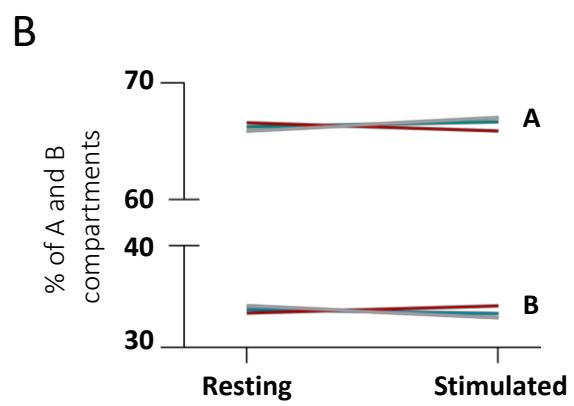
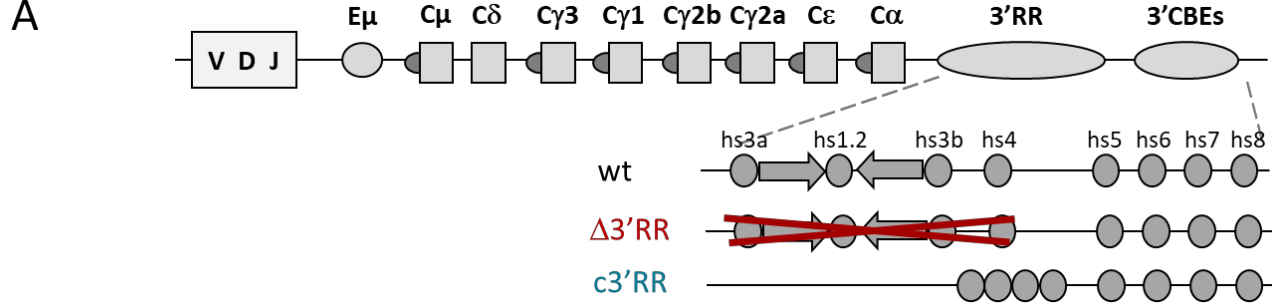
**Supplementary Figure S1:** **A.** Gating strategy to sort transitional B cells from spleen. **B.** Top: gating strategy to identified plasmablasts *in vitro* stimulated splenic B cells from *wt*. Bottom: bargraphs represented the percentage of plasmablast in *wt*,  $\Delta 3'RR$  and *c3'RR* mice (n=6-9).

**Supplementary Figure S2:** Principal Component Analysis of ATAC-Seq of resting (R, n=2) and *in vitro* LPS stimulated (J3, n=1-2) B cells from *wt*,  $\Delta 3'RR$  and *c3'RR* mice.

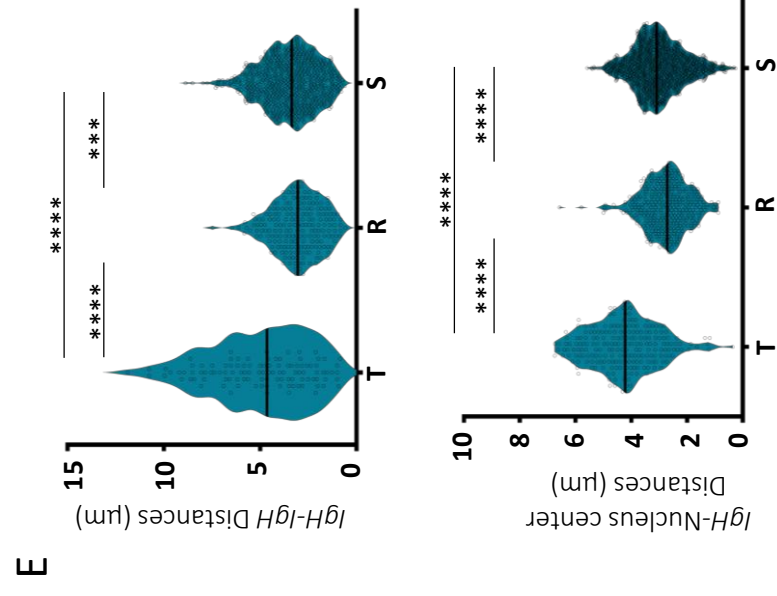
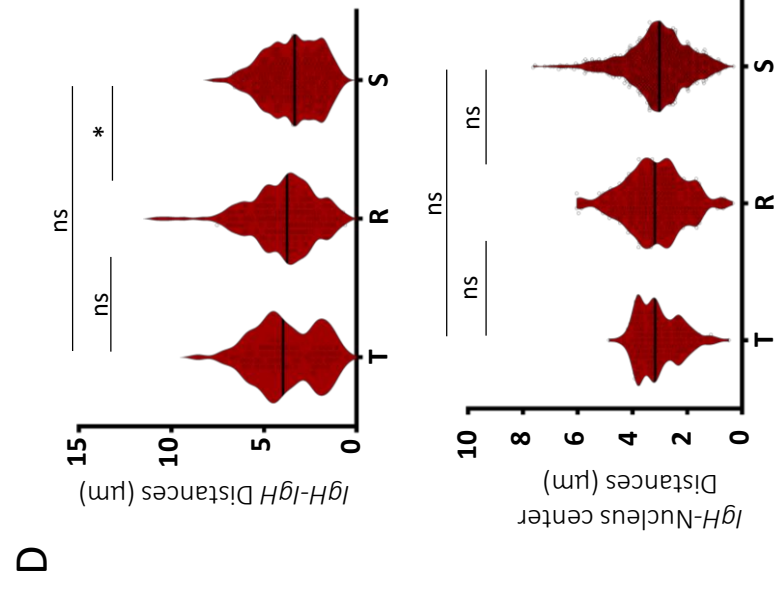
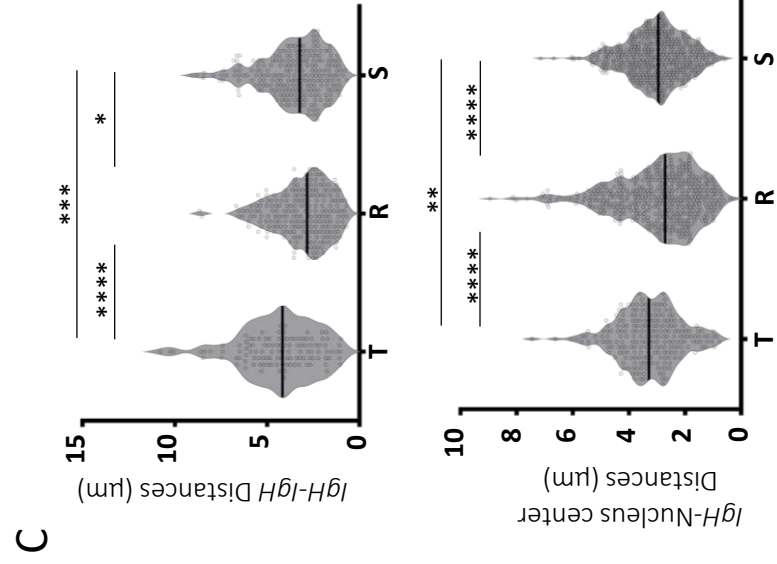
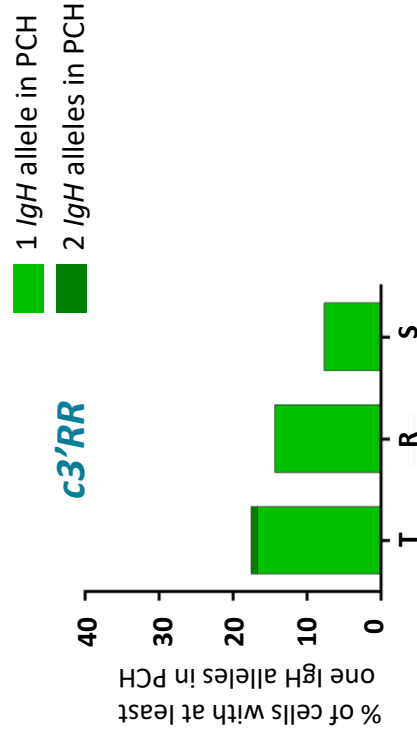
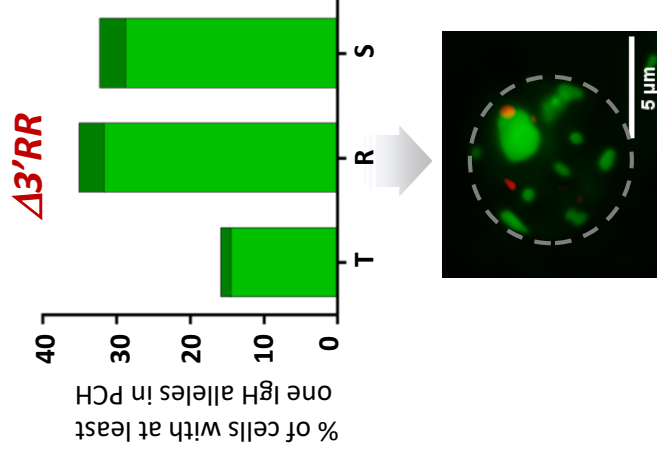
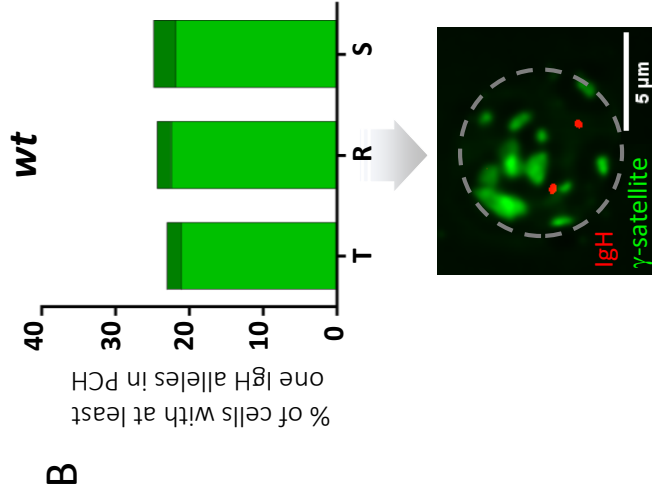
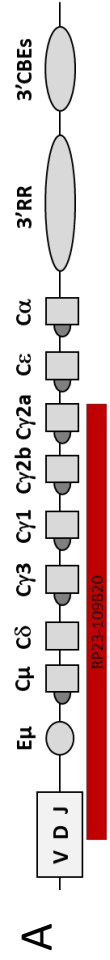
**Supplementary Figure S3:** **Top.** Scheme of murine *IgH* locus.  $I\mu/S\mu$  bait is represented by yellow arrow. **Left.** Bedgraphs, visualized with IGV, of replicates of splenic resting B cells from *wt* (n=5 independent mice),  $\Delta 3'RR$  (n=3 independent mice) and *c3'RR* (n=2 independent mice). **Right.** Same as in B for *in vitro* stimulated B cells from *wt* (n=4 independent mice),  $\Delta 3'RR$  (n=3 independent mice) and *c3'RR* (n=3 independent mice) mice. Scale is adapted to see interactions of interest.

**Supplementary Figure S4:** **A.** Scheme of murine *IgH* locus. 3'CBE bait is represented by yellow arrow. **B.** Bedgraphs, visualized with IGV, of replicates of splenic resting B cells from *wt* (n=2 independent mice),  $\Delta 3'RR$  (n=2 independent mice) and *c3'RR* (n=2 independent mice). **C.** Same as in B for *in vitro* stimulated B cells from *wt* (n=3 independent mice),  $\Delta 3'RR$  (n=2 independent mice) and *c3'RR* (n=2 independent mice) mice. Scale is adapted to see interactions of interest.

**Supplementary Figure S5:** **A.** Scheme of murine *IgH* locus and representation of  $\Delta 3'RR$  and  $\Delta hs567$  mouse models. Deletion are represented by red crosses **B.** 3'RR bait is represented by yellow arrow. **C.** Bedgraphs, visualized with IGV, of replicates of splenic resting B cells from *wt* (n=3 independent mice),  $\Delta 3'RR$  (n=3 independent mice) and  $\Delta hs567$  (n=3 independent mice). **D.** Same as in C for *in vitro* stimulated B cells from *wt* (n=3 independent mice),  $\Delta 3'RR$  (n=2 independent mice) and  $\Delta hs567$  (n=2 independent mice) mice. Scale is adapted to see interactions of interest.



**Figure 1**

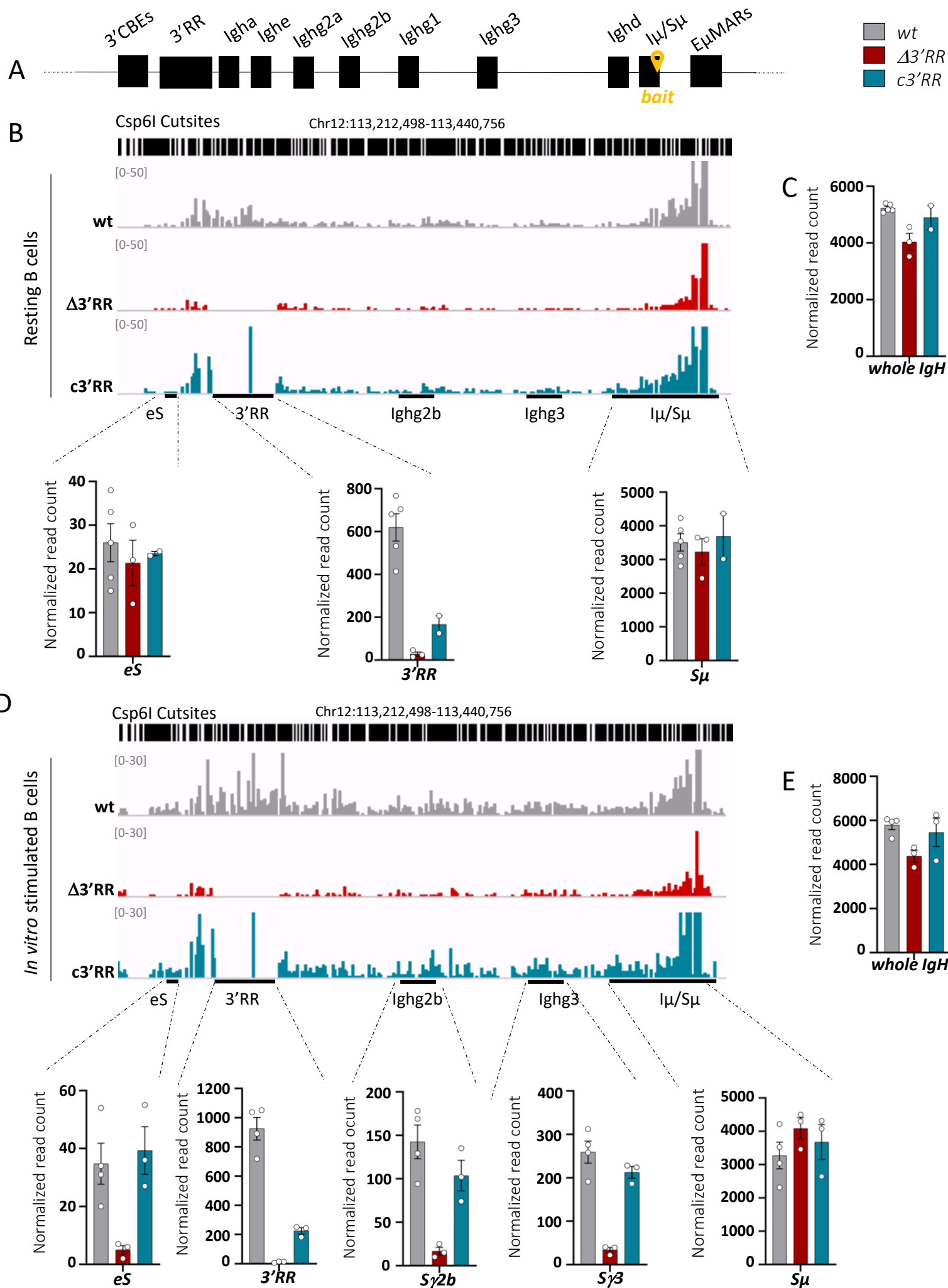


**Legend**

wt █

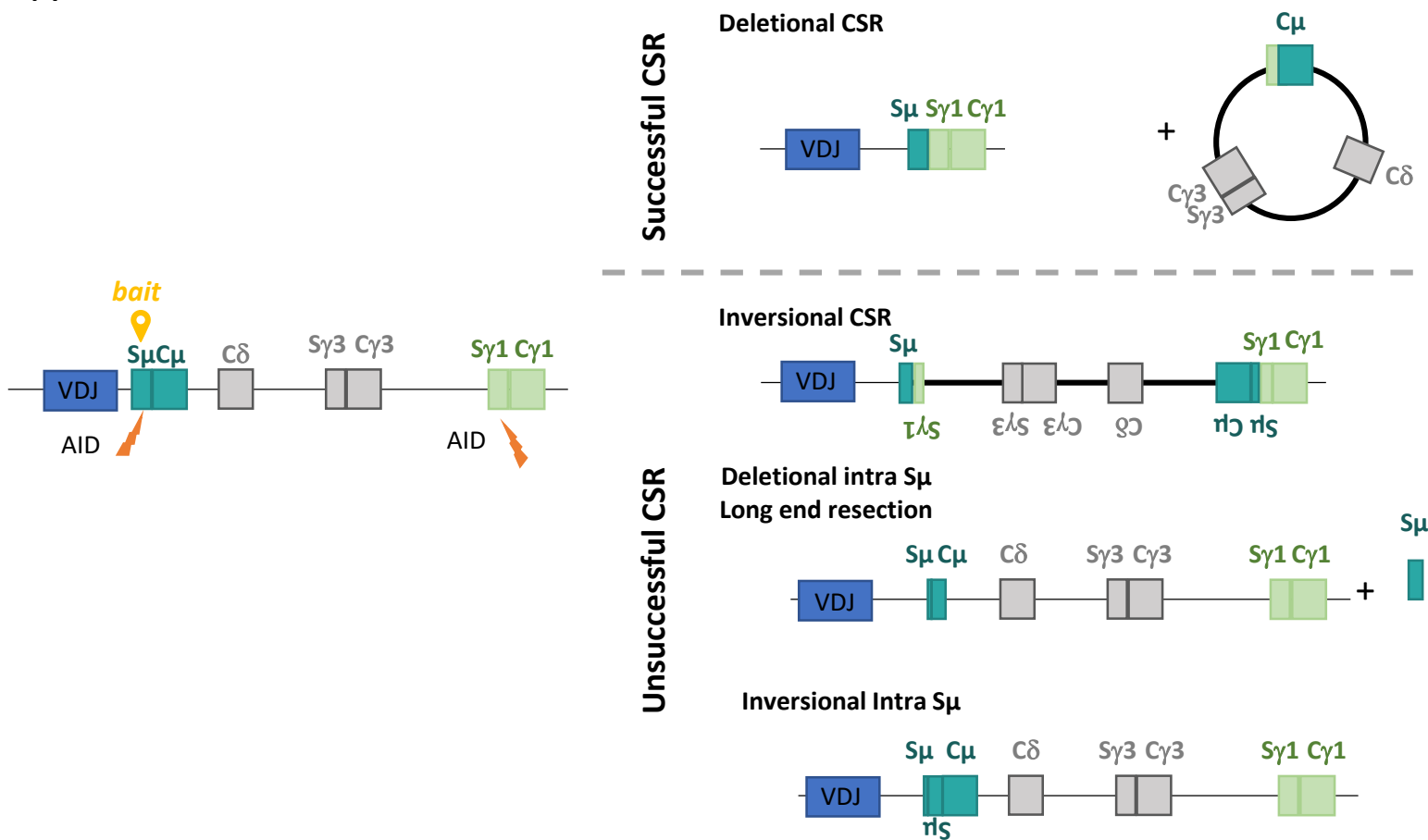
$\Delta 3'RR$  █

$C3'RR$  █

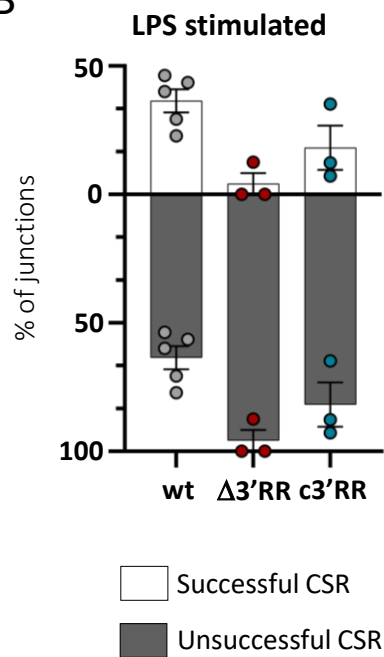




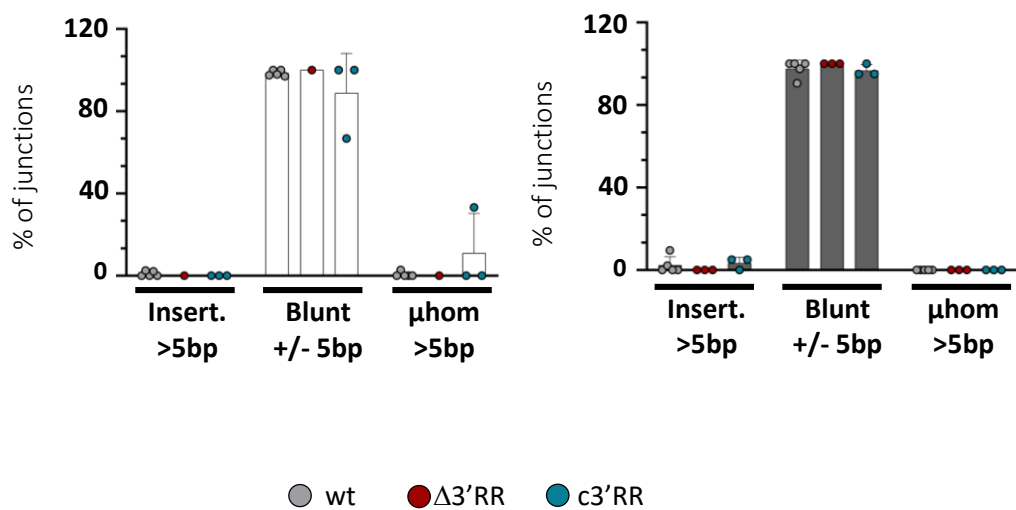
A



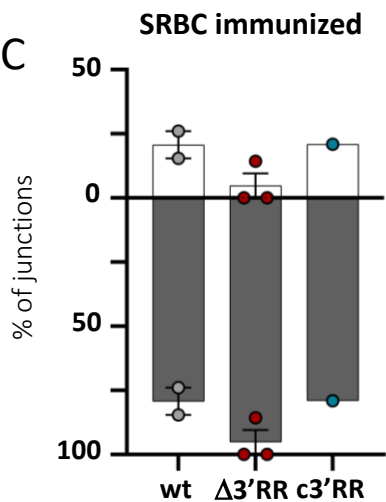
B



D



C



E

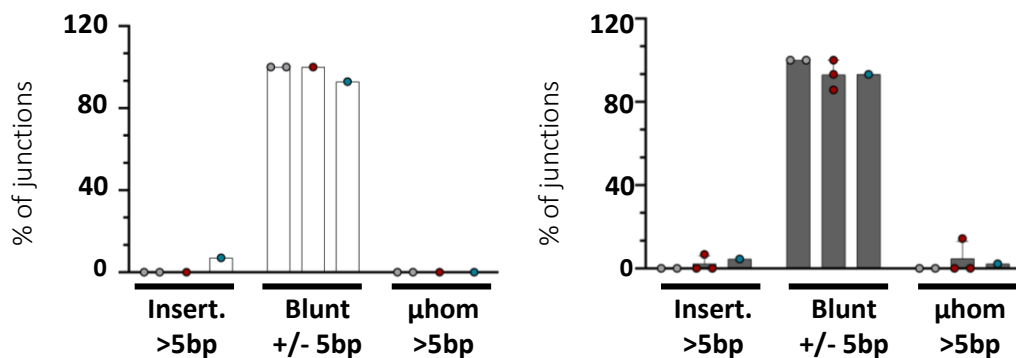
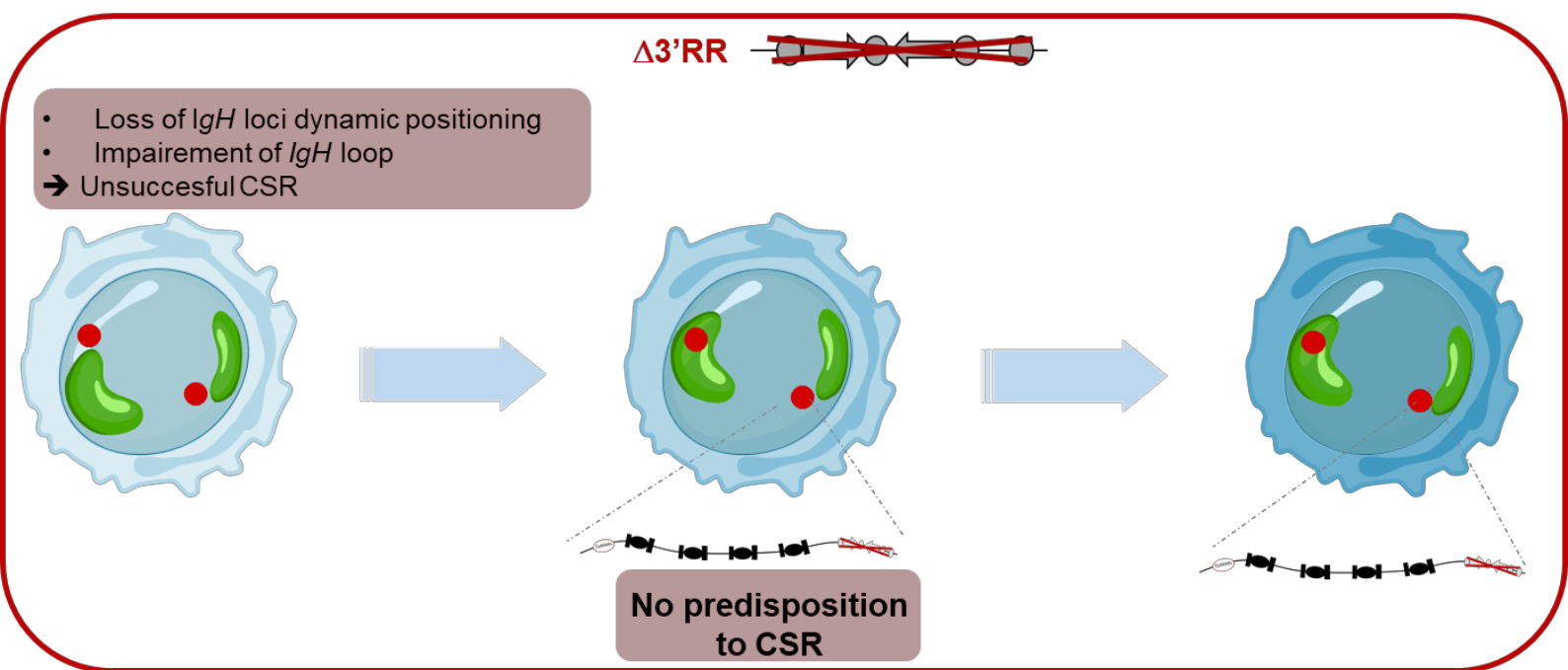
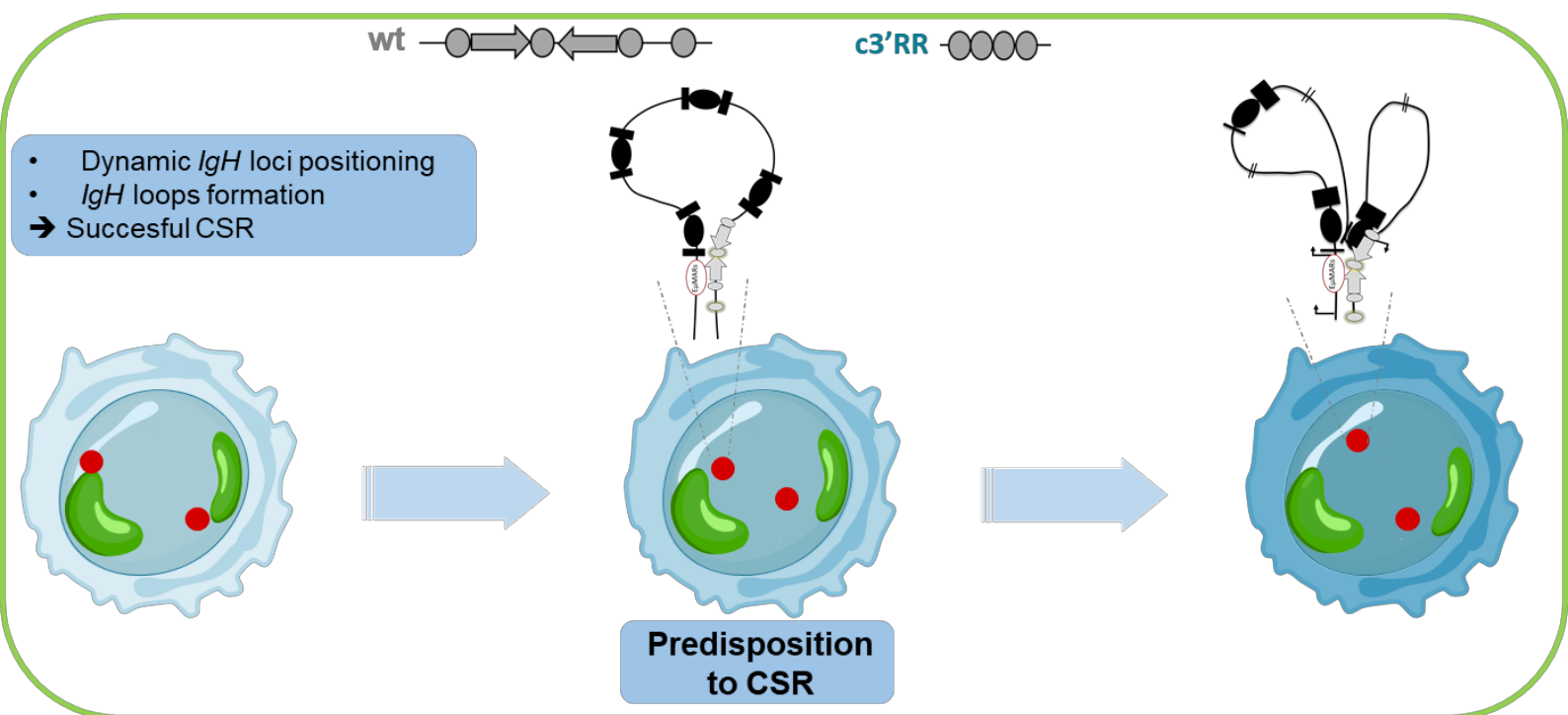


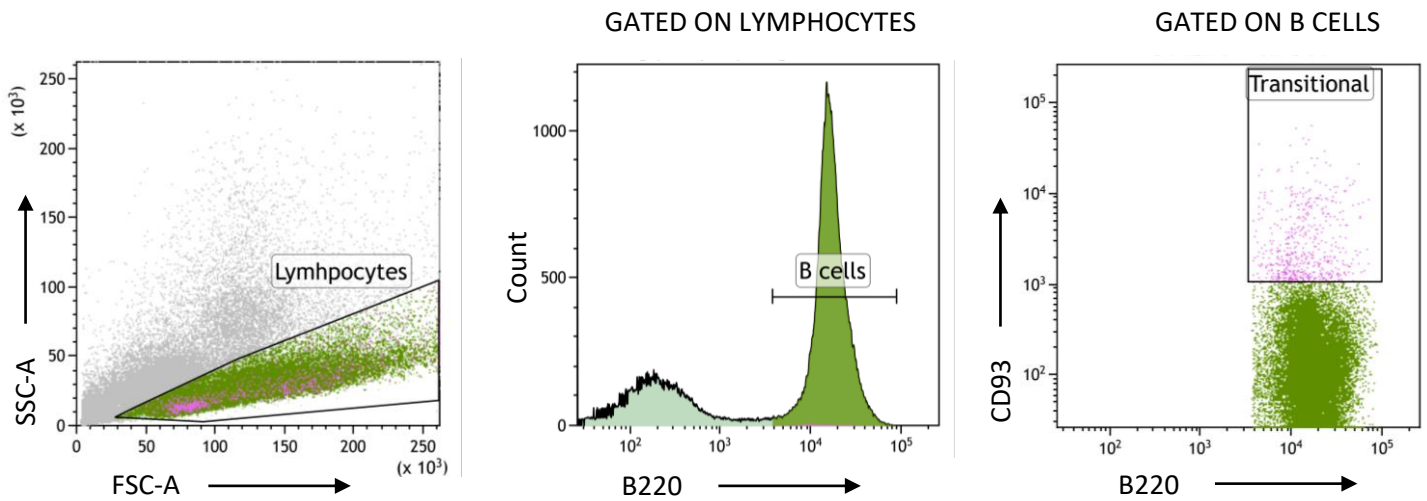
Figure 4



Graphical Abstract

**Figure S1**

**A**



**B**

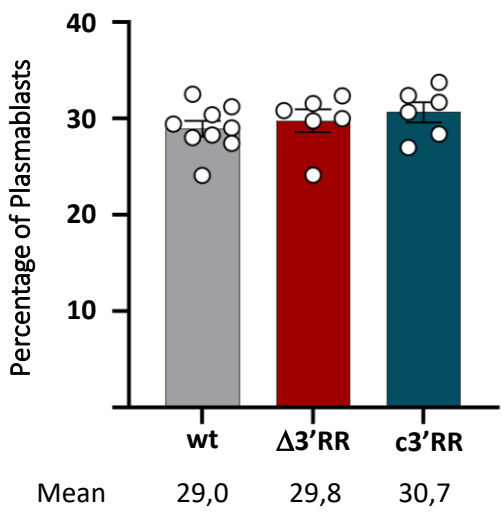
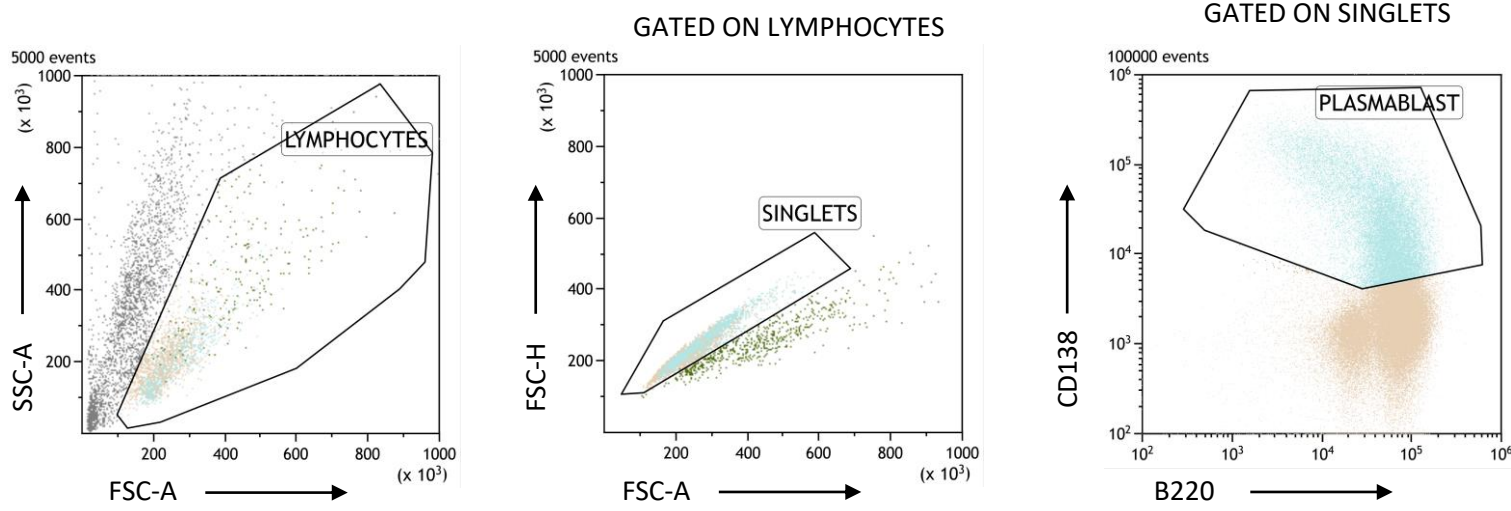
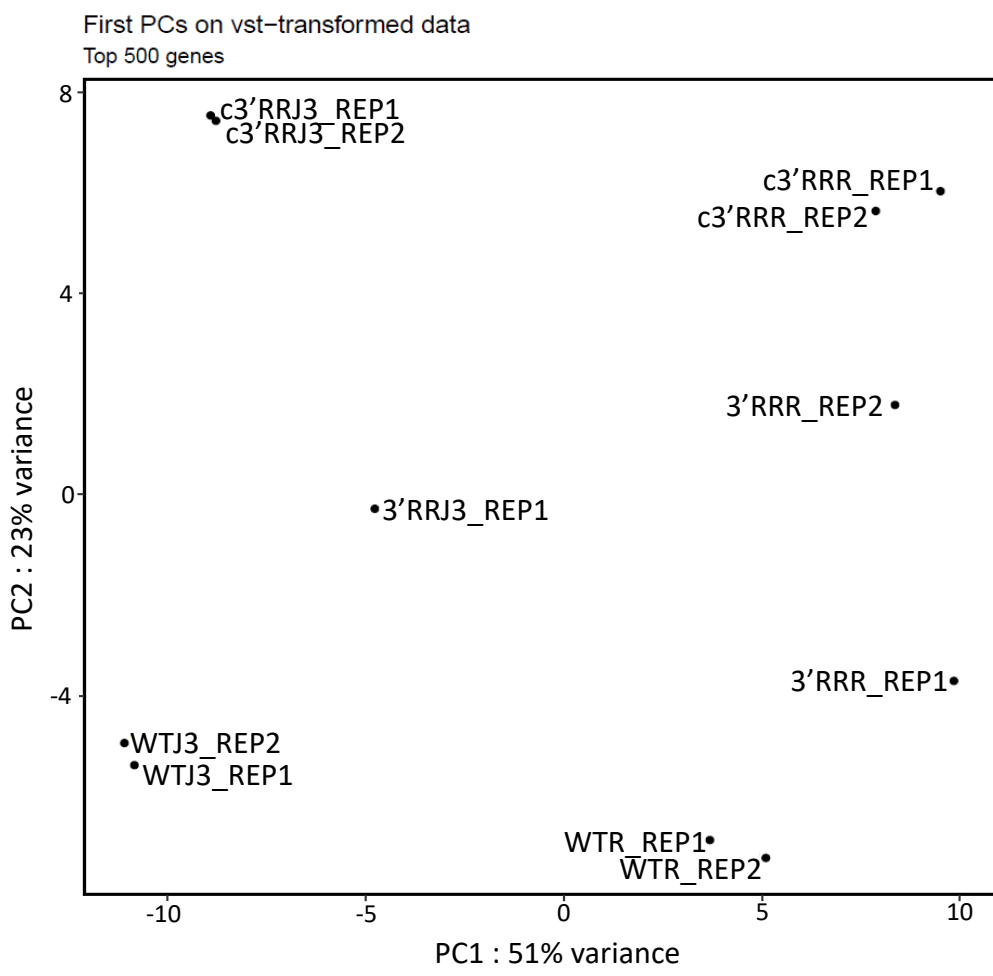


Figure S2



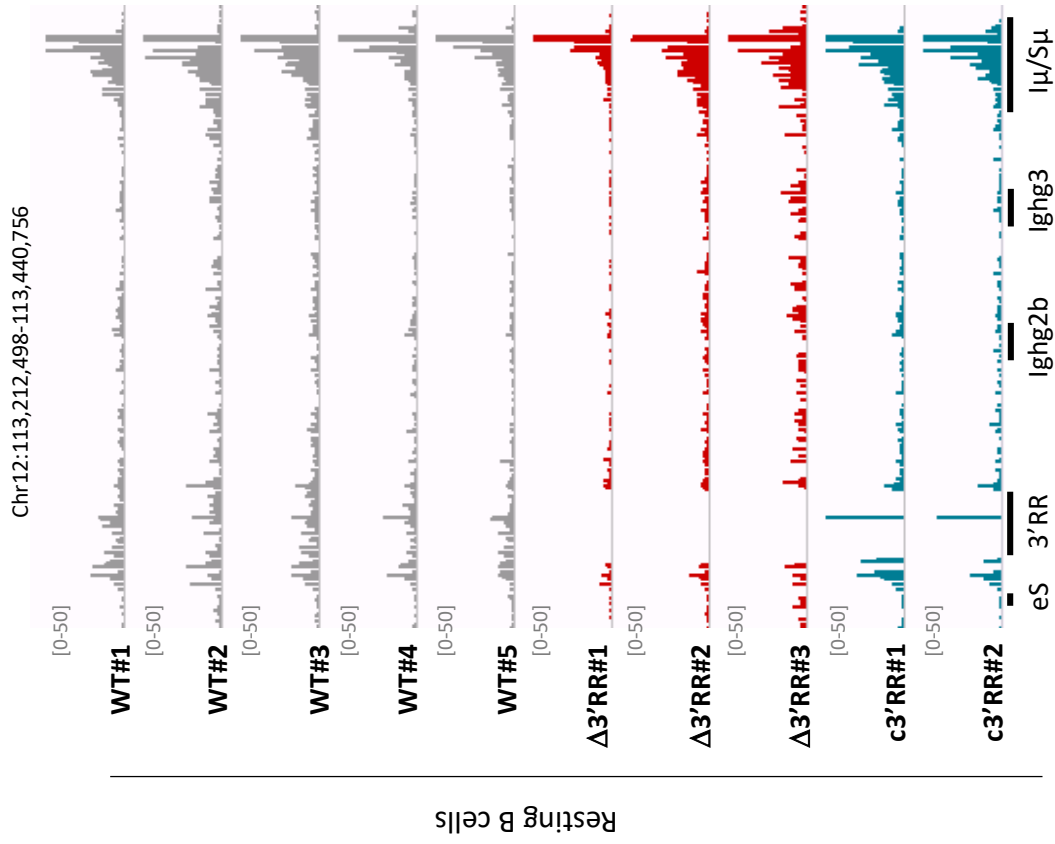
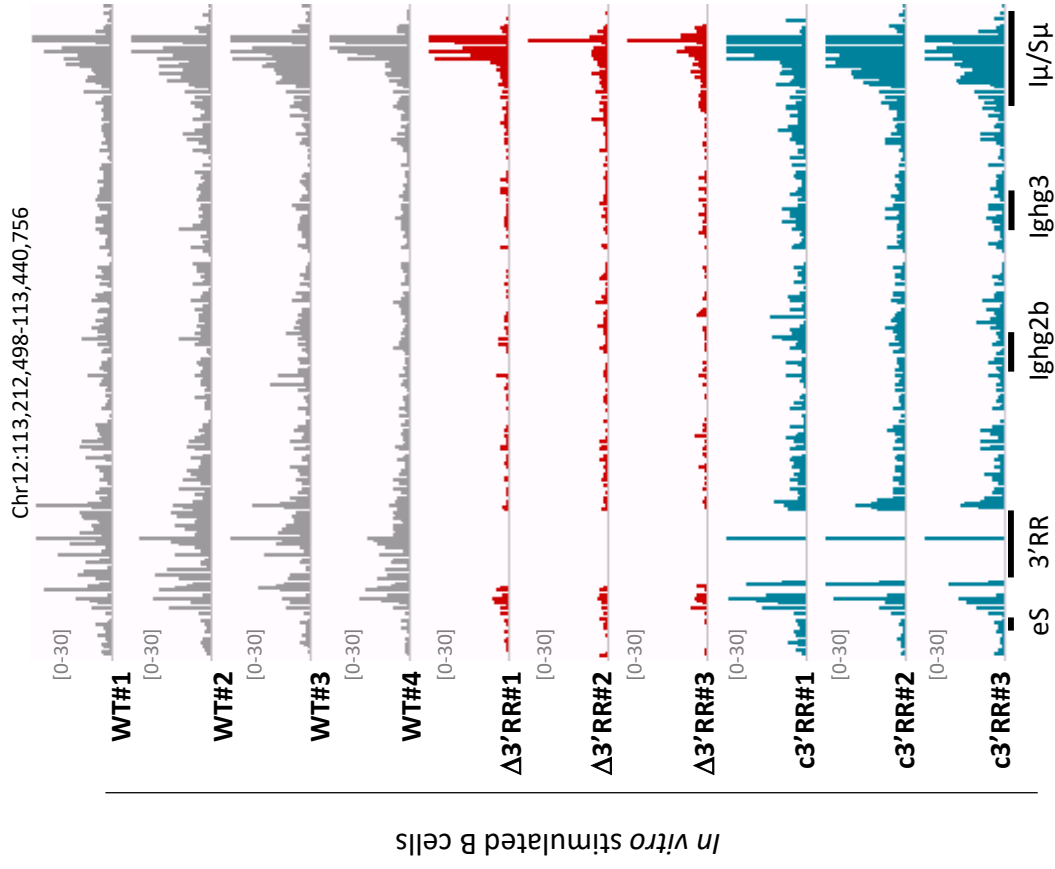
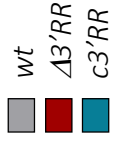
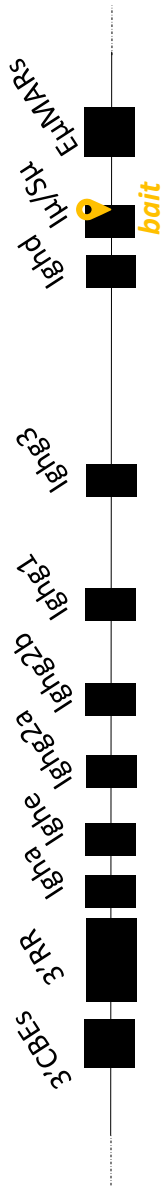
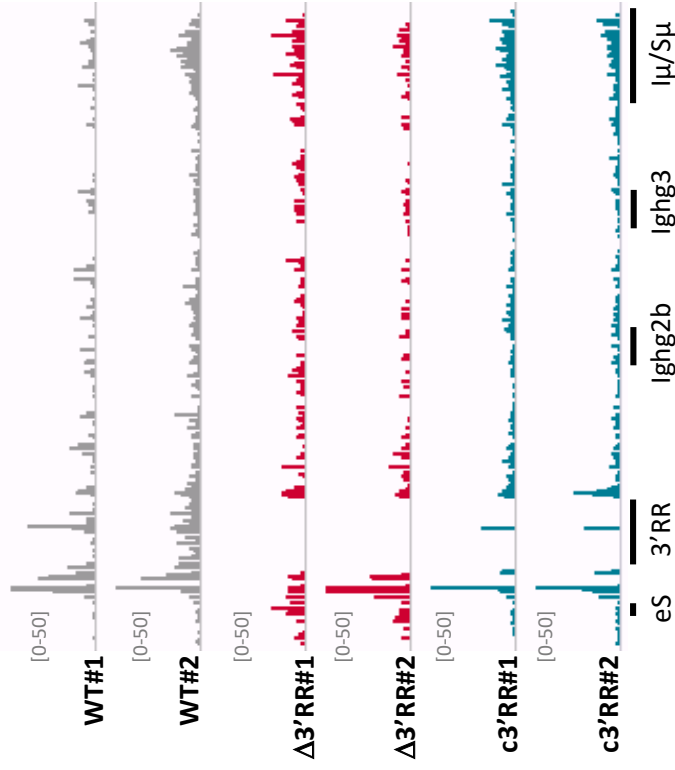


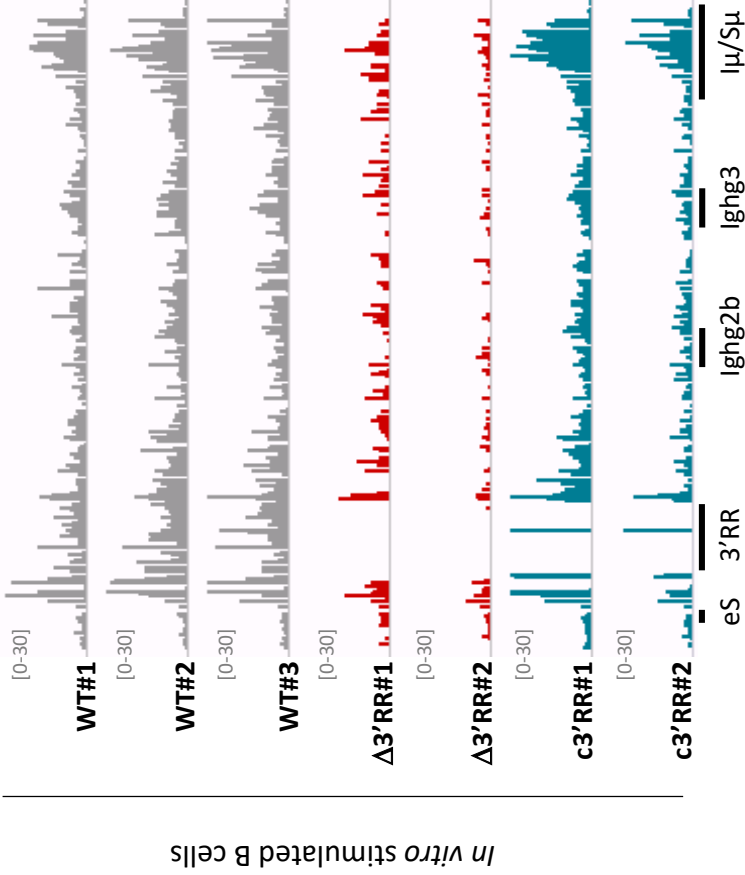
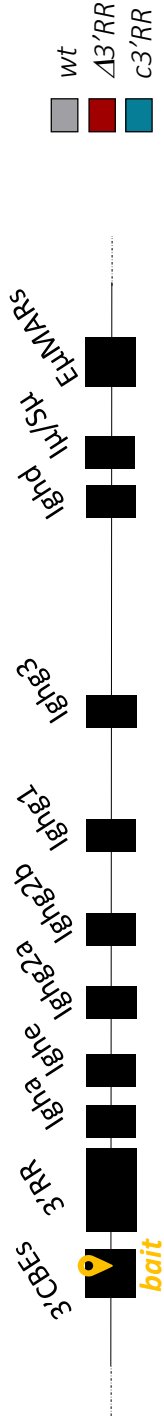
Figure S3

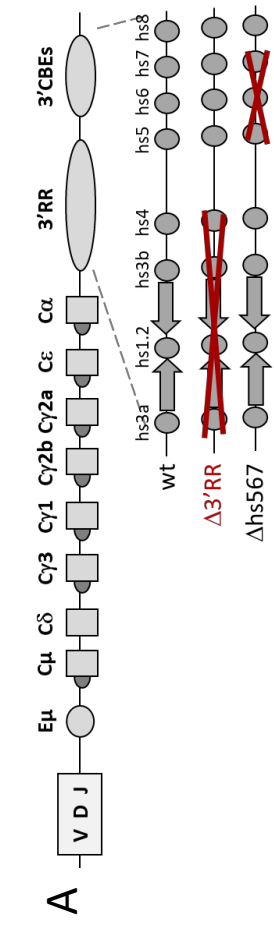
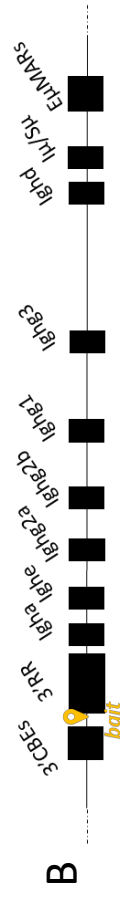
**B**

Chr12:113,212,498-113,440,756

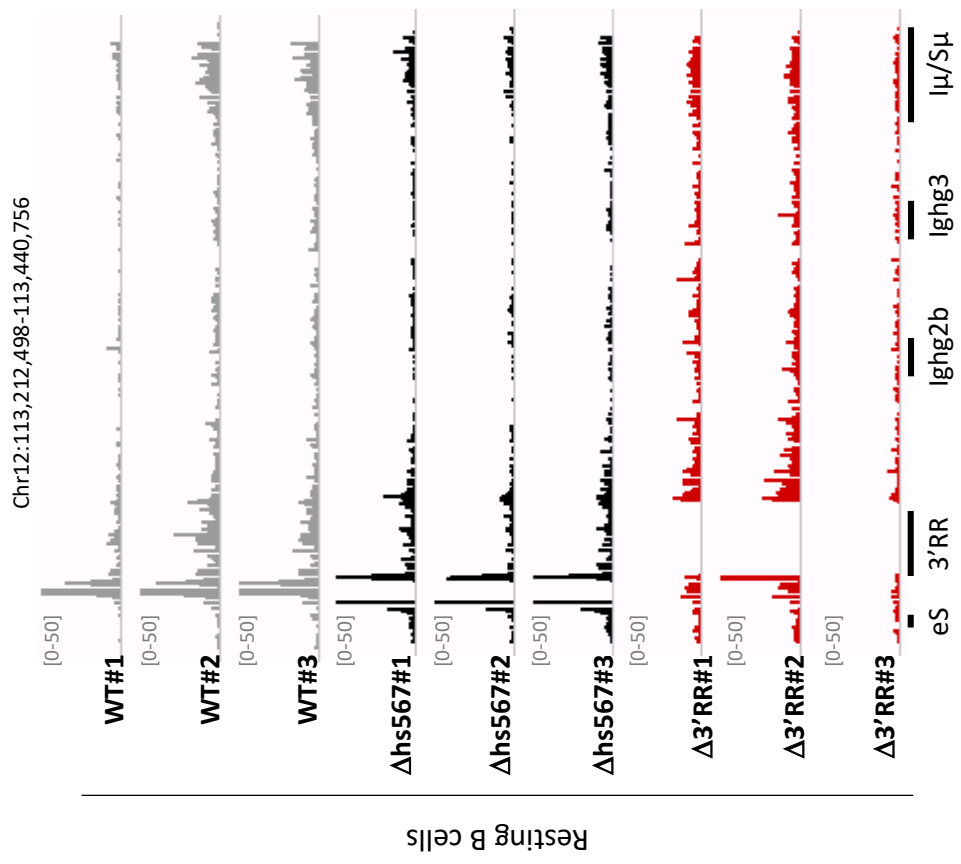
**C**

Chr12:113,212,498-113,440,756

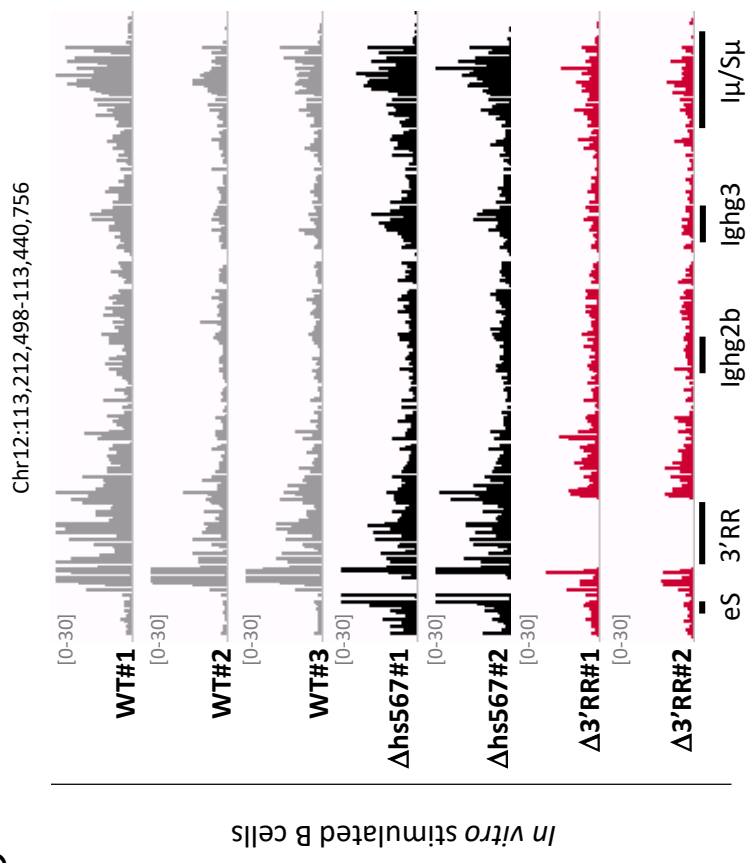
**A**



**C**



**D**



**Figure S5**

**Table S1 : Number of differentially opened regions by ATAC-Seq**

**A**

	Number of differentially opened regions
WT R vs WT S	687
c3'RR R vs c3'RR S	1292
3'RR R vs 3'RR S	384

**B**

	Number of Differentially opened regions	
	Resting	<i>In vitro</i> stimulated
WT vs c3'RR	45	73
WT vs $\Delta$ 3'RR	12	8
c3'RR vs $\Delta$ 3'RR	11	1



**Table S2 : Differentially opened regions, compared to wt , in resting and stimulated B cells from  $\Delta 3'RR$  and  $c3'RR$  mutant mice**

**A**

WT vs  $\Delta 3'RR$  in resting B cells

PeakID	baseMean	log2FoldChange	pvalue	Chr	Start	End	Strand	Annotation	Distance to TSS	Gene Name
Interval_12080	65,32400803	1,382533223	1,79903E-05	11	118218906	118219606	+	promoter-TSS (ENSMUSG00000017132)	-6514	Cyth1
Interval_12599	53,35748369	2,491795186	8,17632E-09	12	20920526	20921709	+	promoter-TSS (ENSMUSG00000085642)	145	Zfp125
Interval_12753	33,5991101	1,760311902	3,70619E-05	12	31973306	31973947	+	intron (ENSMUSG00000002997, intron 5 of 10)	-33219	Hbp1
Interval_15943	149,6198133	0,8875904	5,6026E-05	13	43764957	43765971	+	Intergenic	-19721	Cd83
Interval_30667	55,30642524	-1,652147387	2,11986E-05	19	21912947	21914145	+	Intergenic	-23834	Ldhb-ps
Interval_31323	44,29469979	-2,069512985	4,66191E-07	19	44351580	44352475	+	Intergenic	18935	Scd4
Interval_44362	27,05318006	2,289530272	3,88176E-05	5	23718896	23719554	+	Intergenic	-6559	AI506816
Interval_46728	14,33030532	2,908161161	6,42419E-05	5	122602245	122602675	+	intron (ENSMUSG00000029469, intron 6 of 6)	4909	Iff81
Interval_47330	53,93385657	-1,995496625	5,65951E-07	5	139780930	139782044	+	Intergenic	-5849	Ints1
Interval_47620	28,93190946	2,987157693	7,11953E-09	5	147049813	147050666	+	intron (ENSMUSG00000016520, intron 1 of 9)	-27325	Polr1d
Interval_56539	212,1635314	0,844766457	1,76651E-05	8	83739935	83742697	+	promoter-TSS (ENSMUSG00000002885)	10	Adgre5
Interval_59014	47,91307206	1,827564692	7,51789E-05	9	58510246	58511060	+	Intergenic	21691	Insyn1

**B**

WT vs  $c3'RR$  in resting B cells

PeakID	baseMean	log2FoldChange	pvalue	Chr	Start	End	Strand	Annotation	Distance to TSS	Gene Name
Interval_21965	32,17992055	3,865821078	1,29912E-10	15	79900715	79901440	+	TTS (ENSMUSG00000009585)	735	Apobec3
Interval_56539	212,1635314	1,18583272	7,49072E-10	8	83739935	83742697	+	promoter-TSS (ENSMUSG00000002885)	10	Adgre5
Interval_31323	44,29469979	-2,497613172	1,63031E-09	19	44351580	44352475	+	Intergenic	18935	Scd4
Interval_59014	47,91307206	2,688054574	1,66363E-09	9	58510246	58511060	+	Intergenic	21691	Insyn1
Interval_31829	42,87674911	2,497945969	1,99371E-08	19	61151514	61152486	+	Intergenic	-11183	Zfp950
Interval_26229	208,4196308	2,089885175	3,08772E-08	17	30611699	30612947	+	intron (ENSMUSG00000024026, intron 1 of 2)	246	Glo1
Interval_22720	21,47947807	4,10409096	1,99232E-07	15	100443572	100443968	+	Intergenic	-5914	n-R5s43
Interval_26244	38,10047802	2,071492998	1,14397E-06	17	30901814	30902460	+	intron (ENSMUSG00000024027, intron 1 of 13)	278	Glp1r
Interval_8283	30,33147907	-3,054353917	1,64249E-06	11	6497445	6498063	+	Intergenic	-21837	Purb
Interval_150	26,71177382	2,655475369	2,6277E-06	1	13593256	13593668	+	Intergenic	-3587	Tram1
Interval_26649	137,2364225	1,04371443	3,53146E-06	17	36110320	36111975	+	promoter-TSS (ENSMUSG000000092365)	-279	BC023719
Interval_8250	13,00628344	-7,12937049	4,306E-06	11	6006357	6006741	+	intron (ENSMUSG00000057897, intron 3 of 10)	-22411	Camk2b
Interval_36363	22,79411746	-2,390389571	4,99884E-06	2	173246679	173247222	+	intron (ENSMUSG00000038400, intron 1 of 3)	-28053	Zbp1
Interval_21963	110,4203818	1,252723517	5,92411E-06	15	79891525	79893135	+	promoter-TSS (ENSMUSG00000009585)	2798	AC113595.1
Interval_12599	53,35748369	1,940154387	8,37645E-06	12	20920526	20921709	+	promoter-TSS (ENSMUSG00000085642)	145	Zfp125

Interval_23703	28,86762533	2,160736956	1,1068E-05	16	31066693	31067108	+	intron (ENSMUSG000000047434, intron 1 of 3)	2926 Xyylt1
Interval_26645	96,3291297	1,123886763	1,31793E-05	17	36082985	36084629	+	exon (ENSMUSG000000073406, exon 2 of 6)	416 H2-BI
Interval_22557	63,60686106	1,481887312	1,39924E-05	15	97458206	97459223	+	Intergenic	97512 Pced1b
Interval_3614	11,9978808	6,753681783	1,75376E-05	1	171065212	171065430	+	promoter-TSS (ENSMUSG000000059498)	-669 Mir6546
Interval_49487	37,1492157	-2,247784038	3,50483E-05	6	87457375	87457932	+	TTS (ENSMUSG000000030047)	5994 Arhgap25
Interval_60872	120,6465451	-3,752072151	3,59424E-05	9	123936024	123937162	+	Intergenic	32099 Ccr1
Interval_36364	36,85784676	-1,867822973	5,78303E-05	2	173263590	173264142	+	intron (ENSMUSG000000038400, intron 1 of 3)	12323 Pmepa1
Interval_11216	97,55991131	1,008032889	6,01264E-05	11	99162626	99163659	+	Intergenic	-8065 Ccr7
Interval_57274	25,84206401	1,916226576	6,08076E-05	8	114801723	114802128	+	intron (ENSMUSG000000004637, intron 8 of 8)	89758 Wwox
Interval_21740	187,9637624	-0,804022157	6,61994E-05	15	76195518	76197057	+	promoter-TSS (ENSMUSG000000022565)	10034 Plec
Interval_3570	180,1793892	0,858393433	7,82664E-05	1	170174174	170175521	+	promoter-TSS (ENSMUSG000000026670)	-81 Uap1
Interval_3802	13,31646395	3,454270326	8,95444E-05	1	173905654	173906189	+	intron (ENSMUSG000000026536, intron 3 of 6)	7125 Ifi211
Interval_21964	51,86372983	4,339140464	9,9423E-05	15	79894946	79896164	+	exon (ENSMUSG00000009585, exon 2 of 8)	-4787 Apobec3
Interval_26299	51,59054898	1,459141889	0,000106059	17	31840333	31840838	+	Intergenic	10990 Siki1
Interval_34659	63,48790214	1,286894861	0,000106301	2	119015023	119016216	+	Intergenic	7410 Uqcrh-ps2
Interval_47825	71,08511526	-1,248560772	0,000106463	6	3399079	3399916	+	promoter-TSS (ENSMUSG000000047735)	-14491 Samd9l
Interval_55587	22,88391855	2,573328868	0,000116775	8	35123345	35123955	+	Intergenic	9882 Rpl31-ps23
Interval_31117	77,06605739	1,142616199	0,000120654	19	38025756	38027125	+	intron (ENSMUSG000000048612, intron 1 of 9)	16959 Myof
Interval_60871	127,8743705	-2,985733378	0,000124221	9	123930334	123930943	+	Intergenic	38054 Ccr1
Interval_3914	36,31498783	-1,593335152	0,000140264	1	177984119	177984575	+	intron (ENSMUSG000000091476, intron 1 of 19)	-7089 Catspere2
Interval_23653	14,59206643	2,911240379	0,000144414	16	30155064	30155415	+	Intergenic	-13294 n-R5s32
Interval_376	20,53100598	2,382695576	0,000178836	1	33883063	33883748	+	promoter-TSS (ENSMUSG000000042182)	75 Bend6
Interval_44362	27,05318006	2,064182212	0,000195559	5	23718896	23719554	+	Intergenic	-6559 Al506816
Interval_45629	165,5957232	0,750055988	0,000201062	5	86065126	86066144	+	promoter-TSS (ENSMUSG000000029253)	-84 Cenpc1
Interval_30607	54,45130071	-1,257058279	0,000212272	19	18619187	18620437	+	intron (ENSMUSG000000024725, intron 1 of 10)	11977 Ostf1
Interval_3816	51,29338043	-1,278227131	0,000229145	1	173977544	173978232	+	intron (ENSMUSG000000026535, intron 1 of 6)	4856 Ifi202b
Interval_1702	79,43648872	1,083395583	0,000231854	1	88213636	88214948	+	intron (ENSMUSG000000089960, intron 1 of 4)	2333 Ugt1a1
Interval_20339	188,585929	0,70617437	0,000249264	14	121937672	121939330	+	intron (ENSMUSG000000041765, intron 1 of 3)	-7813 Ubac2
Interval_19401	34,94976441	1,658884581	0,00026309	14	69823411	69823912	+	Intergenic	-16759 Pebp4
Interval_59140	36,41273826	1,56398437	0,000263996	9	61965345	61965995	+	intron (ENSMUSG000000032278, intron 2 of 4)	-40549 Kif23

## C

WT vs Δ3'RR in activated B cells

PeakID	baseMean	log2FoldChange	pvalue	Chr	Start	End	Strand	Annotation	Distance to TSS	Gene Name
Interval_42073	233,3062839	1,864429221	1,03407E-06	4	118547560	118548718	+	Intergenic	-4850	Tmem125
Interval_12753	33,5991101	2,150470647	5,34371E-05	12	31973306	31973947	+	intron (ENSMUSG000000002997, intron 5 of 10)	-33219	Hbp1
Interval_1702	79,43648872	1,396831824	6,974E-05	1	88213636	88214948	+	intron (ENSMUSG000000089960, intron 1 of 4)	2333	Ugt1a1

Interval_5472	132,8493816	-1,208150001	0,000105676	10	44380319	44382062	+	Intergenic	21514	Mir1929
Interval_47232	37,14445921	1,960240443	0,00012695	5	137596735	137597133	+	TTS (ENSMUSG00000037221)	3757	Mospd3
Interval_10679	149,4549722	-1,245289127	0,000140669	11	87442493	87444022	+	promoter-TSS (ENSMUSG00000098943)	20	Rnu3b1
Interval_59722	61,27660221	-1,751427473	0,000154888	9	80057522	80058632	+	Intergenic	-9740	Senp6
Interval_33995	187,3979067	-1,0141174	0,000202174	2	90579301	90581549	+	exon (ENSMUSG00000025314, exon 1 of 24)	222	Ptprj

## D

WT vs c3'RR in activated B cells

PeakID	baseMean	log2FoldChange	pvalue	Chr	Start	End	Strand	Annotation	Distance to TSS	Gene Name
Interval_1702	79,43648872	1,574730358	9,6632E-09	1	88213636	88214948	+	intron (ENSMUSG00000089960, intron 1 of 4)	2333	Ugt1a1
Interval_26244	38,10047802	2,209881998	1,96598E-08	17	30901814	30902460	+	intron (ENSMUSG00000024027, intron 1 of 13)	278	Glp1r
Interval_42073	233,3062839	1,70172397	4,17253E-08	4	118547560	118548718	+	Intergenic	-4850	Tmem125
Interval_21965	32,17992055	2,484747824	6,44651E-08	15	79900715	79901440	+	TTS (ENSMUSG00000009585)	735	Apobec3
Interval_35426	131,2968808	-1,507463092	9,62783E-08	2	148763686	148764765	+	Intergenic	7272	Cst11
Interval_31829	42,87674911	2,155830932	1,16867E-07	19	61151514	61152486	+	Intergenic	-11183	Zfp950
Interval_44094	45,62720152	2,071611603	5,10844E-07	5	14953946	14954607	+	Intergenic	-15847	Speer4e
Interval_44183	32,48779374	2,400333736	6,26839E-07	5	15674456	15675023	+	Intergenic	-5971	Speer4cos
Interval_22720	21,47947807	3,156788772	1,29761E-06	15	100443572	100443968	+	Intergenic	-5914	n-R5s43
Interval_31323	44,29469979	-1,768033742	3,19341E-06	19	44351580	44352475	+	Intergenic	18935	Scd4
Interval_19401	34,94976441	1,869601741	3,68295E-06	14	69823411	69823912	+	Intergenic	-16759	Pebp4
Interval_3816	51,29338043	-1,983993196	3,95923E-06	1	173977544	173978232	+	intron (ENSMUSG00000026535, intron 1 of 6)	4856	Ifi202b
Interval_44187	132,6956757	1,818150838	5,11197E-06	5	15688990	15691064	+	intron (ENSMUSG00000089871, intron 1 of 4)	9317	Speer4cos
Interval_44164	30,06497591	2,159268128	6,03011E-06	5	15612559	15613311	+	intron (ENSMUSG00000094230, intron 2 of 2)	-7227	Speer4d
Interval_44093	27,51772661	2,145276739	8,19656E-06	5	14952075	14952520	+	Intergenic	-13868	Speer4e
Interval_3614	11,9978808	6,648824762	1,01262E-05	1	171065212	171065430	+	promoter-TSS (ENSMUSG00000059498)	-669	Mir6546
Interval_44188	22,32978783	2,395398899	1,16701E-05	5	15697839	15698169	+	intron (ENSMUSG00000089871, intron 1 of 4)	16267	Speer4c
Interval_44186	113,2661305	1,853039056	1,38283E-05	5	15680019	15681848	+	intron (ENSMUSG00000089871, intron 1 of 4)	223	Speer4cos
Interval_6230	124,1870108	-1,071706257	2,06779E-05	10	75222061	75223178	+	intron (ENSMUSG00000033444, intron 2 of 17)	-38348	Specc1
Interval_44184	40,29153996	1,730058639	2,76438E-05	5	15678270	15678776	+	Intergenic	-2187	Speer4cos
Interval_15282	105,7124407	-0,992545515	2,88585E-05	13	21734066	21735952	+	promoter-TSS (ENSMUSG00000067455)	-55	H4c11
Interval_22866	364,4687526	-0,706606264	3,85548E-05	15	103239341	103241098	+	promoter-TSS (ENSMUSG00000046434)	-275	Hnrnpa1
Interval_60870	711,7367054	-2,705222962	4,33377E-05	9	123927808	123929581	+	Intergenic	39998	Ccr1
Interval_3611	55,66941374	1,519453318	4,50646E-05	1	171043984	171045076	+	Intergenic	14873	Fcgr3
Interval_47232	37,14445921	1,666946345	5,42141E-05	5	137596735	137597133	+	TTS (ENSMUSG00000037221)	3757	Mospd3
Interval_45797	131,6538692	-0,897779869	5,51586E-05	5	96221157	96223265	+	intron (ENSMUSG00000029486, intron 2 of 6)	12067	Mrpl1
Interval_15347	82,63239193	-1,068115974	8,69491E-05	13	23534826	23535728	+	promoter-TSS (ENSMUSG00000099517)	-145	H3c8
Interval_30464	317,3291162	-0,660957985	8,88432E-05	19	11964905	11966219	+	promoter-TSS (ENSMUSG00000024687)	-379	Osbp

Interval_26224	26,54437711	1,982402618	9,46021E-05	17	30518826	30519215	+	intron (ENSMUSG000000062202, intron 1 of 3)	5883 Btbd9
Interval_34114	142,9101241	-0,882283421	0,000107796	2	93427054	93428636	+	intron (ENSMUSG000000027215, intron 6 of 6)	-5838 Mir7001
Interval_44302	13,78885203	-2,562278039	0,000108172	5	21558544	21559035	+	intron (ENSMUSG000000048520, intron 9 of 18)	15230 Lrrc17
Interval_29172	243,8779389	-0,779322032	0,000115394	18	60803129	60804071	+	promoter-TSS (ENSMUSG000000024610)	-305 Cd74
Interval_21963	110,4203818	1,001117949	0,000142526	15	79891525	79893135	+	promoter-TSS (ENSMUSG000000009585)	2798 AC113595.1
Interval_21964	51,86372983	4,082461674	0,000153837	15	79894946	79896164	+	exon (ENSMUSG000000009585, exon 2 of 8)	-4787 Apobec3
Interval_1712	25,20582771	1,918411199	0,00017025	1	88300647	88300934	+	intron (ENSMUSG000000036251, intron 1 of 5)	-18123 Trpm8
Interval_14878	130,5275533	-0,971499777	0,000170437	13	3525899	3526891	+	Intergenic	-27644 Gdi2
Interval_15284	161,9871444	-0,782875175	0,000172515	13	21753266	21754892	+	promoter-TSS (ENSMUSG000000063021)	-44 H2bc15
Interval_26454	144,4452703	-0,909899924	0,000189353	17	34257405	34258215	+	exon (ENSMUSG000000073421, exon 1 of 3)	-5440 H2-Ab1
Interval_40863	103,4302046	-1,078185572	0,000189918	4	53277689	53278635	+	Intergenic	-12670 AI427809
Interval_41725	47,00441879	-1,382843908	0,000190004	4	105263481	105264243	+	Intergenic	106515 Plpp3
Interval_50529	109,2901486	-1,061151687	0,000206901	6	125572473	125573699	+	intron (ENSMUSG00000001930, intron 6 of 14)	20123 Vwf
Interval_47180	64,90966257	-1,140198511	0,000217993	5	136640538	136641571	+	Intergenic	-56725 Myl10
Interval_44168	34,46301788	1,728339739	0,000228464	5	15632164	15632812	+	intron (ENSMUSG000000094230, intron 2 of 2)	12326 Speer4d
Interval_40012	21,56651958	-1,879694059	0,00023588	4	9640702	9641342	+	intron (ENSMUSG000000028207, intron 1 of 4)	2695 Asph
Interval_3770	24,14270976	1,789781071	0,000245681	1	173490800	173491426	+	promoter-TSS (ENSMUSG000000037849)	-17187 Ifi206
Interval_55548	45,58801231	1,545998902	0,000250773	8	33930207	33931049	+	promoter-TSS (ENSMUSG000000031586)	-1236 Rbpms
Interval_26245	11,73009879	2,714798357	0,00025438	17	30907467	30907703	+	intron (ENSMUSG000000024027, intron 1 of 13)	5726 Glp1r
Interval_51254	323,8723748	-0,69664298	0,000270884	7	3644569	3645690	+	promoter-TSS (ENSMUSG000000035632)	-139 Cnot3
Interval_5472	132,8493816	-0,813153144	0,000277818	10	44380319	44382062	+	Intergenic	21514 Mir1929
Interval_44088	30,47068406	1,776983366	0,000293044	5	14940262	14940887	+	Intergenic	-2145 Speer4e
Interval_1704	21,35051159	2,094514512	0,000297615	1	88222931	88223288	+	Intergenic	-3877 Mroh2a
Interval_1711	25,49836405	1,767036919	0,00030777	1	88292041	88292380	+	intron (ENSMUSG000000036251, intron 1 of 5)	-21263 Hjurp
Interval_6364	188,3314937	-0,748373791	0,000311505	10	78086406	78087696	+	Intergenic	13712 Icosl
Interval_6903	13,08088675	2,870172895	0,00032234	10	93198355	93198727	+	intron (ENSMUSG000000020015, intron 1 of 2)	35756 Mir1931
Interval_8250	13,00628344	-3,595617214	0,00034183	11	6006357	6006741	+	intron (ENSMUSG000000057897, intron 3 of 10)	-22411 Camk2b
Interval_44167	23,28036799	1,865273345	0,000341887	5	15624001	15625093	+	TTS (ENSMUSG000000070933)	4385 Speer4d
Interval_32299	216,176694	-0,709439392	0,000346953	2	22773296	22775070	+	promoter-TSS (ENSMUSG000000026786)	-207 Appb1ip
Interval_44181	18,0304766	2,268498764	0,000356485	5	15670992	15671420	+	Intergenic	-9504 Speer4cos
Interval_22557	63,60686106	1,253044281	0,000358304	15	97458206	97459223	+	Intergenic	97512 Pced1b
Interval_3919	273,6291358	-0,659543071	0,000390299	1	178186997	178188050	+	promoter-TSS (ENSMUSG000000026502)	106 Desi2
Interval_15370	295,3522521	-0,628231787	0,000394584	13	23760742	23762873	+	promoter-TSS (ENSMUSG000000060093)	-577 H4c1
Interval_15349	95,13161521	-0,923886173	0,000408953	13	23550775	23552006	+	exon (ENSMUSG000000069274, exon 1 of 1)	258 H4c6
Interval_376	20,53100598	2,050888202	0,000449203	1	33883063	33883748	+	promoter-TSS (ENSMUSG000000042182)	75 Bend6
Interval_33989	71,50294514	-1,153223244	0,000461867	2	90547072	90547800	+	intron (ENSMUSG000000025314, intron 1 of 23)	33211 Ptprj
Interval_8579	88,94296747	-0,972741454	0,000484739	11	22935942	22937170	+	intron (ENSMUSG000000051355, intron 1 of 1)	36303 Comm1d
Interval_26229	208,4196308	1,309514607	0,000484765	17	30611699	30612947	+	intron (ENSMUSG000000024026, intron 1 of 2)	246 Glo1

Interval_8283	30,33147907	-1,831873649	0,000490909	11	6497445	6498063	+	Intergenic	-21837 Purb
Interval_37946	97,07367932	-0,937381159	0,000491146	3	87401074	87402162	+	intron (ENSMUSG000000059994, intron 10 of 10)	25158 Fcrl1
Interval_15313	139,2061391	-0,765686035	0,000506773	13	22035040	22036939	+	promoter-TSS (ENSMUSG000000069302)	119 H2bc12
Interval_57167	51,72714388	-1,299461372	0,000515874	8	109750497	109751074	+	Intergenic	-18672 Atxn1l
Interval_51462	210,1594425	-0,682504995	0,000530183	7	12977554	12978703	+	promoter-TSS (ENSMUSG000000033961)	235 Zfp446
Interval_59918	140,8688252	-0,770390845	0,000541605	9	90059740	90061248	+	exon (ENSMUSG000000032359, exon 2 of 12)	5940 Ctsh
Interval_23609	40,12752824	-1,308519993	0,000567463	16	27399071	27400151	+	intron (ENSMUSG000000038127, intron 1 of 8)	10742 Ccdc50

**Table S3 : Coordinates used for quantification of normalized coverage in R1, R2 and R3 regions**

	chr12	
R1	113 216 000	113 225 000
R2	113 255 000	113 407 000
R3	113 407 000	113 430 000

**Table S4 : Normalized reads count of 3C-HTGTS performed with  $I\mu/S\mu$  bait in region of interest within the *IgH* locus in resting and stimulated B cells from *wt* and mutants mice**

**A**

	chr12	
<b>S<math>\mu</math></b>	113 385 000	113 430 000
<b>S<math>\gamma</math>3</b>	113 355 000	113 370 000
<b>S<math>\gamma</math>2b</b>	113 315 000	113 300 000
<b>3'RR</b>	113 255 000	113 225 001
<b>eS</b>	113 215 000	113 209 000
<b>Whole IgH</b>	113 209 000	133 430 000

**B**

	Resting B cells									
	WT#1	WT#2	WT#3	WT#4	WT#5	$\Delta$ 3RR#1	$\Delta$ 3RR#2	$\Delta$ 3RR#3	c3'RR#1	c3'RR#2
<b>S<math>\mu</math></b>	30	58	56	28	43	18	53	41	47	34
<b>3'RR</b>	767	704	532	681	414	46	24	14	207	125
<b>eS</b>	33	15	26	38	18	22	30	12	23	24
<b>Whole IgH</b>	5066	5157	5176	5401	5359	4561	3522	4035	4479	5315

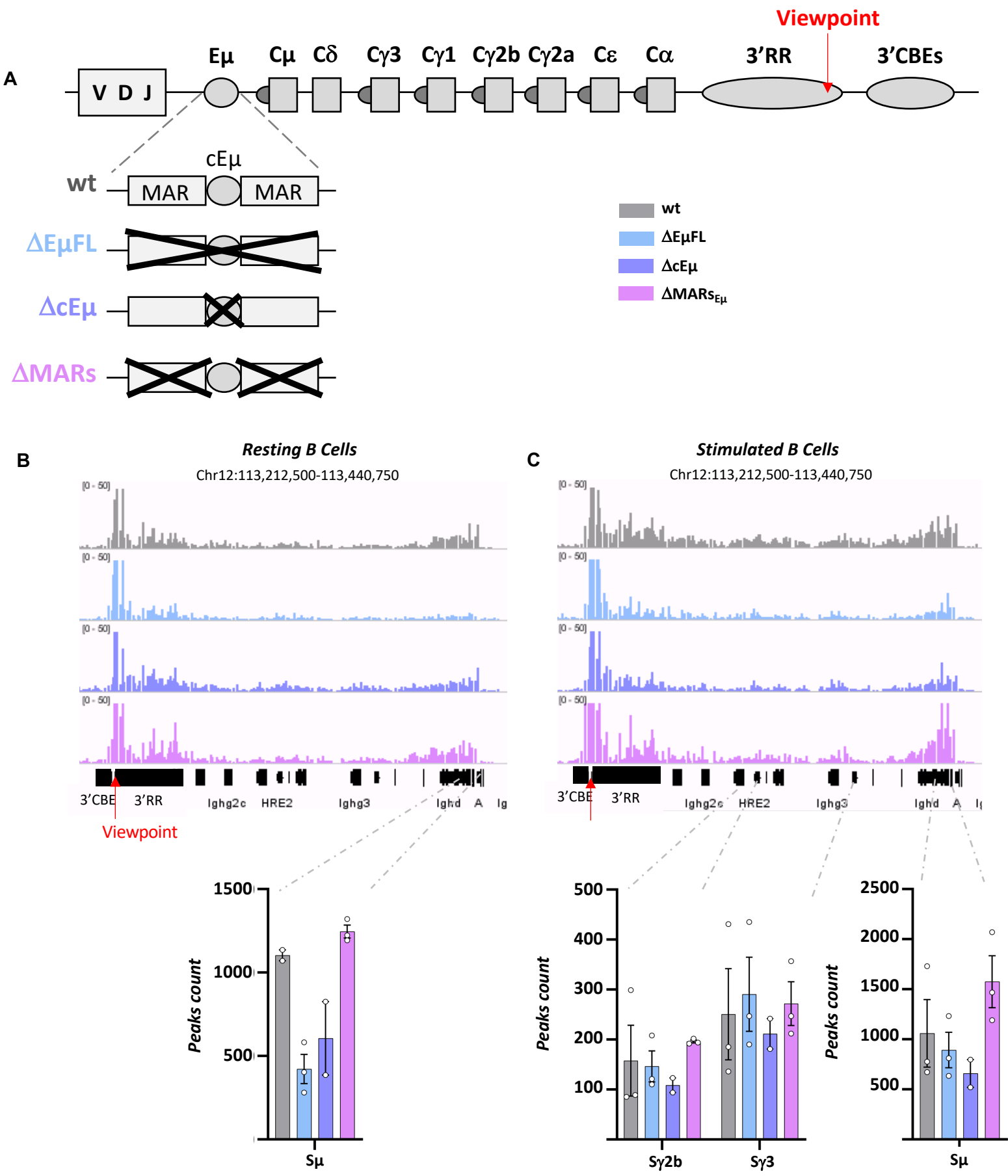
**C**

	<i>In vitro</i> stimulated B cells									
	WT#1	WT#2	WT#3	WT#4	$\Delta$ 3RR#1	$\Delta$ 3RR#2	$\Delta$ 3RR#3	c3'RR#1	c3'RR#2	c3'RR#3
<b>S<math>\mu</math></b>	82	68	50	40	15	11	5	82	47	61
<b>S<math>\gamma</math>3</b>	1228	1247	1067	881	114	113	50	501	484	538
<b>S<math>\gamma</math>2b</b>	169	178	129	94	10	25	15	74	135	102
<b>3'RR</b>	717	891	1037	1051	14	6	15	251	244	182
<b>eS</b>	20	34	31	54	7	2	6	27	36	55
<b>Whole IgH</b>	6057	5941	5987	5177	3870	4508	4771	6252	5951	4168

**Table S5 : Coordinates used for LAM junctions analysis**

	chr12	
<b>S<math>\mu</math></b>	113 403 753	113 426 678
<b>S<math>\gamma</math>3</b>	113 356 233	113 366 412
<b>S<math>\gamma</math>1</b>	113 326 667	113 339 416
<b>S<math>\gamma</math>2b</b>	113 304 074	113 314 828
<b>S<math>\gamma</math>2a</b>	113 284 212	113 296 326
<b>S<math>\epsilon</math></b>	113 269 177	113 277 804
<b>S<math>\alpha</math></b>	113 256 071	113 265 138





**Figure 26 : Involvement of  $E\mu$ MARs region in loop conformation of the  $IgH$  locus,**

**A.** Murine  $IgH$  locus is represented (not to scale).  $Ig$  gene segments are indicated. Deletion of  $E\mu$ MARs is represented for each model used in this study. 3C-HTGTS viewpoint is represented by the red arrow. **B.** 3C-HTGTS bedgraphs from Resting B cells visualized with IGV tool. **C.** Same as B in *in vitro* LPS stimulated B cells. **D.** Quantification of normalized reads count in  $S\mu$ ,  $S\gamma3$  and  $S\gamma2b$  region in resting and stimulated B cells. Error bars represent SEM.

## I.II. Résultats préliminaires : Rôle des régions $MARs_{E\mu}$ dans la formation de boucles de chromatine.

Il est démontré, depuis maintenant de nombreuses années, qu'une boucle de chromatine mettant en interaction les deux éléments *cis*-régulateurs majeurs du locus *IgH* ( $E\mu MARs$  et  $3'RR$ ) est présente dans les cellules B matures au repos (Wuerffel et al., 2007). Cette grande boucle forme une structure appelée « centre de CSR » (CSRC) qui constitue un prérequis essentiel à la CSR (X. Zhang et al., 2019). Lors de l'activation antigénique des cellules B, cette grande boucle persiste et des boucles internes apparaissent. Ces boucles internes mettent en contact les régions  $E\mu MARs$  et  $3'RR$  avec les régions switch donneuse et acceptrices (Wuerffel et al., 2007). Plus récemment, il a été montré que la formation de la grande boucle et des boucles internes est médiée par le complexe de la cohésine *via* le mécanisme d'extrusion de boucle (X. Zhang et al., 2019). Cependant, la régulation de la formation et de la dissociation de ces boucles n'est à ce jour pas complètement décrite.

Etant mises en contact lors de la formation de la grande boucle, nous émettons l'hypothèse que les régions  $E\mu MARs$  et  $3'RR$  participent à la régulation du mécanisme d'extrusion de boucle. Nous avons exploré la formation de la grande boucle dans des cellules B au repos et stimulées *in vitro* pendant 3 jours avec du LPS, issues de différents modèles murins présentant des délétions totales ou partielles de la région  $E\mu MARs$  (modèles  $\Delta E\mu FL$  (Marquet et al., 2014),  $\Delta cE\mu$  (Perlot et al., 2005),  $\Delta MARs_{E\mu}$  (Martin et al., 2023)). La configuration du locus a été évaluée par 3C-HTGTS avec un point de vue (=ancre) situé dans la  $3'RR$  (**Figure 26A**).

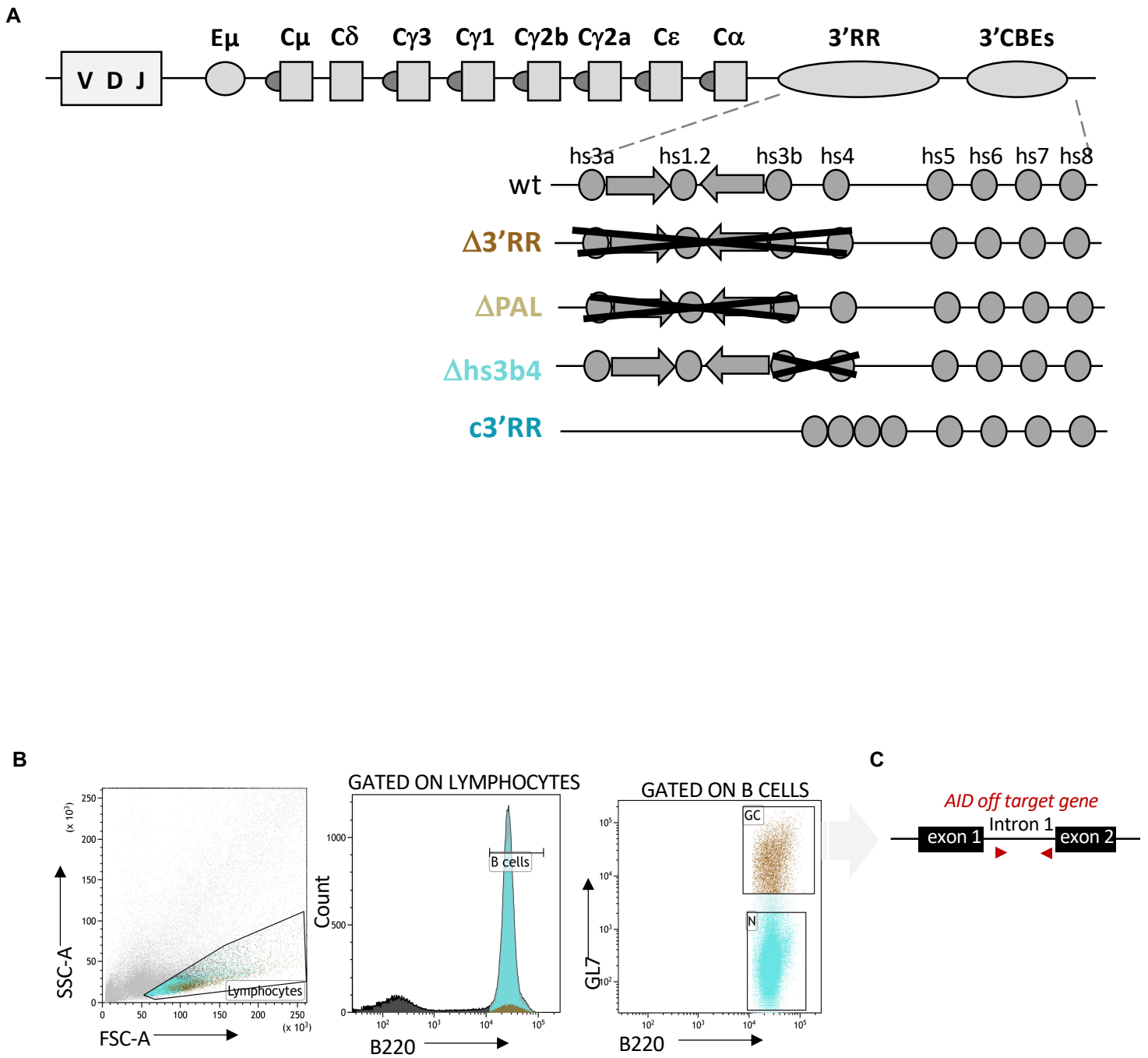
Dans les cellules B au repos *wt*, les interactions sont détectées essentiellement autour de l'ancre (traduisant un certain repliement de la région  $3'RR$ ) et à plus longue distance entre les régions  $3'RR$  et  $E\mu MARs/S\mu$  (révélant la présence de la grande boucle) (**Figure 26B**). De façon attendue, lorsque l'ensemble de la région  $E\mu MARs$  ( $\Delta E\mu FL$ ) est délétée, la grande boucle n'est quasiment plus détectable, seules demeurent les interactions autour de l'ancre  $3'RR$  (**Figure 26B**). Dans le modèle  $\Delta cE\mu$ , modèle dans lequel l'enhancer  $cE\mu$  est absent mais les deux  $MARs$  sont conservées, la fréquence d'interaction entre la  $3'RR$  et la région  $E\mu MARs$  semble également sensiblement diminuée (**Figure 26B**). A l'inverse, dans le modèle  $\Delta MARs$  la fréquence d'interaction entre la  $3'RR$  et  $S\mu$  apparaît sensiblement augmenté par rapport au contexte *wt* (**Figure 26B**). Ces données suggèrent l'implication de l'enhancer  $cE\mu$  dans l'établissement de la grande boucle dans les cellules B au repos. Les régions  $MARs$  flanquant ce même enhancer pourraient, quant à elles, être impliquées dans un mécanisme limitant la formation de cette grande boucle, puisque leur suppression provoque une augmentation de la fréquence d'interaction entre la  $3'RR$  et la région  $E\mu MARs/S\mu$ .

Alors que la grande boucle  $3'RR-E\mu MARs-S\mu$  persiste dans les cellules B *wt* stimulées *in vitro* ; une diminution globale des interactions est observée dans les modèles  $\Delta E\mu FL$  et  $\Delta cE\mu$ . A l'inverse, dans le modèle  $\Delta MARs$ , l'interaction entre la  $3'RR$  et  $S\mu$  persiste et se trouve renforcée (**Figure 26C**) ; cet effet

est observé de façon plus nette que dans les cellules B au repos. Nous supposons que cette dernière observation ne traduit pas nécessairement une augmentation du nombre absolu de boucles mais pourrait être la conséquence d'une stabilité accrue de ces boucles.

Des interactions supplémentaires entre la 3'RR et les régions switch cibles (*S $\gamma$ 3* et *S $\gamma$ 2b*) sont détectées en contexte *wt* (**Figure 26C**). Ces boucles de chromatine internes sont également détectées dans les différents modèles mutants, avec une fréquence similaire à celle observée en contexte *wt*. Ces observations sont en accord avec une recombinaison de CSR s'opérant normalement dans ces modèles (Martin et al., 2023 ; Perlot et al., 2005).

Ces expérimentations de 3C-HTGTS nécessitent d'être répétées afin d'obtenir des triplicats pour chaque modèle et type cellulaire. Nous prévoyons également de compléter ces données en utilisant une ancre complémentaire localisée dans la région *I $\mu$ /S $\mu$* . Les hypothèses mécanistiques relatives à ces données et leur mise en perspective sont discutées dans la partie **Discussion & Perspectives**.



**Figure 27 : Scheme of mouse models and gating strategy to sort GC B cells**

**A.** Murine IgH locus is represented (not to scale). Ig gene segments are indicated; grey arrows represent the palindromic inverted repeats. Deletion of 3'RR is represented for each model used in this study.  
**B.** Gating strategy for GC B cell sorting **C.** AID off target genes zone where mutations are analyzed, PCR primers for SHM analysis are indicated by red arrows.

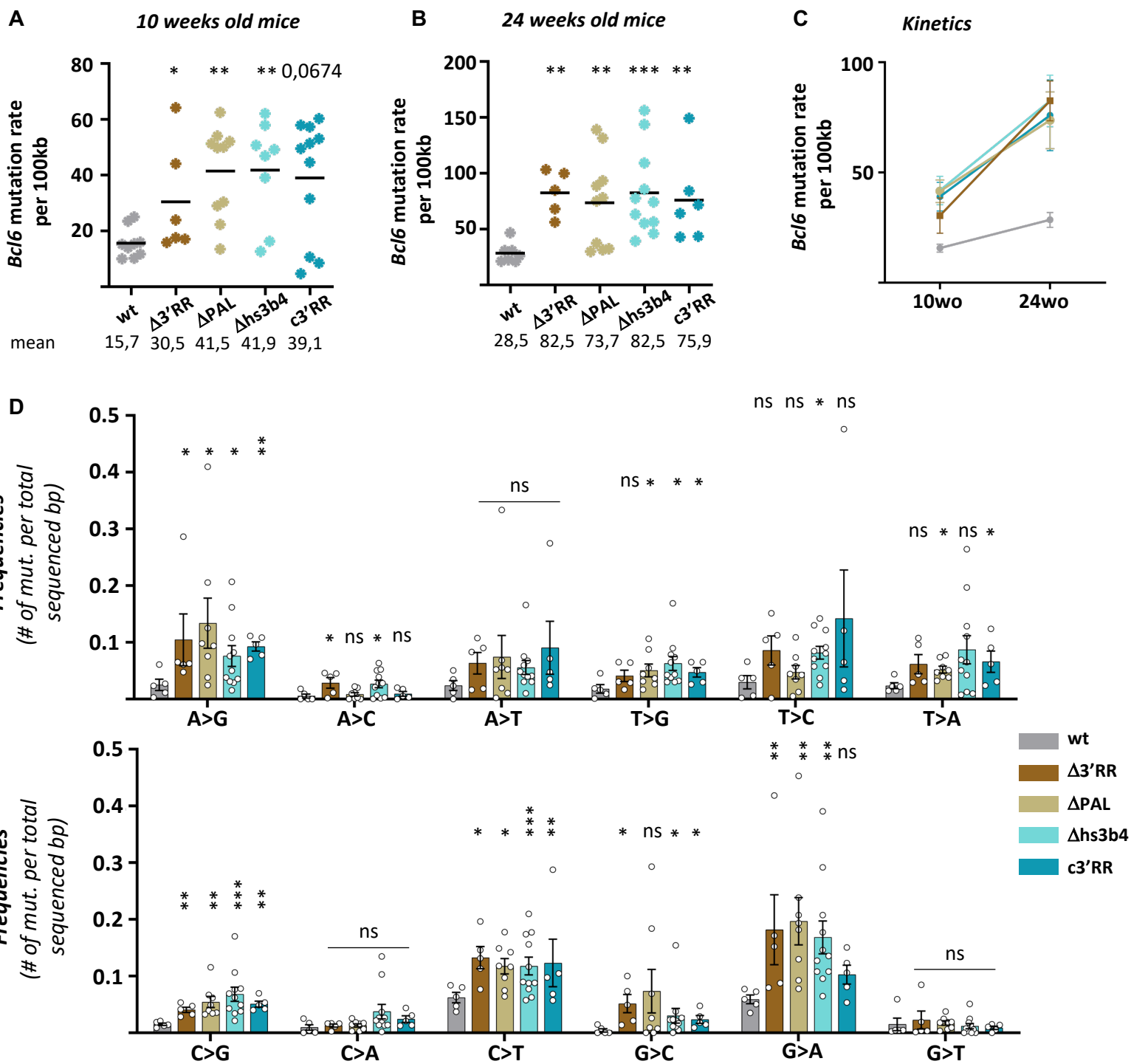
## Partie II. 3'RR, organisation et intégrité du génome de la cellule B.

De façon générale, l'intégrité du génome doit être préservée afin que les cellules maintiennent leur fonction et dans le but d'éviter toute dérive anormale (apoptose, arrêt ou emballement du cycle cellulaire etc...). L'organisation en trois dimensions du génome a récemment émergée comme un niveau de régulation important dans le maintien de son intégrité. En effet, il est montré que la structure de l'ADN dans le noyau est organisée en domaines topologiques, stabilisés par des protéines de structure telles que CTCF et contrôlés notamment par des éléments *cis*-régulateurs (Dixon et al., 2012; Nora et al., 2017). Ces domaines topologiques endossent des rôles importants dans le maintien de l'intégrité du génome en permettant par exemple la réparation de CDB de l'ADN (Arnould et al., 2021). En outre, la position tridimensionnelle des gènes dans le noyau peut avoir aussi un impact sur leur expression et leur stabilité. Par exemple, certains gènes sont positionnés à proximité de régions du génome qui sont plus sensibles aux cassures ou aux mutations de l'ADN, et peuvent donc être plus susceptibles de subir des altérations géniques (Zhang et al., 2012). De plus, la position des gènes dans le noyau influence également la façon dont ces derniers sont réparés (Lemaître et al., 2014).

Les cellules B sont des acteurs clés du système immunitaire qui produisent des anticorps hautement spécifiques nécessaires pour la défense de l'organisme. Cependant, la production de ces anticorps est associée à des remaniements géniques qui rendent la cellule B particulièrement vulnérable face aux dommages de l'ADN. Les CDB et les mutations, qui se produisent respectivement lors de la CSR et la SHM, peuvent engendrer des événements illégitimes tels que des translocations ou des mutations oncogéniques, susceptibles de conduire à la formation de lymphomes (Nussenzweig and Nussenzweig, 2010).

Plusieurs gènes « non-Ig » sont décrits comme étant des cibles illégitimes de AID. Ces gènes sont répartis en plusieurs groupes selon leur ciblage par AID et leur réparation : les gènes dits du groupe 1 et 2 sont des cibles préférentielles de AID. Néanmoins, le groupe 1 subit en plus des réparations infidèles tandis que le groupe 2 est particulièrement bien surveillé et réparé. Enfin, le groupe 3 contient les gènes non ciblés par AID (Liu et al., 2008).

L'organisation en trois dimensions du noyau pourrait jouer un rôle crucial dans le ciblage illégitime de gènes non *Ig* par l'enzyme AID. Il a été montré que la proximité physique entre le locus *IgH* et d'autres régions du génome peut favoriser la survenue de translocation médiée par AID (Rocha et al., 2012). Cette constatation suggère que la proximité entre le locus *IgH* et d'autres gènes pourrait également contribuer à la survenue de mutations dans les gènes « off-target ». Il a aussi été proposé que les éléments régulateurs tels que les super-enhancers peuvent réguler des gènes aussi bien en *cis* qu'en



**Figure 28 : Study of *Bcl6* mutation frequency**

**A.** Analysis of *Bcl6* mutation in 10 weeks old mice. **B.** Same as A in 24 weeks old mice. **C.** Kinetics of *Bcl6* mutation rate. **D.** Mutation pattern by DNA bases.

p-value was determined with two tailed Mann Whitney test, significant differences were indicated by:

\* P<0.05; \*\*P<0.01; \*\*\* P<0.001; \*\*\*\* P<0,0001 and error bars represent SEM

*trans* en interagissant avec les régions promotrices de gènes localisés sur d'autres chromosomes (Spilianakis et al., 2005; Williams et al., 2010).

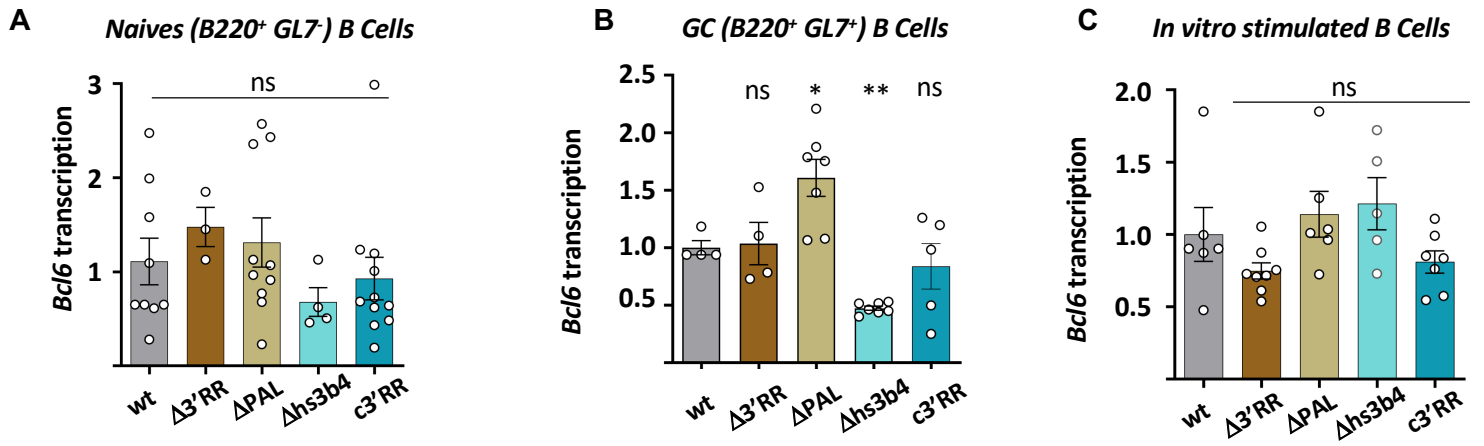
Les gènes d'immunoglobulines (*IgH*, *Igκ* et *Igλ*) possèdent tous des éléments *cis*-régulateurs impliqués dans la régulation des remaniements géniques. Initialement, leurs rôles ont été décrit dans un contexte de régulation transcriptionnelle. Puis, plusieurs études ont argumenté leur implication à un niveau plus global, notamment à l'échelle nucléaire (pour revue Bruzeau et al., 2022).

Précédemment, nous avons montré que la 3'RR, éléments *cis*-régulateur majeur du locus *IgH*, régule l'organisation nucléaire en pré-conformant le noyau des cellules B au repos pour le bon déroulement des étapes clés de remaniements géniques tardifs, tels que la recombinaison de classe (**Article n°1**). Il a également été montré que la 3'RR peut être subdivisée en deux unités fonctionnelles distinctes : le palindrome est indispensable à la SHM tandis que les core enhancers seuls sont suffisants pour permettre la CSR (Le Noir et al., 2017; Saintamand et al., 2016, 2015). Ces études suggèrent que la 3'RR pourrait également être impliquée dans les mécanismes de maintien de l'intégrité du génome. A l'aide d'un éventail de modèles murins (**Figure 27A**) ( $\Delta$ 3'RR (Vincent-Fabert et al., 2010b),  $\Delta$ PAL (Garot et al., 2016),  $\Delta$ hs3b4 (Pinaud et al., 2001) et  $c$ 3'RR (Le Noir et al., 2017)) nous testerons l'hypothèse selon laquelle la 3'RR, dans son intégralité, contribuerait au maintien de l'intégrité du génome des cellules B en organisant le noyau. Pour explorer cette hypothèse nous avons évalué l'introduction de mutations dans différents gènes « AID off-target » dans des cellules B activées de GC triées à partir des plaques de Peyer d'animaux (*wt* et mutants) âgés de 10 et 24 semaines (**Figure 27B**).

## II.1. Etude de *Bcl6*.

### ***La région 3'RR protège Bcl6 des mutations.***

Cette étude s'est concentrée sur la partie intronique située entre les exons 1 et 2 (*ie* : intron 1) du gène *Bcl6*, car décrite comme étant une cible privilégiée de AID (Liu et al., 2008). De façon inattendue, chez les souris âgées de 10 semaines, nos résultats indiquent une augmentation d'environ 2 fois la fréquence de mutation de *Bcl6* chez tous les modèles testés, par comparaison aux souris *wt* (**Figure 28A**). Pour étudier l'effet cumulatif des mutations avec l'âge, la même analyse a été réalisée chez des souris âgée de 24 semaines. A cet âge, la fréquence de mutation de *Bcl6* est significativement augmentée d'environ un facteur 3, dans toutes les souris mutantes, par comparaison aux souris *wt* (**Figure 28B**). La cinétique d'accumulation des mutations dans l'intron 1 de *Bcl6* entre 10 et 24 semaines met en évidence une augmentation plus rapide de la fréquence de mutations de *Bcl6* chez toutes les souris mutantes pour la 3'RR en comparaison aux souris *wt* (**Figure 28C**). Ces résultats suggèrent un rôle de la 3'RR comme « gardien » de l'intégrité génomique de *Bcl6*, le protégeant des mutations.

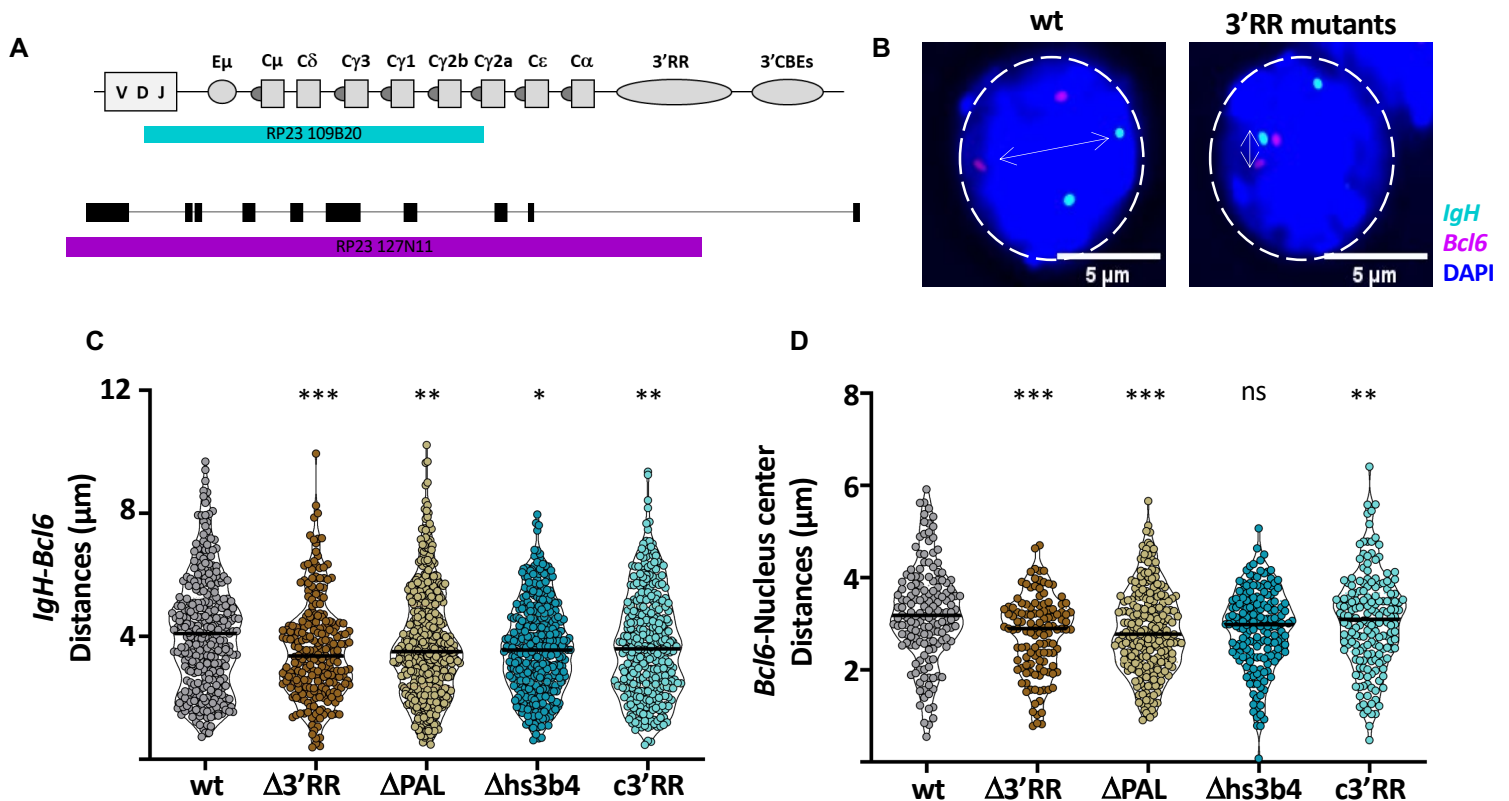


**Figure 29 : Quantification of *Bcl6* primary transcripts**

**A.** *Bcl6* transcription level relative to *Cd19* in naive B cells. **B.** Same as B in GC B Cells. **C.** Same as in B and C in *in vitro* stimulated B cells

p-value was determined with two tailed Mann Whitney test, significant differences were indicated by:

\* P<0.05; \*\*P<0.01; \*\*\* P<0.001; \*\*\*\* P<0,0001 and error bars represent SEM



**Figure 30 : *Bcl6* position in nucleus is modified.**

**A.** Scheme of FISH probe used in this study for *IgH* (RP23 109B20) and *Bcl6* (RP23 127N11). **B.** Representative pictures of DNA 3D-FISH. **C.** Distances between *IgH* and *Bcl6* loci in *in vitro* LPS stimulated B cells. **D.** Distances measured between *Bcl6* and the center of nucleus in *in vitro* LPS stimulated B cells.

p-value was determined with two tailed Mann Whitney test, significant differences were indicated by:

\* P<0.05; \*\* P<0.01; \*\*\* P<0.001; \*\*\*\* P<0,0001



Pour aller plus loin, le profil des mutations dans l'intron 1 de *Bcl6* chez les souris mutantes et *wt* a été analysé. Cette étude révèle un profil mutationnel différent chez tous les mutants, caractérisé notamment par des mutations de type transition C>T, G>A, A>G ainsi que de type transversion C>G, augmentée d'environ 2 fois chez les souris mutantes en comparaison des souris *wt* (**Figure 28D**). Cette signature mutationnelle évoque fortement une activité de l'enzyme AID. Ces résultats viennent suggérer que la perte d'intégrité de la 3'RR conduit à une augmentation du ciblage illégitime de *Bcl6* par l'enzyme AID.

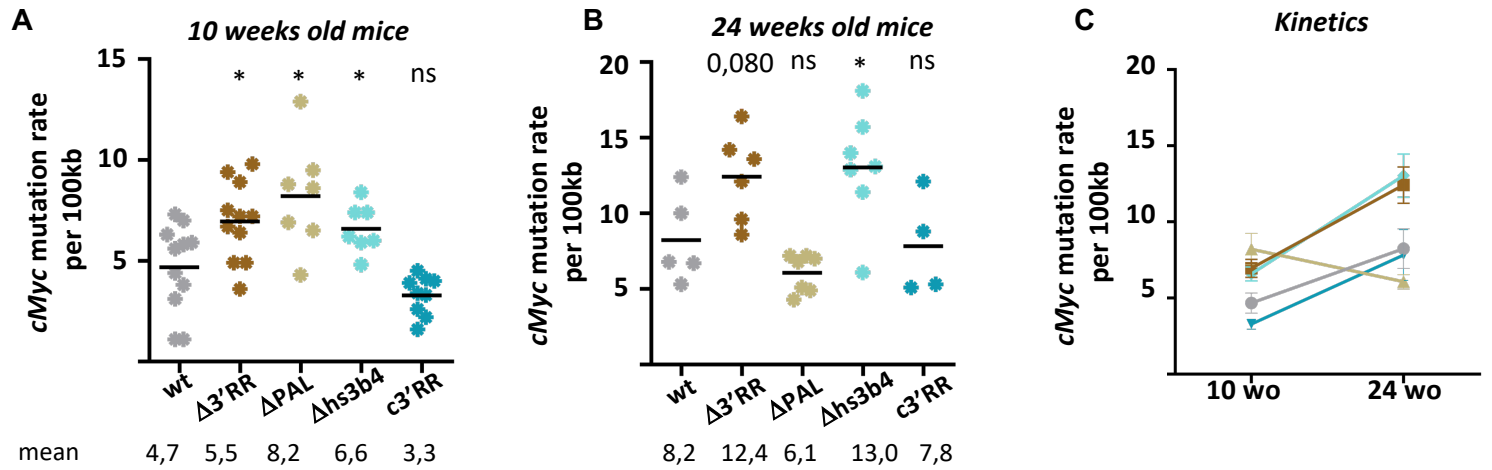
***L'augmentation des mutations de Bcl6 n'est pas corrélée à une augmentation de sa transcription.***

Il est admis que la transcription constitue un prérequis essentiel à la SHM (Fukita et al., 1998). La transcription primaire de *Bcl6* a donc été explorée par RT-qPCR dans des cellules B naïves (B220<sup>+</sup>GL7<sup>-</sup>) et activées du GC (B220<sup>+</sup>GL7<sup>+</sup>) triées à partir des plaques de Peyer afin de déterminer si l'augmentation de sa fréquence de mutation est corrélée à une augmentation de transcription. Dans les cellules B naïves, le niveau de transcription de *Bcl6* est identique entre les souris mutantes et *wt* (**Figure 29A**). Dans les cellules B du GC (qui subissent la SHM), des modifications de la transcription de *Bcl6* sont observées (**Figure 29B**) : si le modèle  $\Delta$ PAL est sujet à une augmentation significative de la transcription, le modèle  $\Delta$ hs3b4 voit quant à lui, une diminution de la transcription primaire de *Bcl6*. Dans les modèles  $\Delta$ 3'RR et  $c3'RR$ , la transcription primaire de *Bcl6* demeure inchangée. En parallèle, la transcription de *Bcl6* a également été étudiée dans des cellules B spléniques stimulées *in vitro* pendant 3 jours avec du LPS. Dans cette population activée, aucune modification significative de la transcription de *Bcl6* n'est observée dans les modèles mutants en comparaison au *wt* (**Figure 29C**).

Bien que les données obtenues à partir des cellules B GC restent à confirmer (en augmentant le nombre d'échantillons), il semble qu'il n'y ait pas de corrélation directe entre l'augmentation de la fréquence de mutation de *Bcl6* et son niveau de transcription. En effet, les modèles  $\Delta$ hs3b4 et  $\Delta$ PAL présentent des niveaux de transcription de *Bcl6* opposés alors que l'accumulation de mutations dans la région intronique de ce même gène est significativement augmentée dans ces deux modèles.

***La délétion de la 3'RR induit un repositionnement de Bcl6 dans le noyau des lymphocytes B activés.***

Afin d'explorer l'hypothèse du lien entre position nucléaire et mutations, nous avons étudié le positionnement de *Bcl6* par une méthode d'hybridation *in situ* fluorescente en 3 dimensions (FISH-3D ADN) dans les noyaux interphasiques des lymphocytes B stimulés *in vitro* (**Figure 30A**). Nous avons dans un premier temps mesuré les distances entre les loci *IgH* (RP23 109B20) et *Bcl6* (RP23 127N11). De façon surprenante, la délétion partielle ( $\Delta$ PAL,  $\Delta$ hs3b4 et  $c3'RR$ ) ou totale ( $\Delta$ 3'RR) de la 3'RR conduit

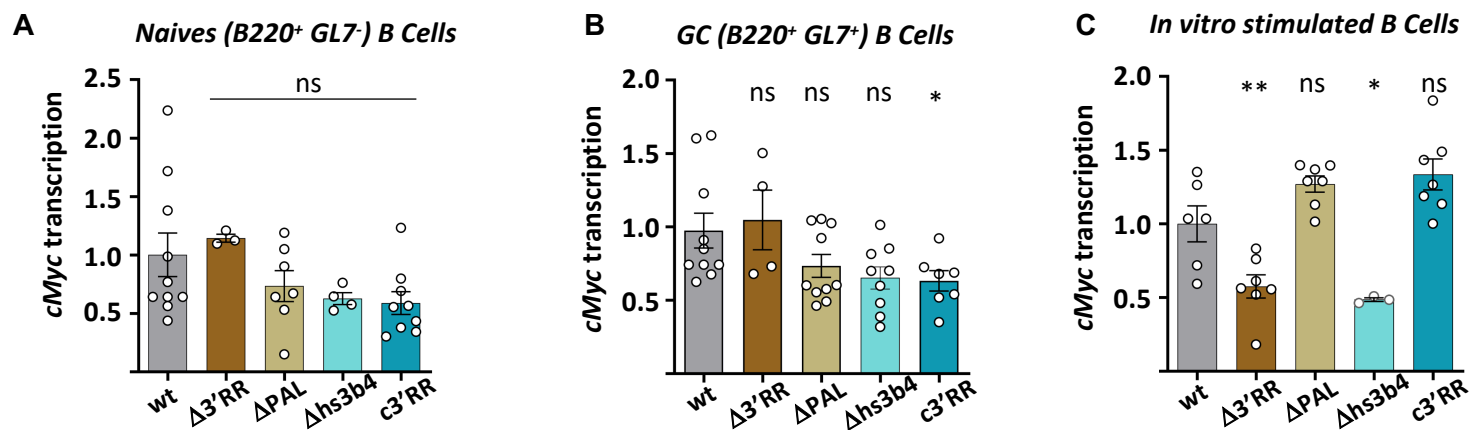


**Figure 31 : Study of cMyc mutations**

**A,B.** Analysis of cMyc mutation frequency in 10 weeks old mice (**A**) and 24 weeks old mice (**B**). **C.** Kinetics of cMyc mutation rate.

p-value was determined with two tailed Mann Whitney test, significant differences were indicated by:

\* P<0.05; \*\*P<0.01; \*\*\* P<0.001; \*\*\*\* P<0,0001 and error bars represent SEM



**Figure 32 : Quantification of cMyc primary transcription**

**A,B,C.** cMyc transcription level relative to Cd19 by RT-qPCR in naive B cells (**A**), GC B cells (**B**) and *in vitro* stimulated B cells (**C**).

p-value was determined with two tailed Mann Whitney test, significant differences were indicated by:

\* P<0.05; \*\*P<0.01; \*\*\* P<0.001; \*\*\*\* P<0,0001 and error bars represent SEM

à une diminution significative de la distance entre les loci *IgH* et *Bcl6*, traduisant un rapprochement de ces deux loci distincts (**Figures 30B et C**).

Ces données suggèrent un rôle protecteur de la 3'RR en maintenant les loci *Bcl6* et *IgH* à distance l'un de l'autre, et ainsi protégeant *Bcl6* des mutations induites par AID.

La localisation d'un gène à la périphérie du noyau est souvent associée à son ancrage à la lamina nucléaire et à son inclusion dans les domaines associés à la lamine (LAD). Les LADs sont des entités généralement considérées comme des régions répressives du génome (van Steensel and Belmont, 2017).

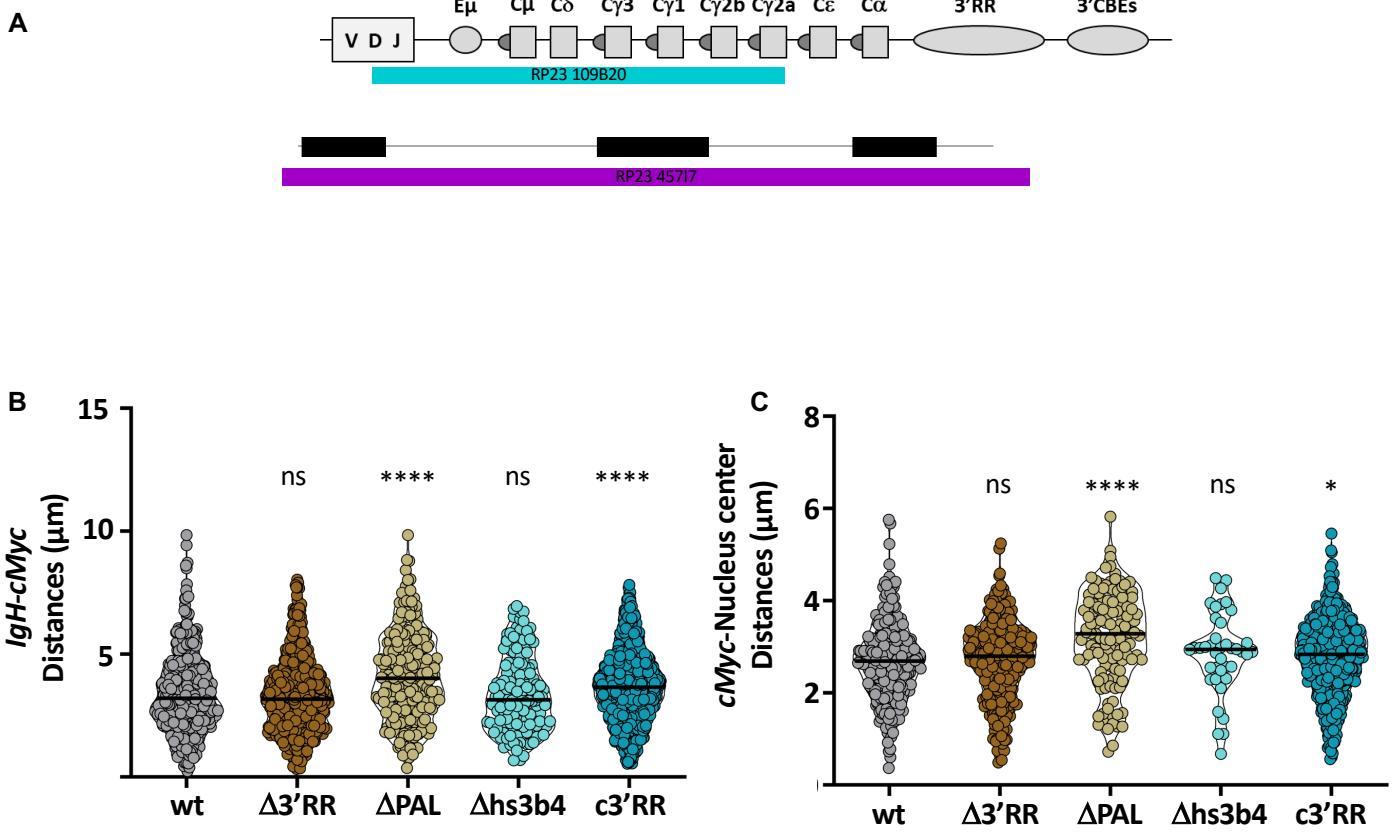
Dans le contexte de l'introduction de mutations, il est possible que la localisation d'un gène à la périphérie puisse influencer sa stabilité génomique, notamment en limitant l'accès à AID ou aux enzymes de réparation de l'ADN. Bien que controversée par l'équipe de Subhajyoti De (Smith et al., 2017), les résultats d'une étude de 2018, menée par le groupe de Braun, suggèrent que la localisation d'un gène à la lamine conduit à une diminution de sa fréquence de mutation (Klymenko et al., 2018). Il aurait donc été pertinent d'évaluer les distances entre loci ciblés par AID et la périphérie nucléaire. Toutefois, notre méthode d'analyse des images de FISH-3D ADN ne permet pas de situer de façon reproductible et précise la zone périphérique du noyau. Pour pallier à cette difficulté, nous avons fait le choix d'évaluer un paramètre inverse : la distance des allèles au centre du noyau. En parallèle des distances intergéniques *IgH-Bcl6*, les distances entre les loci *Bcl6* et le centre du noyau de cellules B stimulées *in vitro* ont donc été mesurées. Une diminution significative de la distance entre *Bcl6* et le centre du noyau est notée dans tous les modèles mutants (**Figure 30D**). Ceci traduit une relocalisation des loci *Bcl6* dans une zone plus centrale du noyau lorsque l'intégrité de la 3'RR est compromise. Ceci vient renforcer l'hypothèse selon laquelle la 3'RR est impliquée dans la protection de l'intégrité génomique en empêchant notamment la localisation de *Bcl6* au sein de zones du noyau qui sont permissives à l'introduction de mutations.

## II.II. Etude de *cMyc*.

En parallèle de l'étude de l'analyse des mutations dans le gène *Bcl6*, nous avons étudié l'introduction de mutations au sein d'une autre cible illégitime de l'enzyme AID : *cMyc*, gène du groupe 2 (Liu et al., 2008).

### ***L'enhancer hs4, gardien de cMyc ?***

A l'âge de 10 semaines, une augmentation significative de la fréquence de mutation de *cMyc* est observée uniquement dans deux modèles :  $\Delta PAL$  et  $\Delta hs3b4$  (**Figure 31A**). Cependant il est intéressant de noter qu'à 24 semaines, la fréquence des mutations de *cMyc* cesse d'augmenter dans le modèle



**Figure 33 : Nuclear localization of *cMyc***

**A.** Scheme of FISH probe used in this study for *IgH* (RP23 109B20) and *cMyc* (RP23 45717). **B,C.** Study of *cMyc* position by DNA-3D-FISH relative to *IgH* (**B**) and nucleus center (**C**).

p-value was determined with two tailed Mann Whitney test, significant differences were indicated by:

\* P<0.05; \*\*P<0.01; \*\*\* P<0.001; \*\*\*\* P<0,0001

$\Delta PAL$ , tandis qu'une augmentation se poursuit dans le modèle  $\Delta hs3b4$  (**Figure 31B**). Les souris  $\Delta 3'RR$  présentent quant à elles une augmentation, toutefois non significative, de la fréquence de mutation de *cMyc* à 24 semaines, (**Figure 31B**). La cinétique des mutations de *cMyc* dans le modèle  $c3'RR$  est comparable à celle observée chez les souris *wt*, et ce à 10 et 24 semaines (**Figures 31A et B**). En présentant un niveau constant de mutations sur le gène *cMyc*, le modèle  $\Delta PAL$  diffère clairement des animaux  $c3'RR$  et *wt* chez lesquels cet oncogène accumule des mutations. Néanmoins, les modèles  $\Delta 3'RR$  et  $\Delta hs3b4$  se distinguent des autres par l'accumulation plus importante de mutations dans l'intron 1 de *cMyc* au cours du temps (**Figure 31C**).

Ces deux derniers modèles sont les seuls modèles dépourvus de l'enhancer *hs4*, suggérant ainsi un rôle spécifique important de ce dernier.

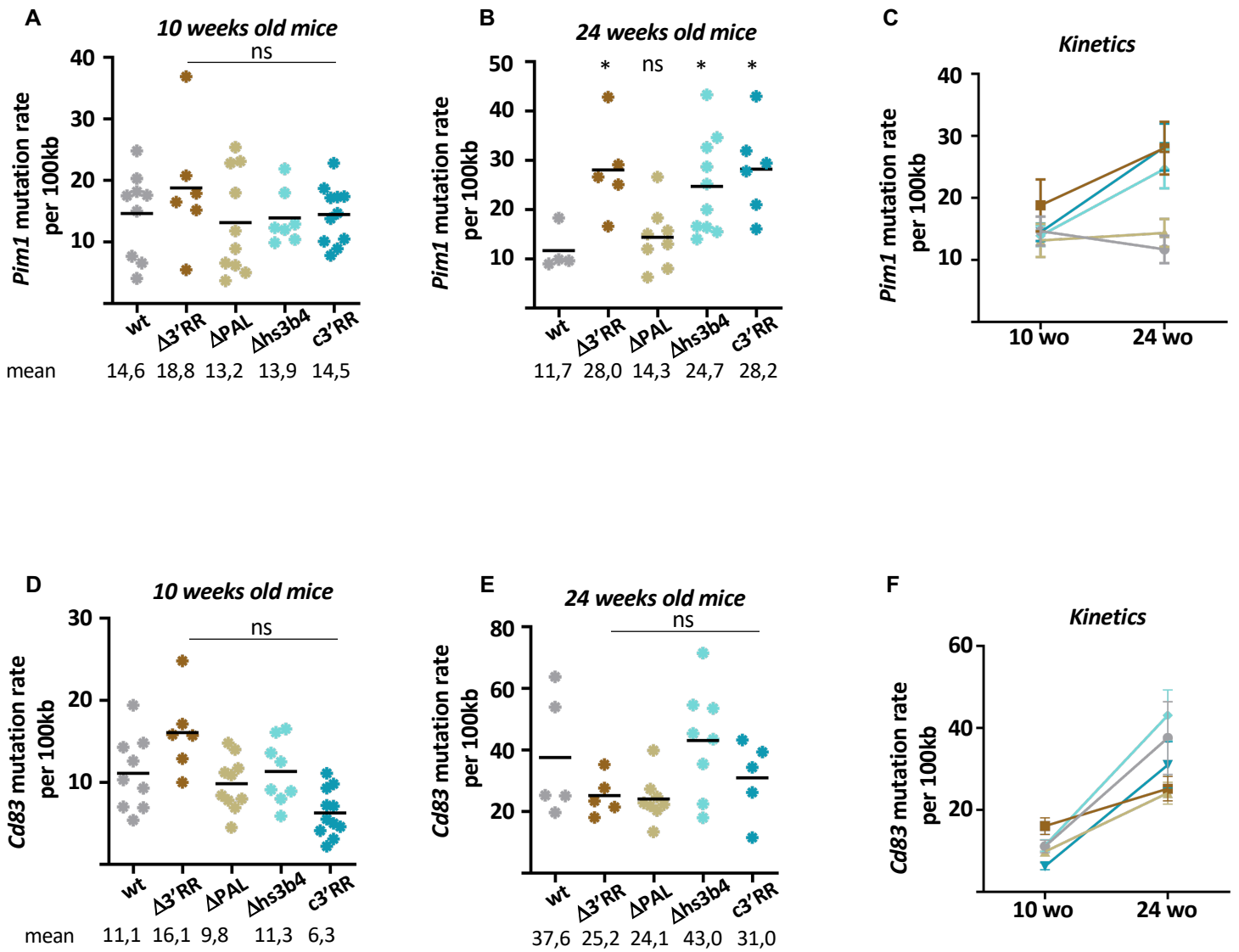
### ***La fréquence de mutation n'est pas corrélée à la transcription***

L'introduction de mutations pendant la SHM étant au moins en partie corrélée à la transcription, nous avons exploré si les différences de fréquence de mutation observée étaient liées au niveau de transcription de *cMyc*. Pour cela, nous avons quantifié la transcription primaire de *cMyc* relative à *Cd19* par RT-qPCR dans différentes populations : les cellules B naïves ( $B220^+GL7^-$ ) et de GC ( $B220^+GL7^+$ ) triées à partir des plaques de Peyer ainsi que dans des cellules B spléniques stimulées *in vitro* pendant trois jours avec du LPS. Les résultats montrent que le niveau de transcription primaire de *cMyc* ne diffère pas entre les modèles mutants et le *wt* dans les cellules B naïves ou de GC (**Figures 32A et B**). En revanche, dans les cellules B stimulées *in vitro*, une diminution de la transcription de *cMyc* est observée dans les modèles  $\Delta 3'RR$  et  $\Delta hs3b4$  alors qu'elle reste inchangée dans les modèles  $\Delta PAL$  et  $c3'RR$ , comparativement au contexte *wt* (**Figure 32C**). Cette observation est surprenante puisque les modèles  $\Delta 3'RR$  et  $\Delta hs3b4$ , qui présentent une diminution de la transcription, sont ceux qui accumulent le plus de mutation au fil du temps.

L'ensemble de ces données suggère que l'accumulation de mutation dans la première région intronique de *cMyc* ne serait pas directement corrélée à son niveau de transcription.

### ***La proximité de cMyc et IgH semble impliquée dans l'introduction des mutations***

Nous avons également examiné la relation entre la position des loci *cMyc* par rapport à *IgH*, à l'instar de notre étude sur l'oncogène *Bcl6*. Nous avons mesuré la distance entre les loci *cMyc* et *IgH*, dans des noyaux interphasiques de cellules B stimulées *in vitro* avec du LPS pendant trois jours, après marquage des loci en FISH-3D ADN (*IgH* : RP23 109B20 et *cMyc* : RP23 45717) (**Figure 33A**). Contrairement aux résultats obtenus pour l'oncogène *Bcl6*, les modèles pour lesquels la fréquence de mutation de *cMyc* est plus élevée ( $\Delta 3'RR$  et  $\Delta hs3b4$ ) ne présentent pas de rapprochement de leurs loci *cMyc* et *IgH* : ces



**Figure 34 : Mutation analysis in intron 1 of *Pim1* and *Cd83***

**A,B.** Analysis of *Pim1* mutation frequency in 10 weeks old mice (**A**) and 24 weeks old mice (**B**). **C.** Kinetics of *Pim1* mutation rate. **D,E.** Analysis of *Cd83* mutation frequency in 10 weeks old mice (**D**) and 24 weeks old mice (**E**). **F.** Kinetics of *Cd83* mutation rate p-value was determined with two tailed Mann Whitney test, significant differences were indicated by: \* P<0.05; \*\*P<0.01; \*\*\* P<0.001; \*\*\*\* P<0,0001

derniers restent à une distance équivalente à celle mesurée en contexte *wt* (**Figure 33B**). En revanche, pour les modèles dans lesquels *cMyc* est muté de façon similaire au contexte *wt* ( $\Delta$ *PAL* et *c3'RR*), les loci *cMyc* et *IgH* s'éloignent (**Figure 33B**). Bien que cela soit étonnant, cette observation corrobore avec des données publiées de 4C-Seq faisant état d'interactions entre *IgH* et *cMyc* dans des cellules B *wt* activées (Rocha et al., 2012). Ces résultats « d'absence d'éloignement » sont également compatibles avec ceux décrits du groupe de Peter Fraser qui montre, par RNA-FISH, que *cMyc* est relocalisé dans la même unité de transcription qu'*IgH* à l'activation des cellules B (Osborne et al., 2007).

Nous avons également étudié la position globale de *cMyc* dans le noyau en mesurant la distance entre les loci *cMyc* et le centre du noyau. Une fois de plus, nos observations révèlent un comportement différent de celui observé pour *Bcl6*. Dans les modèles  $\Delta$ *3'RR* et  $\Delta$ *hs3b4*, chez lesquels la fréquence de mutation de *cMyc* est plus élevée, les loci *cMyc* ne sont pas relocalisés dans le noyau (**Figure 33C**). En revanche, dans les modèles  $\Delta$ *PAL* et *c3'RR*, où la fréquence de mutation de *cMyc* n'est pas plus élevée, les loci *cMyc* sont repositionnés plus en périphérie du noyau en comparaison au *wt*.

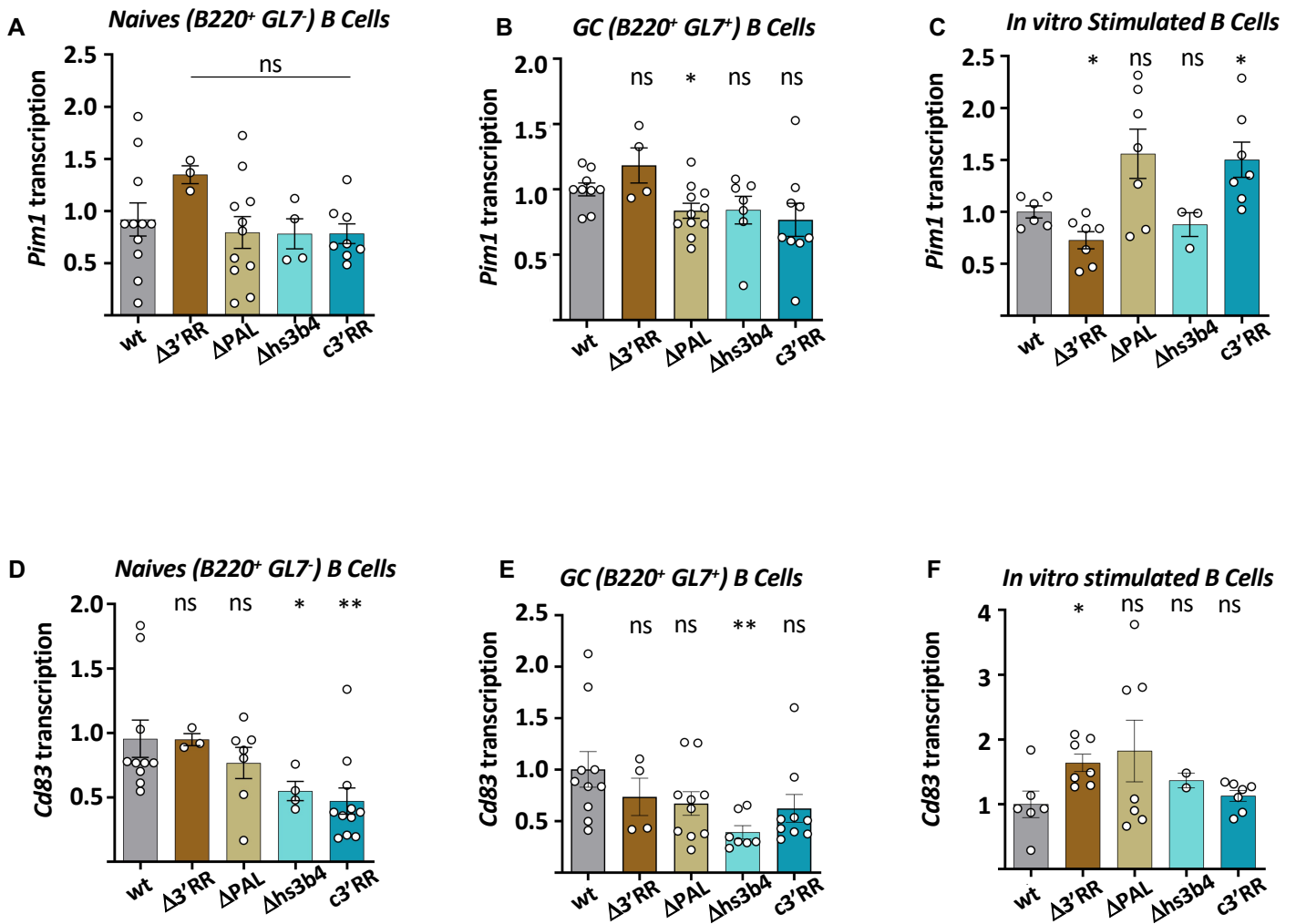
L'ensemble de ces éléments suggère que la cinétique de mutation et de position des loci *cMyc* et *Bcl6* diffèrent lors de l'activation B. Bien que ces modifications semblent être influencées par la *3'RR*, le mécanisme impliqué diffère pour ces deux loci. Ces deux gènes (*Bcl6* et *cMyc*) appartiennent à deux groupes différents : *Bcl6* est fréquemment muté, même en condition *wt*, car il est à la fois cible de AID et des voies de réparation infidèles. En revanche, la fréquence de mutation de *cMyc* est normalement plus faible car ce gène est réparé de façon fidèle. Il pourrait être proposé que la *3'RR*, et plus particulièrement *via* l'enhancer *hs4*, soit impliquée dans le recrutement de facteurs des voies de réparation fidèle.

### **II.III. Etude préliminaire de *Pim1* et *Cd83*.**

Afin d'étayer notre hypothèse selon laquelle la *3'RR* participe à l'organisation du noyau, nous avons étendu notre étude de mutations et de positionnement de loci à deux autres gènes « AID off-target » du groupe 1, *Pim1* et *Cd83*.

#### ***Etude des fréquences de mutation***

Nous avons analysé la fréquence de mutation dans la première région intronique de *Pim1*, décrite comme étant ciblée par AID, chez des cellules B du GC triées à partir de plaques de Peyer d'animaux *wt* et mutants âgés de 10 et 24 semaines. Chez les animaux mutants de 10 semaines, la fréquence de mutation de *Pim1* est identique à celle observé chez les animaux *wt* (**Figure 34A**). A 24 semaines, la fréquence de mutation de *Pim1* apparaît augmentée chez tous les animaux mutants, à l'exception du



**Figure 35 : Quantification of primary transcription**

**A,B,C.** *Pim1* transcription level relative to *Cd19* by RT-qPCR in naive B cells (**A**), GC B cells (**B**) and *in vitro* stimulated B cells (**C**). **D,E,F.** *Cd83* transcription level relative to *Cd19* in naive B cells (**D**), GC B cells (**E**) and *in vitro* stimulated B cells (**F**).

p-value was determined with two tailed Mann Whitney test, significant differences were indicated by:

\*  $P < 0.05$ ; \*\*  $P < 0.01$ ; \*\*\*  $P < 0.001$ ; \*\*\*\*  $P < 0.0001$  and error bars represent SEM



modèle  $\Delta PAL$  (**Figure 34B**). La cinétique d'accumulation des mutations au cours du temps révèle deux cas distincts : le modèle  $\Delta PAL$  se comporte de la même façon que le *wt* tandis que l'ensemble des autres mutants ( $\Delta 3'RR$ ,  $\Delta hs3b4$  et  $c3'RR$ ) accumulent des mutations de façon marquée (**Figure 34C**). La fréquence de mutation de *Cd83* ne présente aucune différence significative chez les animaux mutants âgés de 10 semaines en comparaison aux animaux *wt* (**Figure 34D**) ; il en va de même chez les animaux âgés de 24 semaines (**Figure 34E**). La représentation cinétique montre que les animaux mutants accumulent des mutations dans l'intron 1 de *Cd83* de façon comparable aux animaux *wt* (**Figure 34F**).

Ces données révèlent un impact foncièrement différent de la  $3'RR$  sur les deux cibles *Pim1* et *Cd83*. En effet, l'accumulation de mutations sur *Pim1* est globalement augmentée en absence de la  $3'RR$  (à l'exception de  $\Delta PAL$ ) alors que les mutations dans l'intron 1 de *Cd83* s'accumulent indépendamment de la région régulatrice ; ceci suggère que *Cd83* n'est pas contrôlé par la  $3'RR$ .

#### **Quantification de la transcription primaire**

Pour les mêmes raisons que précédemment, nous avons quantifié la transcription primaire de *Pim1* dans différentes populations de cellules B. Dans les cellules B naïves (B220<sup>+</sup>GL7<sup>-</sup>), aucune différence significative de transcription de *Pim1* n'est observée dans les modèles mutants de la  $3'RR$  (**Figure 35A**). Il en va de même dans les cellules B activées de GC (B220<sup>+</sup>GL7<sup>+</sup>) (**Figure 35B**). L'étude de la transcription de *Pim1* dans des cellules B stimulées *in vitro* révèle quant à elle des disparités : diminuée dans le modèle  $\Delta 3'RR$ , inchangée dans le modèle  $\Delta hs3b4$  et augmentée chez les souris  $c3'RR$  (**Figure 35C**). Les différences marquées de transcription observées dans ces modèles contrastent avec les données de mutations qui, quant à elles, ne varient pas entre ces différents modèles (fréquence de mutation augmentée). Ces observations suggèrent une nouvelle fois, avec pour exemple *Pim1*, qu'il existe une régulation « décorrélée » entre transcription et mutations, nous suspectons ici que le positionnement nucléaire de ce gène pourrait être impliquée.

Nous avons également quantifié la transcription primaire de *Cd83* relative à *Cd19* par RT-qPCR. Dans les cellules B naïves (B220<sup>+</sup>GL7<sup>-</sup>), la transcription de *Cd83* est inchangée dans les modèles  $\Delta 3'RR$  et  $\Delta PAL$  alors qu'elle est significativement diminuée dans les modèles  $\Delta hs3b4$  et  $c3'RR$  (**Figure 35D**). Dans les cellules B activées de GC (B220<sup>+</sup>GL7<sup>+</sup>), une diminution de la transcription n'est observée que dans le modèle  $\Delta hs3b4$  (**Figure 35E**). Dans les cellules B stimulées *in vitro*, la transcription de *Cd83* est inchangée dans la plupart des modèles mutants à l'exception du modèle  $\Delta 3'RR$  pour lequel une augmentation significative des transcrits *Cd83* est observée (**Figure 35F**).

Dans leur ensemble, les données collectées à partir de notre collection de modèles mutants mettent en lumière les effets « à la carte » de la région 3'RR sur la régulation des gènes « AID off-target ». Ces observations suggèrent des mécanismes de régulation spécifiques pour chaque gène.

Afin de déterminer si l'organisation du noyau, et plus particulièrement la position relative des gènes « AID off-target », est corrélée aux mutations subies par ces cibles dans les souris mutantes pour la 3'RR, il nous apparaît nécessaire de poursuivre l'étude de la position relative des gènes *Pim1* et *Cd83* par rapport à *IgH*. Compte tenu de nos précédentes observations, la position relative de *Pim1* par rapport à *IgH* devrait être modifiée dans tous les modèles mutants excepté  $\Delta$ *PAL* tandis que la position de *Cd83* par rapport à *IgH* devrait être inchangée dans tous les mutants. Cette étude pourrait également être étendue à d'autres gènes « AID off-target » tels que *Pax5*, *Bcl2*... afin de tenter d'établir, si possible, une relation « logique » entre position des gènes dans le noyau, intégrité de la 3'RR et mutations.

Ces résultats sont discutés dans leur ensemble dans la partie **Discussion et Perspectives**.

## *Discussion & Perspectives*

---

## Discussion et Perspectives

---

A l'aide de différents modèles murins, le laboratoire s'est illustré dans la description du rôle des régions *EμMARs* et *3'RR* dans le contrôle transcriptionnel des événements de remaniements géniques spécifiques de la cellule B. Mon travail de thèse a consisté à étudier l'implication de ces deux régions *cis*-régulatrices majeures du locus *IgH* dans l'organisation du noyau du lymphocyte B ainsi que dans le maintien de l'intégrité du génome des cellules B matures.

### ***Implication des régions 3'RR et EμMARs dans l'organisation du génome de la cellule B***

La *3'RR*, super enhancer qui s'étend sur 30kb en aval de *Cα*, possède une structure bien particulière : elle est composée de quatre core enhancers (*hs3a*, *hs1.2*, *hs3b* et *hs4*) inclus dans un quasi-palindrome, formé par des séquences inversées répétées bordées par les enhancers *hs3a* et *hs3b* et dont l'élément central est l'enhancer *hs1.2* (Chauveau and Cogné, 1996; Lieberson et al., 1991; Matthias and Baltimore, 1993; Michaelson et al., 1995). Cette structure palindromique (mais pas la séquence) est particulièrement bien conservée au cours de l'évolution, suggérant un rôle important de celle-ci (Sepulveda et al., 2005). Le laboratoire, à l'aide de différents modèles murins, a pu mettre en évidence le rôle fonctionnel des enhancers dans le processus de switch isotypique tandis que l'environnement quasi-palindromique de la *3'RR* est nécessaire au processus de SHM (Garot et al., 2016; Le Noir et al., 2017; Saintamand et al., 2016). Il a été démontré que la *3'RR* est un élément régulateur majeur du locus *IgH* à différents niveaux : (i) transcriptionnel et nucléosomal (Saintamand et al., 2015), (ii) nucléaire (Le Noir et al., 2017) et (iii) supranucléosomal (dans la formation de boucles de chromatine) (Wuerffel et al., 2007; X. Zhang et al., 2019). Toutefois, ces études ne se sont pas intéressées à l'implication de la structure de la *3'RR* (enhancers vs structure palindromique) à ces niveaux de régulation. Nous avons donc exploré le rôle de la structure palindromique et des enhancers de la *3'RR* aux échelles nucléaires (position des gènes et territoires chromosomiques) et supranucléosomales (compartiments A vs B et boucles de chromatine). Nous avons montré, dans l'**article n°1**, que la structure palindromique de la *3'RR* n'est pas impliquée dans la régulation du locus *IgH* à ces deux échelles. En revanche, la seule présence des quatre core enhancers (sans le « packaging » palindromique) est suffisante et nécessaire à l'adressage correcte des loci *IgH* dans le noyau ainsi qu'à la formation de boucles de chromatine requises pour la CSR. Nous avons également montré que l'accessibilité du locus *IgH* ne requiert pas la présence de la structure palindromique.

La diminution de l'accessibilité chromatiniennne du locus *IgH* dans le modèle  $\Delta 3'RR$  pourrait s'expliquer par le fait que la *3'RR* recrute des facteurs régulateurs de la chromatine. En effet, l'étude des marques épigénétiques, par immunoprécipitation de chromatine (ChIP), dans différents modèles murins de délétion/destructuration de la *3'RR*, a mis en évidence une diminution des marques de chromatine

ouverte (H3K4me3, H3K27Ac) au sein de différentes régions du locus *IgH* (Saintamand et al., 2016). La *3'RR* permet notamment le recrutement de BRG1, protéine faisant parti du complexe de remodelage de la chromatine SWI/SNF. Le recrutement de ce facteur à la *3'RR* a initialement été décrit dans les stades immatures de différenciation B (Holley et al., 2014; Osipovich et al., 2009). Plus récemment, le recrutement de BRG1 sur la *3'RR* a été démontré pendant la formation du GC (Schmiedel et al., 2021). Cette dernière étude décrit BRG1 comme se fixant préférentiellement sur les enhancers pour ouvrir la chromatine et ainsi activer la transcription. Ce facteur pourrait donc continuer à jouer son rôle dans le modèle *c3'RR* en se fixant sur les enhancers. D'autres facteurs sont recrutés à la *3'RR*, c'est le cas par exemple de ZMYND8, protéine décrite comme « lecteur de chromatine » avec un rôle prépondérant dans l'initiation de la transcription et, dont la délétion entraîne un phénotype similaire à celui de la délétion de la *3'RR* (diminution de la CSR et de la SHM) (Delgado-Benito et al., 2018). Compte tenu du phénotype normal observé dans le modèle *c3'RR*, la seule présence des enhancers suffirait à permettre le recrutement de ZMYND8.

Ainsi, nous proposons que les core enhancers permettent de préconfigurer le noyau des cellules B au repos afin que ces dernières puissent exercer correctement leurs remaniements géniques (CSR) lors de l'activation. Nos données viennent contredire une étude, publiée par le groupe de R. Pavri cette année, dans laquelle l'implication de la *3'RR* dans la formation du CSRC est (re)mise en cause (Costea et al., 2023). Cette étude est plutôt en faveur de l'implication de la région *3'CBE* (et non de la *3'RR*) dans l'établissement du CSRC. Cependant, cette étude ne dispose pas de modèle de délétion pour la région *3'CBE* pour confirmer ses dires et ne montre cet effet qu'en condition de stimulation LPS+IL4, induisant un switch vers IgG1. A l'inverse, nos données montrent que la formation des boucles d'ADN n'est pas altérée dans un modèle de délétion partielle de la région *3'CBE* (*Δhs567*) chez lequel nous n'avons observé aucun défaut d'interactions, en accord avec le niveau normal de CSR déjà décrit pour ce modèle (Volpi et al., 2012).

Les observations concernant la configuration tridimensionnelle du locus *IgH* (absence de la grande boucle et des boucles internes) sont en accord avec le défaut de CSR observé dans le modèle *Δ3'RR* et le niveau de CSR normal dans le modèle *c3'RR*. En revanche, ces deux modèles présentent une diminution drastique de la SHM (Le Noir et al., 2017). Ceci suggère que dans les cellules B de GC, cellules dans lesquelles la SHM se produit, les mécanismes qui régissent l'organisation du noyau et la configuration tridimensionnelle du locus *IgH* sont différents de ceux agissant dans les cellules B activées *in vitro* (recombinaison de CSR mais pas de SHM). Ce que nous avons appelé la grande boucle, mettant en contact les régions *EμMARS* et *3'RR* ne semble donc pas intervenir dans la régulation du processus de SHM. En outre, la différence de configuration tridimensionnelle du locus *IgH* entre les modèles *Δ3'RR* et *c3'RR* suggère l'implication de la structure palindromique dans la régulation du mécanisme de SHM. Il est donc envisageable qu'une autre boucle se forme afin de mettre en contact

les régions  $3'RR$  et  $E\mu MARS$  avec les promoteurs  $pV_H$  des segments géniques réarrangés. Cependant, l'exploration de cette hypothèse, à ce jour, s'avère techniquement compliquée de par le nombre de cellules requis pour la 3C-HTGTS et le faible nombre de cellules collectable dans un GC de souris. Une manière de contourner ce problème serait d'induire des cellules B de GC *in vitro* à partir de cellules B naïves spléniques. Cette technique permet, grâce à la présence de cellules nourricières délivrant les facteurs d'activation des cellules B, d'obtenir des cellules « GC-like », dont le phénotype se rapproche de celui d'une cellule B du GC, en très grand nombre (Haniuda and Kitamura, 2019).

A plus long terme, il serait intéressant d'étendre l'étude du rôle de la  $3'RR$  sur le positionnement des loci  $IgH$  dans des plasmocytes. En effet, à ce stade, il a été montré par RNA-FISH qu'un allèle de chaque loci contribuant à l'expression de l'Ig ( $IgH$ ,  $Ig\kappa$  et  $IgJ$ , ce dernier codant la chaîne J, polypeptide essentiel pour la sécrétion d'anticorps sous forme multimérique) sont colocalisés « en trio » à la périphérie du noyau. De plus, il a été démontré par ChIP-3C-Seq que les loci  $IgH$  et  $Ig\kappa$  interagissent entre eux et que ces interactions sont médiées par des éléments *cis*-régulateurs du locus  $Ig\kappa$ , notamment  $3'E\kappa$  (Park et al., 2014). La  $3'RR$ , qui peut être considérée comme l'analogue de  $3'E\kappa$ , pourrait donc elle aussi être impliquée dans ce processus de régulation.

La région  $E\mu MARS$ , localisée sur moins d'1kb entre le dernier segment  $J_H$  et  $S\mu$ , est constituée de l'enhancer  $cE\mu$  flanqué de part et d'autre par les régions  $MARS_{E\mu}$ . A l'image de la  $3'RR$ , une dualité fonctionnelle de ces éléments a été mise en lumière. En effet, alors que l'enhancer  $cE\mu$  joue un rôle majeur pendant la recombinaison VDJ, le rôle des  $MARS_{E\mu}$  prend place à un stade plus tardif, notamment lors de la SHM (Marquet et al., 2014; Martin et al., 2023; Perlot et al., 2005). Il a également été montré à plusieurs reprises que la région  $E\mu MARS$  est mise en contact avec la  $3'RR$  par la formation d'une boucle de chromatine (Wuerffel et al., 2007; X. Zhang et al., 2019). Cependant, l'implication précise de  $cE\mu$  et des  $MARS_{E\mu}$  dans l'organisation tridimensionnelle du locus  $IgH$  reste peu explorée.

L'étude des interactions, par 3C-HTGTS, dans les modèles  $\Delta E\mu FL$  et  $\Delta cE\mu$  montre une diminution importante de la fréquence de contact entre la  $3'RR$  et la région  $E\mu MARS$ , soulignant ainsi le rôle primordial de l'enhancer  $cE\mu$  dans la formation de cette grande boucle dans les cellules B au repos. Cette observation vient en partie contredire l'étude princeps effectuée en 2007 par le groupe de Amy L. Kenter dans laquelle la délétion de l'enhancer  $cE\mu$  n'entraîne aucun défaut de la fréquence d'interaction déterminée par méthode de 3C-qPCR (Wuerffel et al., 2007). Cette différence de résultat peut s'expliquer par des limites techniques, notamment imposées par l'utilisation de l'enzyme *HindIII* qui possède un site de restriction de 6pb et permet donc une résolution moindre que notre approche avec *Csp6I*, enzyme reconnaissant un site de restriction de 4pb. Une autre limite pourrait aussi concerner le phénotype des modèles murins eux-mêmes. En effet, les modèles  $\Delta E\mu FL$  et  $\Delta cE\mu$  présentent une proportion d'allèle non productif supérieure à celle rencontrée en contexte *wt*

(Marquet et al., 2014; Perlot et al., 2005). La formation de boucles de chromatine au sein de ces allèles non productifs pose donc question et pourrait alors expliquer la diminution observée puisque la technique de 3C-HTGTS révèle les évènements prenant place indifféremment sur les deux allèles. Cependant, la survenue d'évènements de CSR inter-allélique (mise en lumière dans notre laboratoire (Dougier et al., 2006; Le Noir et al., 2017)) laisse suggérer que les deux allèles, productifs ou non, sont capables de former des boucles de chromatine afin de réaliser la CSR. De plus, il a également été démontré que la SHM se produit sur l'allèle non productif à un niveau similaire à celui de l'allèle productif (Delpy et al., 2004), suggérant là encore que la CSR et les boucles d'ADN associées sont présentes sur les deux allèles. Enfin, bien que débattu, il semblerait que la transcription des deux allèles *IgH* soient équivalentes ce qui sous-entend que les deux allèles peuvent être ciblés par AID (Delpy et al., 2003; Skok et al., 2001; Tinguely et al., 2012).

L'étude inédite de la formation de boucles de chromatine dans le modèle  $\Delta MARS_{E\mu}$  révèle une augmentation de toutes les fréquences de contact entre l'ancre et l'ensemble des points d'intérêts du locus *IgH* (*3'RR* et *S* donneur). Bien que ces données nécessitent d'être confirmées par un échantillonnage plus important et par l'utilisation d'un autre point de vue (ancre  $I\mu/S\mu$ ), nous émettons l'hypothèse qu'elles ne sont pas synonymes d'un plus grand nombre absolu de boucles de chromatine mais qu'elles révèlent une stabilité accrue de la grande boucle *3'RR-S $\mu$*  et donc du CSRC. Ceci pourrait expliquer l'augmentation des évènements de CSR observés dans ce modèle (Martin et al., 2023).

Décrit en 2019 par l'équipe de F. Alt, la formation de boucle au locus *IgH* est médiée par le mécanisme d'extrusion de boucle (X. Zhang et al., 2019). En temps normal, le complexe de la cohésine est chargé sur l'ADN par la protéine NIPBL puis dissocié de l'ADN grâce à l'action de WAPL. L'augmentation des fréquences de contact entre la *3'RR* et *S $\mu$*  observée dans le modèle  $\Delta MARS_{E\mu}$  nous conduit à suspecter que la stabilité des boucles est augmentée dans ce modèle. Cette stabilité accrue fait émerger l'hypothèse selon laquelle les régions *MARS $E\mu$*  constituent une plateforme de recrutement de facteurs nécessaires au mécanisme de déchargement du complexe de la cohésine de l'ADN. Pour tester cette hypothèse, nous envisageons de tester le recrutement des acteurs de l'extrusion de boucle dans des modèles  $\Delta MARS_{E\mu}$  en comparaison au *wt*. La technique de Cut&Run, plus sensible et connue pour avoir un « bruit de fond » inférieur à celui de la technique traditionnelle de CHIP, est en cours de mise au point dans l'équipe B-NATION afin notamment de répondre à la question du recrutement de WAPL et de la présence augmentée de la cohésine (de sa sous-unité Rad21 notamment) dans le modèle  $\Delta MARS_{E\mu}$ . Si notre hypothèse selon laquelle WAPL est recrutée par les *MARS $E\mu$*  est valide, nous devrions observer une diminution du recrutement de cette protéine avec en parallèle une accumulation du complexe de la cohésine au locus *IgH* en contexte KO.

D'autre part, nos données montrent une absence totale de boucle en contexte  $\Delta 3'RR$ , et ce même avec un point de vue localisé dans la région  $I\mu/S\mu$  (**Article n°1**). Nous émettons alors l'hypothèse qu'à l'inverse des  $MARs_{E\mu}$ , la  $3'RR$  constituerait le point d'entrée du complexe de la cohésine. Nous prévoyons donc de rechercher le recrutement de NIPBL, protéine responsable du chargement de la cohésine, par du Cut&Run, en contexte mutant et *wt*. Nous imaginons alors observer une diminution drastique du recrutement de NIPBL, et par conséquent de la cohésine, au locus *IgH* en contexte  $\Delta 3'RR$ . Ces données viendraient préciser le rôle des régions  $E\mu MARs$  et  $3'RR$  dans le mécanisme de l'extrusion de boucle, contribuant ainsi à mieux comprendre les processus sous-jacents à la réponse antigénique lymphocytaire B.

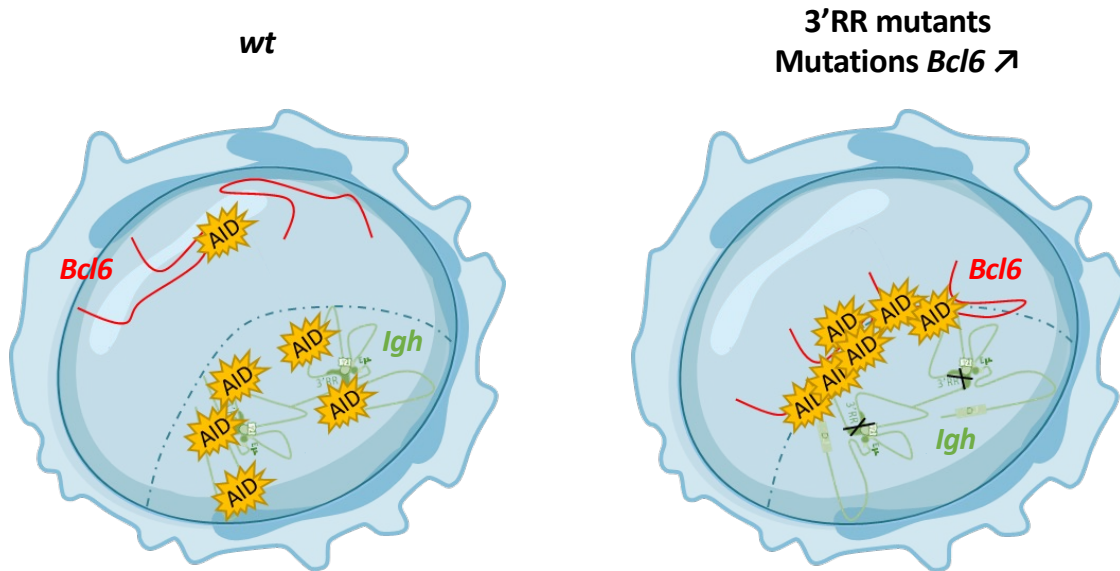
### ***3'RR, organisation et intégrité du génome de la cellule B***

Le second axe de mon projet de thèse portait sur l'étude de l'implication de la  $3'RR$  dans l'intégrité du génome du lymphocyte B. Nous nous sommes donc intéressés à l'introduction de mutations au sein de cibles illégitimes d'AID nommés gènes « AID off-target ». Selon l'équipe de Schatz, ces gènes sont classés en trois groupes distincts en fonction du ciblage par AID qu'ils subissent et des voies de réparations impliquées dans la prise en charge du mésappariement dU:dG induit par l'action d'AID (Liu et al., 2008).

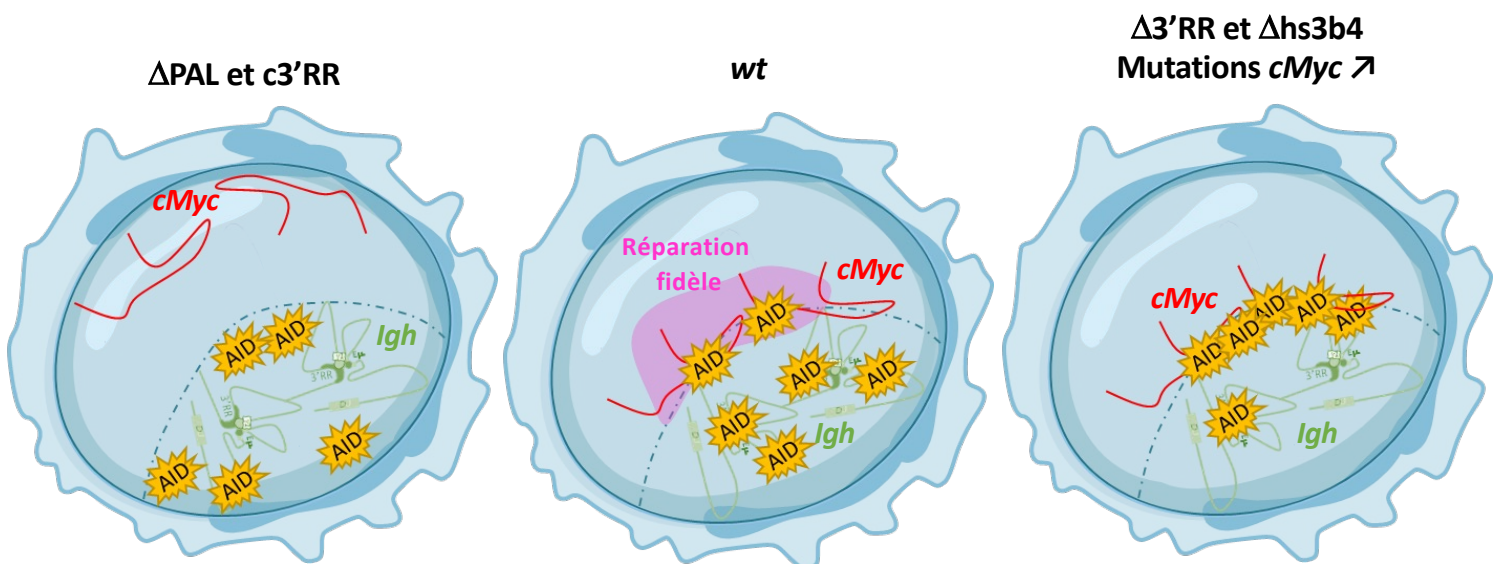
Nous avons sélectionné plusieurs gènes, majoritairement du groupe 1 incluant *Bcl6*, *Pim1* et *Cd83* ainsi que *cMyc*, qui fait quant à lui partie du groupe 2 (Liu et al., 2008). L'analyse des fréquences de mutations de ces différents gènes dans nos animaux mutants pour la  $3'RR$  a révélé des différences importantes qui dépendent soit du gène impliqué, soit du modèle mutant ; ces observations témoignent de toute la complexité du maintien de l'intégrité génomique de la cellule B (**Résultats préliminaires Partie II**). Néanmoins, l'ensemble des dérégulations observées possède un point commun : il semble qu'il n'y a pas de lien direct entre l'augmentation de la fréquence de mutation et la transcription primaire de ces gènes.

La proximité d'un gène avec les loci *IgH* est un facteur favorisant l'action de AID en dehors des gènes d'*Ig* (Rocha et al., 2012). Chez les modèles mutants pour la  $3'RR$ , nous avons donc étudié la distance entre les loci *Bcl6* ou *cMyc* et *IgH* ainsi que leur distance respective au centre du noyau. Notre étude révèle que les deux oncogènes ne sont pas soumis aux mêmes contraintes, un constat probablement lié à leur position nucléaire en condition normale : *Myc* est très fréquemment retrouvé à proximité d'*IgH* tandis que *Bcl6* est normalement localisé dans une autre zone du noyau (Osborne et al., 2007; Rocha et al., 2012). L'intron 1 de *Bcl6* est significativement plus muté dans l'ensemble des modèles mutants de la  $3'RR$  et nous observons une relocalisation de *Bcl6* plus proche d'*IgH* et du centre du noyau. Cette dernière observation fait écho à une étude dans laquelle il est montré que le gène de la





**Figure 36 : Localisation de *Bcl6* en contexte *wt* et *3'RR*-mutant dans des cellules B activées**  
 A gauche, en contexte *wt*, *Bcl6* (en rouge) est localisé dans une zone en périphérie du noyau et éloigné d'*IgH* (en vert), et est faiblement ciblé par AID (en jaune). A droite, en contexte *3'RR* mutant, *Bcl6* est situé plus au centre du noyau et plus proche d'*IgH*. Le ciblage de *Bcl6* par AID est plus important.



**Figure 37 : Localisation de *cMyc* en contexte *wt* et *3'RR*-mutant dans des cellules B activées**  
 Au centre, en contexte *wt*, *cMyc* (en rouge) est localisé proche d'*IgH* (en vert), subit l'action d'AID mais est réparé de façon fidèle (Liu et al, 2008). A gauche, dans les modèles  $\Delta$ PAL et *c3'RR*, *cMyc* serait relocalisé plus loin d'*IgH* et plus en périphérie. A droite, en contexte  $\Delta$ *3'RR* et  $\Delta$ *hs3b4*, *cMyc* reste à proximité d'*IgH* mais ne serait plus réparé de façon fidèle.

cycline D1, transloqué au locus *IGH* est relocalisé dans une zone plus centrale du noyau, à proximité du nucléole (Allinne et al., 2014). De même, une étude a rapporté que la périphérie nucléaire, notamment les domaines associés à la lamine, est une zone répressive de la SHM dans des lignées cellulaires B activées (Klymenko et al., 2018). Cependant, l'existence d'une relation de cause à effet entre position péri-nucléaire et introduction de mutation est débattue dans la littérature. En effet, une étude menée sur plusieurs pathologies malignes montre que les chromosomes plus fréquemment mutés se trouvent plus en périphérie du noyau (Smith et al., 2017).

Nous émettons l'hypothèse que, en l'absence des éléments régulateurs de la 3'RR, *Bcl6* se trouve relocalisé dans une zone centrale du noyau, riche en AID et prône à son action. Les mutants de la 3'RR ne recrutent pas AID au locus *IgH* de façon optimale (Saintamand et al., 2015), AID agirait de façon plus importante sur l'oncogène *Bcl6* nouvellement positionné à proximité de *IgH* (**Figure 36**). Un scénario compatible peut être envisagé dans les mutants présentant une augmentation de la fréquence de mutation de *cMyc* ( $\Delta 3'RR$  et  $\Delta hs3b4$ ) dans lesquels nous observons que les loci *cMyc* et *IgH* restent proches. A l'inverse, dans les mutants  $\Delta PAL$  et  $c3'RR$ , chez lesquels la fréquence de mutation de *cMyc* n'est pas augmentée, *cMyc* s'éloigne de *IgH* et se trouve repositionné plus en périphérie du noyau (**Figure 37**). Il semblerait que la proximité entre *IgH* et *Bcl6* ou *cMyc* ait donc un lien avec l'augmentation de leur mutation bien que ce lien de cause à effet « plus proche d'*IgH*, plus susceptible de subir l'action d'AID » soit discuté dans la littérature (Gramlich et al., 2012).

En plus de leur positionnement initial par rapport à *IgH*, le fait que ces gènes appartiennent à des groupes distincts selon la classification de Schatz (Liu et al., 2008) pourrait aussi expliquer les différentes contraintes que ces derniers subissent. *Bcl6* est un gène du groupe 1, c'est-à-dire qu'il est fortement ciblé par AID et n'est pas réparé de façon fidèle. En revanche, *cMyc* est un gène du 2<sup>ème</sup> groupe, il subit l'action d'AID mais, au contraire de *Bcl6*, les dommages collatéraux qu'il subit sont très fidèlement réparés. Les données que nous avons obtenues suggèrent que *Bcl6* et *cMyc* subissent des mécanismes de déamination et de réparation différents. Alors que *Bcl6* accumule « normalement » des mutations en condition *wt*, son rapprochement avec *IgH* en l'absence de la 3'RR induirait une augmentation de son ciblage par AID et naturellement une surcharge de mutations. Pour *cMyc*, nous pouvons imaginer que la délétion de l'enhancer *hs4* (enhancer communément absent dans les modèles  $\Delta 3'RR$  et  $\Delta hs3b4$  mais présent dans les modèles  $\Delta PAL$  et  $c3'RR$ ) conduit à une modification du recrutement des voies de réparation de l'ADN dans la zone nucléaire des gènes d'Ig (à la faveur de la réparation infidèle), conduisant ainsi à la persistance des mutations en aval de l'action d'AID (**Figure 37**).

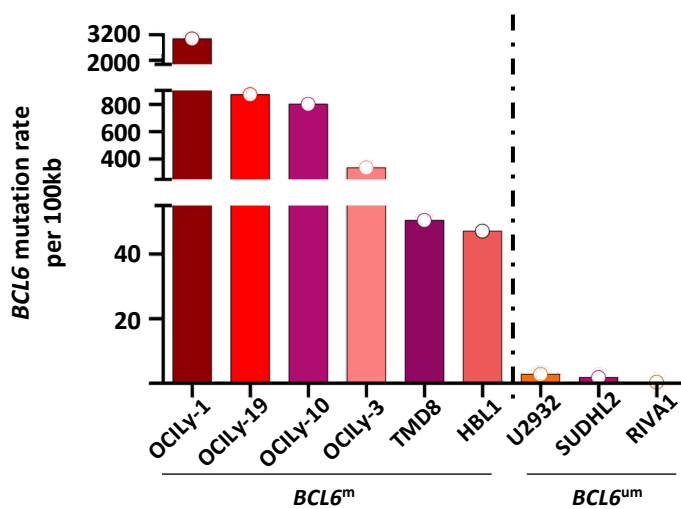
Toutefois, il est important de noter que nos données de localisation de *Bcl6* et *cMyc* par rapport à *IgH* ont été obtenues dans des cellules B activées *in vitro* et constituent donc des données préliminaires qui restent à confirmer dans des cellules B de GC. Une telle exploration se justifie d'autant plus que les

cellules B voient leur architecture nucléaire et génomique remaniée pendant la formation du GC (Azagra et al., 2020; Scourzic et al., 2021).

Pour confirmer ces hypothèses, il serait pertinent de réaliser la mesure des distances par FISH-3D pour les deux autres gènes que nous avons étudié, à savoir *Pim1* et *Cd83*. Une autre possibilité de vérifier nos hypothèses serait d'explorer à l'échelle du noyau, par exemple en réalisant de l'immuno-FISH, la présence des gènes « AID off-target » dans des zones nucléaires riches en protéines actrices des voies de réparation ou riches en AID. Ce dernier point semble complexe à mettre en œuvre de par la régulation très dynamique d'AID qui est constamment importée et exportée du noyau des cellules B activées (Ito et al., 2004; Rada et al., 2002). Pour aller plus loin, des expériences Cut&Run pourraient aussi être envisagées afin de rechercher un éventuel recrutement des protéines actrices des voies de réparations à l'échelle de l'ADN, sur les gènes « AID off target ».

Il serait également intéressant d'explorer si les différences observées entre *Bcl6* et *cMyc* sont communes à l'ensemble des gènes du groupe 1 et du groupe 2 en étendant l'étude à d'autres gènes inclus dans chacun de ces deux groupes. Nous pourrions également envisager de réaliser une stimulation antigénique des souris (immunisation) afin de déterminer si une réponse immune B à un antigène spécifique provoque, dans des cellules de GC induites, des remaniements nucléaires comparables à ceux observés dans les cellules de GC spontanés issus des plaques de Peyer. Sur le plus long terme, l'exploration des voies de réparation en aval de la déamination par AID pourrait être explorée dans les modèles mutants de la 3'RR en réalisant des croisements en fond génétique déficient pour les voies BER (*Ung* KO) et/ou MMR (*Msh2/Msh6* KO) comme cela a précédemment été réalisé pour le modèle  $\Delta$ MARSE $\mu$  et l'étude de la SHM au locus *IgH* (Martin et al., 2023). Cette stratégie permettrait de déterminer si ces voies de réparation sont incriminées dans l'augmentation des mutations de *cMyc*.

Enfin, la fonction protectrice de la 3'RR semble être cantonnée aux oncogènes impliqués dans les hémopathies B. En effet, il est aisé de constater que nos modèles mutants n'affectent aucunement la fréquence de mutation de *Cd83* qui n'est pas un oncogène B. A l'inverse, les trois autres gènes de notre étude (*Bcl6*, *cMyc* et *Pim1*), qui sont des oncogènes fréquemment impliqués dans les lymphomes B (Pasqualucci and Dalla-Favera, 2018), voient leurs fréquences de mutations augmenter dans certains modèles mutants de la 3'RR. Cette observation est corrélée avec la littérature qui souligne depuis longtemps l'implication de la 3'RR dans la lymphomagenèse. En effet, de nombreuses translocations positionnent des oncogènes sous le contrôle de la 3'RR, conduisant à leur expression anormalement élevée et conférant souvent des propriétés prolifératives ou anti-apoptotiques aux cellules lymphomateuses (Duan et al., 2007; Gostissa et al., 2009; Vincent-Fabert et al., 2010a). Parmi les oncogènes dérégulés dans les hémopathies malignes B humaines, nous pouvons citer *BCL6*, *CMYC*, *PIM1*, *BCL2*, *PAX5*...



**Figure 38 : Taux de mutation de *Bcl6* dans des lignées cellulaires de DLBCL.**

Les lignées humaines de DLBCL ont été divisés en deux groupes :  $BCL6^m$  (OCILy-1, OCILy-19, OCILy-10, OCILy-3, TMD8 et HBL1) and  $BCL6^{um}$  (U2932, SUDHL2 et RIVA1).

Sous-type GCB : OCILy-1 et OCILy-19

Sous-type ABC : SUDHL2, RIVA1, OCILy-3, U2932, HBL1, OCILy-10 et TMD8

Les oncogènes *BCL2*, *CMYC* ou encore *PIM1* sont fréquemment transloqués au locus *IGH* dans le lymphome diffus à grande cellules B (DLBCL).

*BCL6*, quant à lui, est muté dans près de 75% des cas de DLBCL (Pasqualucci et al., 1998). Le DLBCL est un des lymphomes non hodgkiniens le plus fréquent chez l'adulte (Swerdlow et al., 2016). C'est un lymphome agressif caractérisé par la prolifération anarchique de cellules B du GC (pour le sous-type GCB) ou post-GC (pour le sous-type ABC). Les DLBCL présentent une diversité génétique considérable et certaines altérations moléculaires jouent un rôle clé dans le développement et la progression du lymphome (Pasqualucci and Dalla-Favera, 2018). En conditions physiologiques, *BCL6* est un facteur de transcription essentiel à la mise en place du GC, normalement réprimé pour permettre aux cellules B de continuer à se différencier en plasmocytes ou cellules B mémoires (Cattoretti et al., 1995; Shaffer et al., 2002). En revanche, dans le contexte du DLBCL, des mutations sont souvent introduites dans la partie 5' du gène, entraînant une altération de sa régulation (Khodabakhshi et al., 2012; Pasqualucci et al., 2003). Il est notamment fait état que, l'introduction de mutations dans *BCL6*, et surtout au niveau de son super-enhancer, est une caractéristique unique et propre au DLBCL (Bal et al., 2022). Ces mutations peuvent alors conduire à une expression persistante de ce gène, perturbant le processus normal de différenciation des cellules B et entraînant leur prolifération excessive.

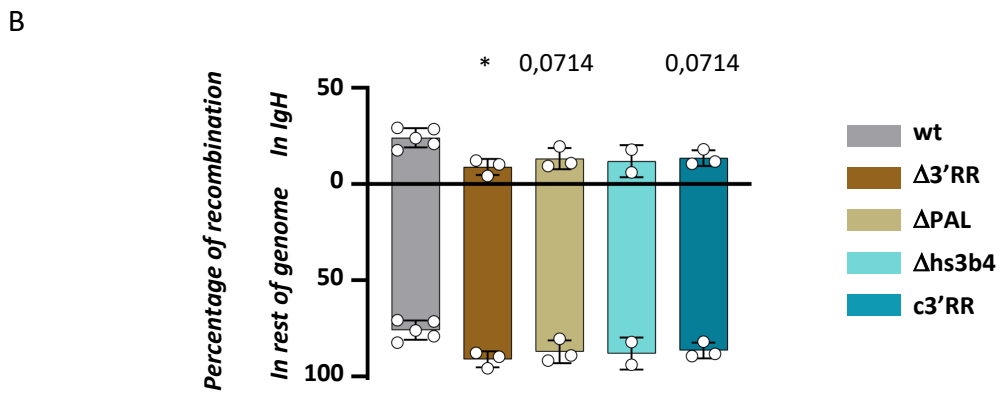
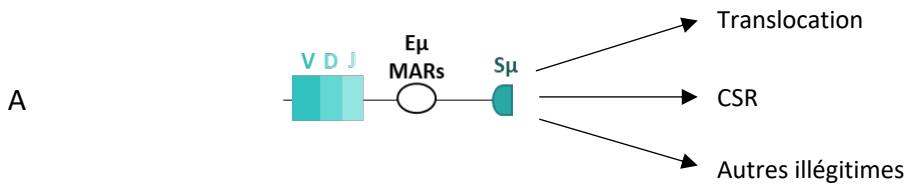
Suite à nos résultats préliminaires (**Résultats préliminaires Partie II.1**), nous avons émis l'hypothèse selon laquelle les loci *BCL6* et *IGH* seraient plus proches l'un de l'autre chez les patients présentant des mutations au sein de *BCL6*. Pour explorer cette hypothèse, nous avons commencé à étudier la fréquence de mutation dans différentes lignées cellulaires humaines de DLBCL, dans le premier intron de l'oncogène *BCL6* humain, région décrite comme fréquemment mutée (Iqbal et al., 2007). Nous avons ainsi quantifié la fréquence de mutations dans 2 lignées de sous-type GCB (OCILy-1 et OCILy-19) et 7 lignées de sous-type ABC (SUDHL2, RIVA-1, OCILy-3, U2932, HBL1, OCILy-10 et TMD8). Ces lignées ont été sélectionnées car elles ne présentent pas de translocations des loci *IGH* ou *BCL6*. De façon attendue, la majorité des lignées (6 sur 9, soit 66%) montrent une fréquence de mutation exacerbée de *BCL6* (**Figure 38**), ce qui est en accord avec le taux de mutation de *BCL6* dans les DLBCL au diagnostic. Nous avons alors constitué deux groupes selon le statut mutationnel de *BCL6* : d'une part le groupe *BCL6<sup>m</sup>* (m pour « mutated ») incluant OCILy-1, OCILy-19, OCILy-10, OCILy-3, TMD8 et HBL1 et d'autre part le groupe *BCL6<sup>um</sup>* (um pour « unmutated ») incluant U2932, SUDHL2 et RIVA1.

Nous avons en parallèle également initié l'étude du positionnement relatif de *BCL6* par rapport à *IGH* par FISH-3D, mais de nombreuses contraintes techniques inattendues ont retardé l'obtention de résultats ; cette partie est donc toujours en cours de réalisation. Cette étude positionnelle nous permettra de déterminer si les lignées du groupe *BCL6<sup>m</sup>* ont leur loci *IGH* et *BCL6* plus proche l'un de l'autre que ceux des lignées du groupe *BCL6<sup>um</sup>*. Nous prévoyons d'étendre cette étude à des échantillons de patients au diagnostic, dont la sélection est en cours dans la tumorothèque du service

de cytogénétique du centre Henri Becquerel à Rouen (Dominique Penther) et en prospection avec le CHU de Limoges (Nathalie Gachard). Nous prévoyons d'analyser la fréquence de mutation de *BCL6* et de procéder à la répartition des patients en deux groupes distincts (*BCL6<sup>m</sup>* vs *BCL6<sup>um</sup>*, comme précédemment réalisé pour les lignées de DLBCL) selon la fréquence de mutation de *BCL6*. Nous réaliserons en parallèle l'étude du positionnement relatif de *BCL6* vis-à-vis d'*IGH* afin de déterminer si les observations faites chez la souris sont transposables à la pathologie.

Enfin, nous avons également demandé l'accès à des données de *Whole Genome Sequencing* (WGS) issues de 73 patients DLBCL recrutés au Canada et incluant le tissu sain et le tissu tumoral pour chacun d'entre eux (Arthur et al., 2018). Cette demande est en cours d'examen au comité d'éthique du CHU de Limoges et auprès du *British Columbia Cancer Agency*. Nous souhaitons dans un premier temps réaliser deux groupes de patients (*BCL6<sup>m</sup>* vs *BCL6<sup>um</sup>*) en analysant la fréquence de mutation de *BCL6*; ceci est rendu possible car nous disposons des données de séquençage du tissu sain de chaque patient qui permettra de distinguer les « vraies » mutations des polymorphismes. Afin de valider, dans un contexte pathologique nos observations initiales faites chez les modèles murins mutants pour la 3'RR, nous prévoyons ensuite d'analyser l'intégrité des régions 3'RR humaines dans cette cohorte de DLBCL en fonction du statut mutationnel de *BCL6*. Ce dernier point représente un réel challenge de par la structure très répétée de cette région *cis*-régulatrice. Cette analyse bio-informatique sera réalisée en collaboration avec Vera Pancaldi (CRCT Toulouse).

Si l'intégrité des 3'RR n'est pas modifiée ou si leur analyse est impossible nous pourrions aussi envisager de récupérer des données de RNA-Seq sur lesquelles nous pourrions analyser l'activité de la 3'RR, notamment par la présence d'« eARN » (e pour enhancer). Il semblerait que les eARN, et plus largement les longs ARN non codants (catégorie à laquelle les eARN appartiennent), puissent jouer des rôles importants à la fois dans l'organisation du génome et son intégrité (Laffleur et al., 2021; Rothschild and Basu, 2017). Ces ARN seraient en effet capables d'agir aussi bien en *cis* qu'en *trans* sur d'autres gènes localisés sur le même chromosome, voir même sur d'autres chromosomes en « guidant » certaines protéines de remodelage de la chromatine (Guttman and Rinn, 2012; Khalil et al., 2009). De plus, en *cis*, la présence de l'eARN facilite les contacts à longue distance entre promoteur et enhancer notamment (Kim et al., 2015). Nous pouvons alors imaginer, bien qu'ayant conservé son intégrité, la 3'RR voit son activité modifiée, entraînant une diminution de la présence d'eARN et donc une perte de contact promoteur-enhancer. L'interactant (partenaire de la 3'RR) se retrouvant alors « seul » se mettrait à interagir avec des cibles « AID off-target » comme *Bcl6* par exemple. Explorer cette piste pourrait donc s'avérer intéressant afin de mieux comprendre l'implication de la 3'RR dans les phénomènes précédemment décrits. L'hypothèse précédemment formulée semble d'autant plus plausible qu'il est de plus en plus évident que les enhancers, et surtout les super-enhancers, tels que la 3'RR, sont impliqués dans l'organisation et le maintien de l'intégrité du génome et contribuent à des



**Figure 39 : Répartition des évènements de recombinaisons dans *IgH* vs. le reste du génome.**  
 p-value was determined with two tailed Mann Whitney test, significant differences were indicated by:  
 \* P<0.05; \*\*P<0.01; \*\*\* P<0.001; \*\*\*\* P<0,0001 and error bars represent SEM

fonctions primordiales notamment dans le cadre des hémopathies malignes (Dębek and Juszczynski, 2022).

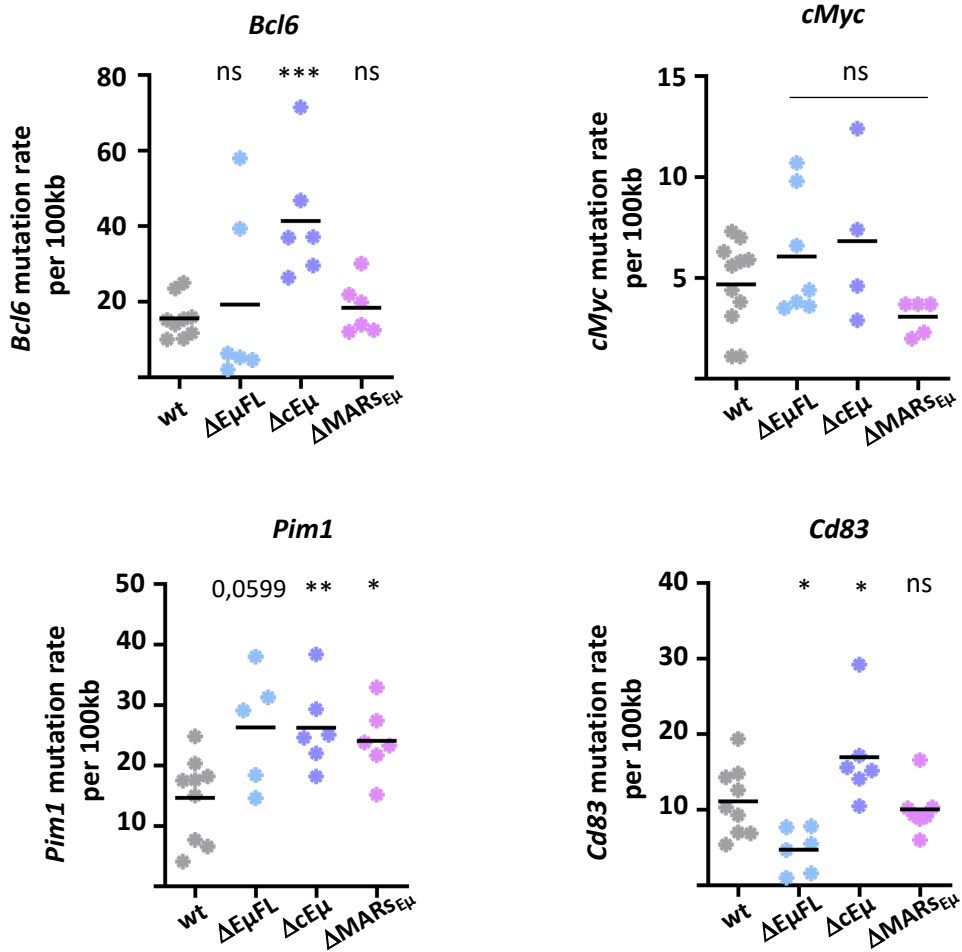
Pour explorer plus en détails le lien entre la 3'RR, la localisation relative des oncogènes par rapport à *IGH* et l'introduction de mutations, il pourrait être intéressant d'endommager l'intégrité de la 3'RR par technique de CRISPR/Cas9 dans des lignées cellulaires B ou des cellules B primaires humaines et d'étudier les conséquences sur la position de *BCL6* et *cMYC* par rapport à *IGH*. L'utilisation d'une lignée capable d'hypermuation somatique, comme la lignée Ramos par exemple, pourrait constituer un outil utile pour comprendre le lien de cause à effet entre la 3'RR, la mutation de gènes cibles illégitimes et la position dans le noyau. L'établissement d'un tel modèle, utilisé par ailleurs pour étudier l'implication des enhanceurs ainsi que le rôle de la structure du génome dans les processus de SHM (Komori et al., 2006; Senigl et al., 2019b), s'avérerait particulièrement utile pour confirmer l'implication de la 3'RR dans le positionnement des oncogènes et l'intégrité du génome.

Pour comprendre si l'activité de la 3'RR joue un rôle dans ces processus, nous pourrions également envisager de traiter nos cellules avec certains agents chimiques capables de moduler l'activité de la 3'RR. Cette méthode permettrait également de faire le lien avec les eARN, puisque certains de ces agents, notamment ceux liant le récepteur AhR (*Aryl hydrocarbon Receptor*), comme la dioxine, sont capables de moduler l'activité de la 3'RR et donc la présence d'eARN (D'Addabbo et al., 2023; Henseler et al., 2009; Salisbury and Sulentic, 2015).

### ***Pour conclure sur la 3'RR***

Mes travaux mettent en avant le rôle des core enhanceurs dans la pré-conformation du noyau des cellules B au repos afin de garantir la survenue d'une CSR délétionnelle lors de l'activation. Nous aurions alors pu suggérer que la seule présence des core enhanceurs, sans la structure quasi-palindromique de la 3'RR, permettrait de maintenir une organisation nucléaire suffisante au maintien de l'intégrité génomique. Pourtant, l'étude de l'introduction de mutations au sein de *Bcl6* révèle que ce n'est pas le cas. De même, si l'on s'intéresse à la proportion de recombinaisons entre le locus *IgH* (depuis *Sμ*, **Figure 39A**) et le reste du génome par LAM-HTGTS (*Linear Amplification Mediated – High Throughput Genome-wide Translocation Sequencing*, technique maîtrisée au laboratoire) (Hu et al., 2016), les travaux que j'ai initiés montrent que ces recombinaisons « hors locus » sont augmentées dans tous les modèles mutants de la 3'RR (**Figure 39B**). A l'avenir nous envisageons d'étudier la structure des jonctions de ces recombinaisons entre *Sμ* et le reste du génome afin de déterminer si les voies de réparation impliquées diffèrent de celles recrutées lors de la CSR, dans un contexte *wt* et chez les différents mutants de la 3'RR.





**Figure 40 : Fréquence de mutation des gènes « AID off target » dans les mutants *EμMARs*.** Fréquence de mutation des gènes «AID off target» chez des animaux âgés de 10 semaines, déterminée avec l’outil DeMinEr  
 Les p-value sont déterminées avec un test Mann Whitney, les différences significatives sont indiquées par : \* P<0.05; \*\*P<0.01; \*\*\* P<0.001; \*\*\*\* P<0,0001

Ces données laissent supposer une dualité fonctionnelle de la 3'RR dans le maintien de l'intégrité du génome : bien que les core enhancers sont suffisants et nécessaires pour réguler l'adressage nucléaire du locus *IgH*, la structure palindromique de la 3'RR semble indispensable pour maintenir l'intégrité du génome.

### ***Pour aller plus loin avec la région EμMARs***

Des travaux similaires avec les modèles mutants de la région EμMARs ont été initiés pendant ma thèse. L'étude des mutations a pour le moment été réalisée chez des animaux âgés de 10 semaines et l'échantillonnage reste insuffisant (**Figure 40**). À l'image de ce que nous observons chez les mutants de la 3'RR, il semble que les modifications de fréquences de mutations dépendent des gènes et des modèles. Nous prévoyons donc de réitérer les expériences chez des souris de 24 semaines. Ceci nous permettra dans un premier temps d'analyser la cinétique d'accumulation de mutations dans les différents gènes « AID off target » chez les mutants EμMARs puis de comparer ces cinétiques avec celles obtenues pour les mutants de la 3'RR. Nous envisageons ensuite d'étudier également le niveau de transcription primaire et la position de chacun de ces gènes dans le noyau des cellules B issues des différents modèles murins.

### ***Pour conclure de façon générale***

De façon plus globale, l'ensemble des modèles mutants pour la 3'RR et pour la région EμMARs ont démontré l'implication de ces deux régions dans la régulation du locus *IgH* à la fois à l'échelle nucléaire et au niveau supranucléosomale. Nos résultats révèlent le rôle de la région 3'RR en tant que gardien du génome du lymphocyte B et suggèrent également une implication de la région EμMARs dans le maintien de l'intégrité génomique. Pour aller plus loin et explorer la fonction des régions EμMARs et 3'RR à une échelle combinée noyau-génome, nous envisageons de réaliser des expériences de GPseq (*Genomic loci Positioning by sequencing*) (Girelli et al., 2020). Cette technique permet d'étudier la position radiale de l'ensemble des gènes en alliant une réaction de restriction enzymatique graduelle suivie du séquençage haut débit des fragments générés. De cette façon, nous pourrions comparer la distribution des gènes d'Ig (*IgH*, *Igκ*, *Igλ*, *IgJ*) ainsi que des gènes AID off-target (*Bcl6*, *cMyc*, *Pim1*, *Cd83*...) dans le noyau des cellules B et comparer ces positions dans l'ensemble de nos modèles murins par rapport au contexte *wt*.

## *Références Bibliographiques*

---

## Références bibliographiques

---

### - A -

- Adolfsson, J., Borge, O.J., Bryder, D., Theilgaard-Mönch, K., Astrand-Grundström, I., Sitnicka, E., Sasaki, Y., Jacobsen, S.E., 2001. Upregulation of Flt3 expression within the bone marrow Lin(-)Sca1(+)c-kit(+) stem cell compartment is accompanied by loss of self-renewal capacity. *Immunity* 15, 659–669. [https://doi.org/10.1016/s1074-7613\(01\)00220-5](https://doi.org/10.1016/s1074-7613(01)00220-5)
- Alessandrini, A., Desiderio, S.V., 1991. Coordination of immunoglobulin DJH transcription and D-to-JH rearrangement by promoter-enhancer approximation. *Mol. Cell. Biol.* 11, 2096–2107. <https://doi.org/10.1128/mcb.11.4.2096-2107.1991>
- Allen, C.D.C., Ansel, K.M., Low, C., Lesley, R., Tamamura, H., Fujii, N., Cyster, J.G., 2004. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat. Immunol.* 5, 943–952. <https://doi.org/10.1038/ni1100>
- Allinne, J., Pichugin, A., Iarovaia, O., Klibi, M., Barat, A., Zlotek-Zlotkiewicz, E., Markozashvili, D., Petrova, N., Camara-Clayette, V., Ioudinkova, E., Wiels, J., Razin, S.V., Ribrag, V., Lipinski, M., Vassetzky, Y.S., 2014. Perinucleolar relocalization and nucleolin as crucial events in the transcriptional activation of key genes in mantle cell lymphoma. *Blood* 123, 2044–2053. <https://doi.org/10.1182/blood-2013-06-510511>
- Allman, D., Lindsley, R.C., DeMuth, W., Rudd, K., Shinton, S.A., Hardy, R.R., 2001. Resolution of three nonproliferative immature splenic B cell subsets reveals multiple selection points during peripheral B cell maturation. *J. Immunol. Baltim. Md 1950* 167, 6834–6840. <https://doi.org/10.4049/jimmunol.167.12.6834>
- Allman, D.M., Ferguson, S.E., Cancro, M.P., 1992. Peripheral B cell maturation. I. Immature peripheral B cells in adults are heat-stable antigenhi and exhibit unique signaling characteristics. *J. Immunol. Baltim. Md 1950* 149, 2533–2540.
- Alt, F.W., Oltz, E.M., Young, F., Gorman, J., Taccioli, G., Chen, J., 1992. VDJ recombination. *Immunol. Today* 13, 306–314. [https://doi.org/10.1016/0167-5699\(92\)90043-7](https://doi.org/10.1016/0167-5699(92)90043-7)
- Alt, F.W., Yancopoulos, G.D., Blackwell, T.K., Wood, C., Thomas, E., Boss, M., Coffman, R., Rosenberg, N., Tonegawa, S., Baltimore, D., 1984. Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J.* 3, 1209–1219. <https://doi.org/10.1002/j.1460-2075.1984.tb01955.x>
- Angelin-Duclos, C., Cattoretto, G., Lin, K.-I., Calame, K., 2000. Commitment of B Lymphocytes to a Plasma Cell Fate Is Associated with Blimp-1 Expression In Vivo. *J. Immunol.* 165, 5462–5471. <https://doi.org/10.4049/jimmunol.165.10.5462>
- Arakawa, H., Shimizu, T., Takeda, S., 1996. Re-evaluation of the probabilities for productive arrangements on the kappa and lambda loci. *Int. Immunol.* 8, 91–99. <https://doi.org/10.1093/intimm/8.1.91>
- Arnould, C., Rocher, V., Finoux, A.-L., Clouaire, T., Li, K., Zhou, F., Caron, P., Mangeot, P.E., Ricci, E.P., Mourad, R., Haber, J.E., Noordermeer, D., Legube, G., 2021. Loop extrusion as a mechanism for formation of DNA damage repair foci. *Nature* 590, 660–665. <https://doi.org/10.1038/s41586-021-03193-z>
- Arthur, S.E., Jiang, A., Grande, B.M., Alcaide, M., Cojocar, R., Rushton, C.K., Mottok, A., Hilton, L.K., Lat, P.K., Zhao, E.Y., Culibrk, L., Ennishi, D., Jessa, S., Chong, L., Thomas, N., Pararajalingam, P., Meissner, B., Boyle, M., Davidson, J., Bushell, K.R., Lai, D., Farinha, P., Slack, G.W., Morin, G.B., Shah, S., Sen, D., Jones, S.J.M., Mungall, A.J., Gascoyne, R.D., Audas, T.E., Unrau, P., Marra, M.A., Connors, J.M., Steidl, C., Scott, D.W., Morin, R.D., 2018. Genome-wide discovery of somatic regulatory variants in diffuse large B-cell lymphoma. *Nat. Commun.* 9, 4001. <https://doi.org/10.1038/s41467-018-06354-3>
- Azagra, A., Marina-Zárate, E., Ramiro, A.R., Javierre, B.M., Parra, M., 2020. From Loops to Looks: Transcription Factors and Chromatin Organization Shaping Terminal B Cell Differentiation. *Trends Immunol.* 41, 46–60. <https://doi.org/10.1016/j.it.2019.11.006>

### - B -

- Ba, Z., Lou, J., Ye, A.Y., Dai, H.-Q., Dring, E.W., Lin, S.G., Jain, S., Kyritsis, N., Kieffer-Kwon, K.-R., Casellas, R., Alt, F.W., 2020. CTCF orchestrates long-range cohesin-driven V(D)J recombinational scanning. *Nature* 586, 305–310. <https://doi.org/10.1038/s41586-020-2578-0>

- Bahjat, M., Guikema, J.E.J., 2017. The Complex Interplay between DNA Injury and Repair in Enzymatically Induced Mutagenesis and DNA Damage in B Lymphocytes. *Int. J. Mol. Sci.* 18, 1876. <https://doi.org/10.3390/ijms18091876>
- Bakhshi, A., Wright, J.J., Graninger, W., Seto, M., Owens, J., Cossman, J., Jensen, J.P., Goldman, P., Korsmeyer, S.J., 1987. Mechanism of the t(14;18) chromosomal translocation: structural analysis of both derivative 14 and 18 reciprocal partners. *Proc. Natl. Acad. Sci. U. S. A.* 84, 2396–2400. <https://doi.org/10.1073/pnas.84.8.2396>
- Bal, E., Kumar, R., Hadigol, M., Holmes, A.B., Hilton, L.K., Loh, J.W., Dreval, K., Wong, J.C.H., Vlasevska, S., Corinaldesi, C., Soni, R.K., Basso, K., Morin, R.D., Khiabani, H., Pasqualucci, L., Dalla-Favera, R., 2022. Super-enhancer hypermutation alters oncogene expression in B cell lymphoma. *Nature* 607, 808–815. <https://doi.org/10.1038/s41586-022-04906-8>
- Balázs, M., Martin, F., Zhou, T., Kearney, J.F., 2002. Blood Dendritic Cells Interact with Splenic Marginal Zone B Cells to Initiate T-Independent Immune Responses. *Immunity* 17, 341–352. [https://doi.org/10.1016/S1074-7613\(02\)00389-8](https://doi.org/10.1016/S1074-7613(02)00389-8)
- Banerji, J., Olson, L., Schaffner, W., 1983. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* 33, 729–740. [https://doi.org/10.1016/0092-8674\(83\)90015-6](https://doi.org/10.1016/0092-8674(83)90015-6)
- Barajas-Mora, E.M., Kleiman, E., Xu, J., Carrico, N.C., Lu, H., Oltz, E.M., Murre, C., Feeney, A.J., 2019. A B-Cell-Specific Enhancer Orchestrates Nuclear Architecture to Generate a Diverse Antigen Receptor Repertoire. *Mol. Cell* 73, 48-60.e5. <https://doi.org/10.1016/j.molcel.2018.10.013>
- Barajas-Mora, E.M., Lee, L., Lu, H., Valderrama, J.A., Bjanes, E., Nizet, V., Feeney, A.J., Hu, M., Murre, C., 2023. Enhancer-instructed epigenetic landscape and chromatin compartmentalization dictate a primary antibody repertoire protective against specific bacterial pathogens. *Nat. Immunol.* 24, 320–336. <https://doi.org/10.1038/s41590-022-01402-z>
- Baranello, L., Kouzine, F., Levens, D., 2013. DNA Topoisomerases: Beyond the standard role. *Transcription* 4, 232–237. <https://doi.org/10.4161/trns.26598>
- Barreto, V., Reina-San-Martin, B., Ramiro, A.R., McBride, K.M., Nussenzweig, M.C., 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)
- Basso, K., Dalla-Favera, R., 2015. Germinal centres and B cell lymphomagenesis. *Nat. Rev. Immunol.* 15, 172–184. <https://doi.org/10.1038/nri3814>
- Basu, U., Meng, F.-L., Keim, C., Grinstead, V., Pefanis, E., Eccleston, J., Zhang, T., Myers, D., Wasserman, C.R., Wesemann, D.R., Janusz, K., Gregory, R.I., Deng, H., Lima, C.D., Alt, F.W., 2011. The RNA Exosome Targets the AID Cytidine Deaminase to Both Strands of Transcribed Duplex DNA Substrates. *Cell* 144, 353–363. <https://doi.org/10.1016/j.cell.2011.01.001>
- Baumgarth, N., 2017. A Hard(y) Look at B-1 Cell Development and Function. *J. Immunol. Baltim. Md* 1950 199, 3387–3394. <https://doi.org/10.4049/jimmunol.1700943>
- Baumgarth, N., 2011. The double life of a B-1 cell: self-reactivity selects for protective effector functions. *Nat. Rev. Immunol.* 11, 34–46. <https://doi.org/10.1038/nri2901>
- Bemark, M., Neuberger, M.S., 2003. By-products of immunoglobulin somatic hypermutation. *Genes Chromosomes Cancer* 38, 32–39. <https://doi.org/10.1002/gcc.10241>
- Bergsagel, P.L., Chesi, M., Nardini, E., Brents, L.A., Kirby, S.L., Kuehl, W.M., 1996. Promiscuous translocations into immunoglobulin heavy chain switch regions in multiple myeloma. *Proc. Natl. Acad. Sci. U. S. A.* 93, 13931–13936. <https://doi.org/10.1073/pnas.93.24.13931>
- Blagodatski, A., Batrak, V., Schmidl, S., Schoetz, U., Caldwell, R.B., Arakawa, H., Buerstedde, J.-M., 2009. A cis-Acting Diversification Activator Both Necessary and Sufficient for AID-Mediated Hypermutation. *PLoS Genet.* 5, e1000332. <https://doi.org/10.1371/journal.pgen.1000332>
- Blasquez, V.C., Sperry, A.O., Cockerill, P.N., Garrard, W.T., 1989. Protein:DNA interactions at chromosomal loop attachment sites. *Genome* 31, 503–509. <https://doi.org/10.1139/g89-098>

Boboila, C., Alt, F.W., Schwer, B., 2012. Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. *Adv. Immunol.* 116, 1–49. <https://doi.org/10.1016/B978-0-12-394300-2.00001-6>

Boboila, C., Yan, C., Wesemann, D.R., Jankovic, M., Wang, J.H., Manis, J., Nussenzweig, A., Nussenzweig, M., Alt, F.W., 2010. Alternative end-joining catalyzes class switch recombination in the absence of both Ku70 and DNA ligase 4. *J. Exp. Med.* 207, 417–427. <https://doi.org/10.1084/jem.20092449>

Bolland, D.J., Wood, A.L., Afshar, R., Featherstone, K., Oltz, E.M., Corcoran, A.E., 2007. Antisense intergenic transcription precedes Igh D-to-J recombination and is controlled by the intronic enhancer Emu. *Mol. Cell. Biol.* 27, 5523–5533. <https://doi.org/10.1128/MCB.02407-06>

Bolland, D.J., Wood, A.L., Johnston, C.M., Bunting, S.F., Morgan, G., Chakalova, L., Fraser, P.J., Corcoran, A.E., 2004. Antisense intergenic transcription in V(D)J recombination. *Nat. Immunol.* 5, 630–637. <https://doi.org/10.1038/ni1068>

Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G.L., Lubling, Y., Xu, X., Lv, X., Hugnot, J.-P., Tanay, A., Cavalli, G., 2017. Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* 171, 557–572.e24. <https://doi.org/10.1016/j.cell.2017.09.043>

Bransteitter, R., Pham, P., Scharff, M.D., Goodman, M.F., 2003. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl. Acad. Sci.* 100, 4102–4107. <https://doi.org/10.1073/pnas.0730835100>

Bredemeyer, A.L., Sharma, G.G., Huang, C.-Y., Helmink, B.A., Walker, L.M., Khor, K.C., Nuskey, B., Sullivan, K.E., Pandita, T.K., Bassing, C.H., Sleckman, B.P., 2006. ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. *Nature* 442, 466–470. <https://doi.org/10.1038/nature04866>

Bruzeau, C., Cook-Moreau, J., Pinaud, E., Le Noir, S., 2022. Contribution of Immunoglobulin Enhancers to B Cell Nuclear Organization. *Front. Immunol.* 13, 877930. <https://doi.org/10.3389/fimmu.2022.877930>

Bruzeau, C., Moreau, J., Le Noir, S., Pinaud, E., 2021. Panorama of stepwise involvement of the IgH 3' regulatory region in murine B cells. *Adv. Immunol.* 149, 95–114. <https://doi.org/10.1016/bs.ai.2021.03.004>

Bruzeau, C., Pollet, J., Thomas, M., Ba, Z., Roulois, D., Pinaud, E., Le Noir, S., 2023. Core enhancers of the 3'RR optimize IgH nuclear position and loop conformation for oriented CSR (preprint). *Immunology*. <https://doi.org/10.1101/2023.07.11.548507>

Buchanan, K.L., Smith, E.A., Dou, S., Corcoran, L.M., Webb, C.F., 1997. Family-specific differences in transcription efficiency of Ig heavy chain promoters. *J. Immunol. Baltim. Md 1950* 159, 1247–1254.

Busslinger, G.A., Stocsits, R.R., van der Lelij, P., Axelsson, E., Tedeschi, A., Galjart, N., Peters, J.-M., 2017. Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* 544, 503–507. <https://doi.org/10.1038/nature22063>

Busslinger, M., 2004. Transcriptional control of early B cell development. *Annu. Rev. Immunol.* 22, 55–79. <https://doi.org/10.1146/annurev.immunol.22.012703.104807>

- C -

Calado, D.P., Sasaki, Y., Godinho, S.A., Pellerin, A., Köchert, K., Sleckman, B.P., de Alborán, I.M., Janz, M., Rodig, S., Rajewsky, K., 2012. The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat. Immunol.* 13, 1092–1100. <https://doi.org/10.1038/ni.2418>

Calame, K.L., Lin, K.-I., Tunyaplin, C., 2003. Regulatory mechanisms that determine the development and function of plasma cells. *Annu. Rev. Immunol.* 21, 205–230. <https://doi.org/10.1146/annurev.immunol.21.120601.141138>

Carson, S., Wu, G.E., 1989. A linkage map of the mouse immunoglobulin lambda light chain locus. *Immunogenetics* 29, 173–179. <https://doi.org/10.1007/BF00373642>

Cattoretti, G., Chang, C.C., Cechova, K., Zhang, J., Ye, B.H., Falini, B., Louie, D.C., Offit, K., Chaganti, R.S., Dalla-Favera, R., 1995. BCL-6 protein is expressed in germinal-center B cells. *Blood* 86, 45–53.

Cerutti, A., Cols, M., Puga, I., 2013. Marginal zone B cells: virtues of innate-like antibody-producing lymphocytes. *Nat. Rev. Immunol.* 13, 118–132. <https://doi.org/10.1038/nri3383>

- Chakraborty, T., Chowdhury, D., Keyes, A., Jani, A., Subrahmanyam, R., Ivanova, I., Sen, R., 2007. Repeat organization and epigenetic regulation of the DH-Cmu domain of the immunoglobulin heavy-chain gene locus. *Mol. Cell* 27, 842–850. <https://doi.org/10.1016/j.molcel.2007.07.010>
- Chang, L.-H., Ghosh, S., Noordermeer, D., 2020. TADs and Their Borders: Free Movement or Building a Wall? *J. Mol. Biol.* 432, 643–652. <https://doi.org/10.1016/j.jmb.2019.11.025>
- Chao, M.P., Seita, J., Weissman, I.L., 2008. Establishment of a Normal Hematopoietic and Leukemia Stem Cell Hierarchy. *Cold Spring Harb. Symp. Quant. Biol.* 73, 439–449. <https://doi.org/10.1101/sqb.2008.73.031>
- Chaumeil, J., Skok, J.A., 2012. The role of CTCF in regulating V(D)J recombination. *Curr. Opin. Immunol.* 24, 153–159. <https://doi.org/10.1016/j.coi.2012.01.003>
- Chauveau, C., Cogné, M., 1996. Palindromic structure of the IgH 3' locus control region. *Nat. Genet.* 14, 15–16. <https://doi.org/10.1038/ng0996-15>
- Chen, J., Young, F., Bottaro, A., Stewart, V., Smith, R.K., Alt, F.W., 1993. Mutations of the intronic IgH enhancer and its flanking sequences differentially affect accessibility of the JH locus. *EMBO J.* 12, 4635–4645. <https://doi.org/10.1002/j.1460-2075.1993.tb06152.x>
- Chesi, M., Bergsagel, P.L., Shonukan, O.O., Martelli, M.L., Brents, L.A., Chen, T., Schröck, E., Ried, T., Kuehl, W.M., 1998. Frequent Dysregulation of the c-maf Proto-Oncogene at 16q23 by Translocation to an Ig Locus in Multiple Myeloma. *Blood* 91, 4457–4463. <https://doi.org/10.1182/blood.V91.12.4457>
- Chiang, M.Y., Monroe, J.G., 2001. Role for transcription Pax5A factor in maintaining commitment to the B cell lineage by selective inhibition of granulocyte-macrophage colony-stimulating factor receptor expression. *J. Immunol. Baltim. Md* 1950 166, 6091–6098. <https://doi.org/10.4049/jimmunol.166.10.6091>
- Chiang, M.Y., Monroe, J.G., 1999. BSAP/Pax5A expression blocks survival and expansion of early myeloid cells implicating its involvement in maintaining commitment to the B-lymphocyte lineage. *Blood* 94, 3621–3632.
- Chowdhury, D., Sen, R., 2003. Transient IL-7/IL-7R Signaling Provides a Mechanism for Feedback Inhibition of Immunoglobulin Heavy Chain Gene Rearrangements. *Immunity* 18, 229–241. [https://doi.org/10.1016/S1074-7613\(03\)00030-X](https://doi.org/10.1016/S1074-7613(03)00030-X)
- Chowdhury, D., Sen, R., 2001. Stepwise activation of the immunoglobulin mu heavy chain gene locus. *EMBO J.* 20, 6394–6403. <https://doi.org/10.1093/emboj/20.22.6394>
- Chung, J.B., Sater, R.A., Fields, M.L., Erikson, J., Monroe, J.G., 2002. CD23 defines two distinct subsets of immature B cells which differ in their responses to T cell help signals. *Int. Immunol.* 14, 157–166. <https://doi.org/10.1093/intimm/14.2.157>
- Cinamon, G., Zachariah, M.A., Lam, O.M., Foss, F.W., Cyster, J.G., 2008. Follicular shuttling of marginal zone B cells facilitates antigen transport. *Nat. Immunol.* 9, 54–62. <https://doi.org/10.1038/ni1542>
- Ciosk, R., Shirayama, M., Shevchenko, A., Tanaka, T., Toth, A., Shevchenko, A., Nasmyth, K., 2000. Cohesin's binding to chromosomes depends on a separate complex consisting of Scc2 and Scc4 proteins. *Mol. Cell* 5, 243–254. [https://doi.org/10.1016/s1097-2765\(00\)80420-7](https://doi.org/10.1016/s1097-2765(00)80420-7)
- Cobaleda, C., Schebesta, A., Delogu, A., Busslinger, M., 2007. Pax5: the guardian of B cell identity and function. *Nat. Immunol.* 8, 463–470. <https://doi.org/10.1038/ni1454>
- Corcoran, A.E., 2005. Immunoglobulin locus silencing and allelic exclusion. *Semin. Immunol.* 17, 141–154. <https://doi.org/10.1016/j.smim.2005.01.002>
- Costea, J., Schoeberl, U.E., Malzl, D., von der Linde, M., Fitz, J., Gupta, A., Makharova, M., Goloborodko, A., Pavri, R., 2023. A de novo transcription-dependent TAD boundary underpins critical multiway interactions during antibody class switch recombination. *Mol. Cell* S1097-2765(23)00037-0. <https://doi.org/10.1016/j.molcel.2023.01.014>
- Cremer, M., Küpper, K., Wagler, B., Wizelman, L., Hase, J. v., Weiland, Y., Kreja, L., Diebold, J., Speicher, M.R., Cremer, T., 2003. Inheritance of gene density-related higher order chromatin arrangements in normal and tumor cell nuclei. *J. Cell Biol.* 162, 809–820. <https://doi.org/10.1083/jcb.200304096>

Cremer, M., von Hase, J., Volm, T., Brero, A., Kreth, G., Walter, J., Fischer, C., Solovej, I., Cremer, C., Cremer, T., 2001. Non-random radial higher-order chromatin arrangements in nuclei of diploid human cells. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* 9, 541–567. <https://doi.org/10.1023/a:1012495201697>

Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., Bickmore, W.A., 1999. Differences in the localization and morphology of chromosomes in the human nucleus. *J. Cell Biol.* 145, 1119–1131. <https://doi.org/10.1083/jcb.145.6.1119>

- D -

D'Addabbo, P., Frezza, D., Sulentic, C.E.W., 2023. Evolutive emergence and divergence of an Ig regulatory node: An environmental sensor getting cues from the aryl hydrocarbon receptor? *Front. Immunol.* 14, 996119. <https://doi.org/10.3389/fimmu.2023.996119>

Dai, H.-Q., Hu, H., Lou, J., Ye, A.Y., Ba, Z., Zhang, X., Zhang, Y., Zhao, L., Yoon, H.S., Chapdelaine-Williams, A.M., Kyritsis, N., Chen, H., Johnson, K., Lin, S., Conte, A., Casellas, R., Lee, C.-S., Alt, F.W., 2021. Loop extrusion mediates physiological Igh locus contraction for RAG scanning. *Nature* 590, 338–343. <https://doi.org/10.1038/s41586-020-03121-7>

Davies, D.R., Metzger, H., 1983. Structural Basis of Antibody Function. *Annu. Rev. Immunol.* 1, 87–115. <https://doi.org/10.1146/annurev.iy.01.040183.000511>

de Wit, E., Vos, E.S.M., Holwerda, S.J.B., Valdes-Quezada, C., Verstegen, M.J.A.M., Teunissen, H., Splinter, E., Wijchers, P.J., Krijger, P.H.L., de Laat, W., 2015. CTCF Binding Polarity Determines Chromatin Looping. *Mol. Cell* 60, 676–684. <https://doi.org/10.1016/j.molcel.2015.09.023>

Dębek, S., Juszczynski, P., 2022. Super enhancers as master gene regulators in the pathogenesis of hematologic malignancies. *Biochim. Biophys. Acta Rev. Cancer* 1877, 188697. <https://doi.org/10.1016/j.bbcan.2022.188697>

Degner, S.C., Verma-Gaur, J., Wong, T.P., Bossen, C., Iverson, G.M., Torkamani, A., Vettermann, C., Lin, Y.C., Ju, Z., Schulz, D., Murre, C.S., Birshtein, B.K., Schork, N.J., Schlissel, M.S., Riblet, R., Murre, C., Feeney, A.J., 2011. CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. *Proc. Natl. Acad. Sci.* 108, 9566–9571. <https://doi.org/10.1073/pnas.1019391108>

Degner, S.C., Wong, T.P., Jankevicius, G., Feeney, A.J., 2009. Cutting edge: developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. *J. Immunol. Baltim. Md 1950* 182, 44–48. <https://doi.org/10.4049/jimmunol.182.1.44>

Dekker, J., Heard, E., 2015. Structural and functional diversity of Topologically Associating Domains. *FEBS Lett.* 589, 2877–2884. <https://doi.org/10.1016/j.febslet.2015.08.044>

DeKoter, R.P., Lee, H.-J., Singh, H., 2002. PU.1 regulates expression of the interleukin-7 receptor in lymphoid progenitors. *Immunity* 16, 297–309. [https://doi.org/10.1016/s1074-7613\(02\)00269-8](https://doi.org/10.1016/s1074-7613(02)00269-8)

Delgado-Benito, V., Rosen, D.B., Wang, Q., Gazumyan, A., Pai, J.A., Oliveira, T.Y., Sundaravinayagam, D., Zhang, W., Andreani, M., Keller, L., Kieffer-Kwon, K.-R., Pękowska, A., Jung, S., Driesner, M., Subbotin, R.I., Casellas, R., Chait, B.T., Nussenzweig, M.C., Di Virgilio, M., 2018. The Chromatin Reader ZMYND8 Regulates Igh Enhancers to Promote Immunoglobulin Class Switch Recombination. *Mol. Cell* 72, 636–649.e8. <https://doi.org/10.1016/j.molcel.2018.08.042>

Delogu, A., Schebesta, A., Sun, Q., Aschenbrenner, K., Perlot, T., Busslinger, M., 2006. Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. *Immunity* 24, 269–281. <https://doi.org/10.1016/j.immuni.2006.01.012>

Delpy, L., Le Bert, M., Cogné, M., Khamlichi, A.A., 2003. Germ-line transcription occurs on both the functional and the non-functional alleles of immunoglobulin constant heavy chain genes. *Eur. J. Immunol.* 33, 2108–2113. <https://doi.org/10.1002/eji.200323969>

Delpy, L., Sirac, C., Le Morvan, C., Cogné, M., 2004. Transcription-dependent somatic hypermutation occurs at similar levels on functional and nonfunctional rearranged IgH alleles. *J. Immunol. Baltim. Md 1950* 173, 1842–1848. <https://doi.org/10.4049/jimmunol.173.3.1842>

Desiderio, S.V., Yancopoulos, G.D., Paskind, M., Thomas, E., Boss, M.A., Landau, N., Alt, F.W., Baltimore, D., 1984. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. *Nature* 311, 752–755. <https://doi.org/10.1038/311752a0>



Di Noia, J., Neuberger, M.S., 2002. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature* 419, 43–48. <https://doi.org/10.1038/nature00981>

Dickerson, S.K., Market, E., Besmer, E., Papavasiliou, F.N., 2003. AID Mediates Hypermutation by Deaminating Single Stranded DNA. *J. Exp. Med.* 197, 1291–1296. <https://doi.org/10.1084/jem.20030481>

Dixon, J.R., Gorkin, D.U., Ren, B., 2016. Chromatin Domains: The Unit of Chromosome Organization. *Mol. Cell* 62, 668–680. <https://doi.org/10.1016/j.molcel.2016.05.018>

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., Diao, Y., Liang, J., Zhao, H., Lobanenkov, V.V., Ecker, J.R., Thomson, J.A., Ren, B., 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature* 518, 331–336. <https://doi.org/10.1038/nature14222>

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., Ren, B., 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380. <https://doi.org/10.1038/nature11082>

Dogan, I., Bertocci, B., Vilmont, V., Delbos, F., Mégrét, J., Storck, S., Reynaud, C.-A., Weill, J.-C., 2009. Multiple layers of B cell memory with different effector functions. *Nat. Immunol.* 10, 1292–1299. <https://doi.org/10.1038/ni.1814>

Dougier, H.-L., Reynaud, S., Pinaud, E., Carrion, C., Delpy, L., Cognø, M., 2006. Interallelic class switch recombination can reverse allelic exclusion and allow trans-complementation of an IgH locus switching defect. *Eur J Immunol* 11.

Duan, H., Heckman, C.A., Boxer, L.M., 2007. The immunoglobulin heavy-chain gene 3' enhancers deregulate bcl-2 promoter usage in t(14;18) lymphoma cells. *Oncogene* 26, 2635–2641. <https://doi.org/10.1038/sj.onc.1210061>

- E -

Early, P., Huang, H., Davis, M., Calame, K., Hood, L., 1980. An immunoglobulin heavy chain variable region gene is generated from three segments of DNA: VH, D and JH. *Cell* 19, 981–992. [https://doi.org/10.1016/0092-8674\(80\)90089-6](https://doi.org/10.1016/0092-8674(80)90089-6)

Ebert, A., McManus, S., Tagoh, H., Medvedovic, J., Salvaggio, G., Novatchkova, M., Tamir, I., Sommer, A., Jaritz, M., Busslinger, M., 2011. The Distal VH Gene Cluster of the Igh Locus Contains Distinct Regulatory Elements with Pax5 Transcription Factor-Dependent Activity in Pro-B Cells. *Immunity* 34, 175–187. <https://doi.org/10.1016/j.immuni.2011.02.005>

Edelman, G.M., 1973. Antibody structure and molecular immunology. *Science* 180, 830–840. <https://doi.org/10.1126/science.180.4088.830>

Emorine, L., Kuehl, M., Weir, L., Leder, P., Max, E.E., 1983. A conserved sequence in the immunoglobulin Jκ–Cκ intron: possible enhancer element. *Nature* 304, 447–449. <https://doi.org/10.1038/304447a0>

Engel, H., Rolink, A., Weiss, S., 1999. B cells are programmed to activate kappa and lambda for rearrangement at consecutive developmental stages. *Eur. J. Immunol.* 29, 2167–2176. [https://doi.org/10.1002/\(SICI\)1521-4141\(199907\)29:07<2167::AID-IMMU2167>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1521-4141(199907)29:07<2167::AID-IMMU2167>3.0.CO;2-H)

Ernst, P., Smale, S.T., 1995. Combinatorial regulation of transcription II: the immunoglobulin μ heavy chain gene. *Immunity* 2, 427–438. [https://doi.org/10.1016/1074-7613\(95\)90024-1](https://doi.org/10.1016/1074-7613(95)90024-1)

- F -

Featherstone, K., Wood, A.L., Bowen, A.J., Corcoran, A.E., 2010. The mouse immunoglobulin heavy chain V-D intergenic sequence contains insulators that may regulate ordered V(D)J recombination. *J. Biol. Chem.* 285, 9327–9338. <https://doi.org/10.1074/jbc.M109.098251>

Ferguson, A.R., 2004. Marginal zone B cells transport and deposit IgM-containing immune complexes onto follicular dendritic cells. *Int. Immunol.* 16, 1411–1422. <https://doi.org/10.1093/intimm/dxh142>

Ferrier, P., Krippel, B., Blackwell, T.K., Furley, A.J., Suh, H., Winoto, A., Cook, W.D., Hood, L., Costantini, F., Alt, F.W., 1990. Separate elements control DJ and VDJ rearrangement in a transgenic recombination substrate. *EMBO J.* 9, 117–125. <https://doi.org/10.1002/j.1460-2075.1990.tb08087.x>

Fortini, P., Dogliotti, E., 2007. Base damage and single-strand break repair: Mechanisms and functional significance of short- and long-patch repair subpathways. *DNA Repair* 6, 398–409. <https://doi.org/10.1016/j.dnarep.2006.10.008>

Fugmann, S.D., Lee, A.I., Shockett, P.E., Villey, I.J., Schatz, D.G., 2000. The RAG Proteins and V(D)J Recombination: Complexes, Ends, and Transposition. *Annu. Rev. Immunol.* 18, 495–527. <https://doi.org/10.1146/annurev.immunol.18.1.495>

Fukita, Y., Jacobs, H., Rajewsky, K., 1998. Somatic Hypermutation in the Heavy Chain Locus Correlates with Transcription. *Immunity* 9, 105–114. [https://doi.org/10.1016/S1074-7613\(00\)80592-0](https://doi.org/10.1016/S1074-7613(00)80592-0)

- G -

Gandhi, R., Gillespie, P.J., Hirano, T., 2006. Human Wapl is a cohesin-binding protein that promotes sister-chromatid resolution in mitotic prophase. *Curr. Biol.* CB 16, 2406–2417. <https://doi.org/10.1016/j.cub.2006.10.061>

Garot, A., Marquet, M., Saintamand, A., Bender, S., Le Noir, S., Rouaud, P., Carrion, C., Oruc, Z., Bébin, A.-G., Moreau, J., Lebrigand, K., Denizot, Y., Alt, F.W., Cogné, M., Pinaud, E., 2016. Sequential activation and distinct functions for distal and proximal modules within the IgH 3' regulatory region. *Proc. Natl. Acad. Sci.* 113, 1618–1623. <https://doi.org/10.1073/pnas.1514090113>

Gassler, J., Brandão, H.B., Imakaev, M., Flyamer, I.M., Ladstätter, S., Bickmore, W.A., Peters, J.-M., Mirny, L.A., Tachibana, K., 2017. A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *EMBO J.* 36, 3600–3618. <https://doi.org/10.15252/embj.201798083>

Genetta, T., Ruezinsky, D., Kadesch, T., 1994. Displacement of an E-box-binding repressor by basic helix-loop-helix proteins: implications for B-cell specificity of the immunoglobulin heavy-chain enhancer. *Mol. Cell. Biol.* 14, 6153–6163. <https://doi.org/10.1128/mcb.14.9.6153-6163.1994>

Giallourakis, C.C., Franklin, A., Guo, C., Cheng, H.-L., Yoon, H.S., Gallagher, M., Perlot, T., Andzelm, M., Murphy, A.J., Macdonald, L.E., Yancopoulos, G.D., Alt, F.W., 2010. Elements between the IgH variable (V) and diversity (D) clusters influence antisense transcription and lineage-specific V(D)J recombination. *Proc. Natl. Acad. Sci. U. S. A.* 107, 22207–22212. <https://doi.org/10.1073/pnas.1015954107>

Gillies, S.D., Morrison, S.L., Oi, V.T., Tonegawa, S., 1983. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell* 33, 717–728. [https://doi.org/10.1016/0092-8674\(83\)90014-4](https://doi.org/10.1016/0092-8674(83)90014-4)

Girelli, G., Custodio, J., Kallas, T., Agostini, F., Wernersson, E., Spanjaard, B., Mota, A., Kolbeinsdottir, S., Gelali, E., Crosetto, N., Bienko, M., 2020. GPSeq reveals the radial organization of chromatin in the cell nucleus. *Nat. Biotechnol.* 38, 1184–1193. <https://doi.org/10.1038/s41587-020-0519-y>

Giudicelli, V., Chaume, D., Lefranc, M.-P., 2005. IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res.* 33, D256–261. <https://doi.org/10.1093/nar/gki010>

Gordon, M.S., Kanegai, C.M., Doerr, J.R., Wall, R., 2003. Somatic hypermutation of the B cell receptor genes *B29* (*Ig*  $\beta$ , CD79b) and *mb1* (*Ig*  $\alpha$ , CD79a). *Proc. Natl. Acad. Sci.* 100, 4126–4131. <https://doi.org/10.1073/pnas.0735266100>

Gostissa, M., Yan, C.T., Bianco, J.M., Cogné, M., Pinaud, E., Alt, F.W., 2009. Long-range oncogenic activation of *Igh-c-myc* translocations by the *Igh* 3' regulatory region. *Nature* 462, 803–807. <https://doi.org/10.1038/nature08633>

Gramlich, H.S., Reisbig, T., Schatz, D.G., 2012. AID-Targeting and Hypermutation of Non-Immunoglobulin Genes Does Not Correlate with Proximity to Immunoglobulin Genes in Germinal Center B Cells. *PLoS ONE* 7, e39601. <https://doi.org/10.1371/journal.pone.0039601>

Grasser, F., Neusser, M., Fiegler, H., Thormeyer, T., Cremer, M., Carter, N.P., Cremer, T., Müller, S., 2008. Replication-timing-correlated spatial chromatin arrangements in cancer and in primate interphase nuclei. *J. Cell Sci.* 121, 1876–1886. <https://doi.org/10.1242/jcs.026989>

Grawunder, U., Leu, T.M.J., Schatz, D.G., Werner, A., Rolink, A.G., Melchers, F., Winkler, T.H., 1995. Down-regulation of RAG1 and RAG2 gene expression in PreB cells after functional immunoglobulin heavy chain rearrangement. *Immunity* 3, 601–608. [https://doi.org/10.1016/1074-7613\(95\)90131-0](https://doi.org/10.1016/1074-7613(95)90131-0)

Gruber, S., Haering, C.H., Nasmyth, K., 2003. Chromosomal cohesin forms a ring. *Cell* 112, 765–777. [https://doi.org/10.1016/s0092-8674\(03\)00162-4](https://doi.org/10.1016/s0092-8674(03)00162-4)

Guo, C., Gerasimova, T., Hao, H., Ivanova, I., Chakraborty, T., Selimyan, R., Oltz, E.M., Sen, R., 2011. Two Forms of Loops Generate the Chromatin Conformation of the Immunoglobulin Heavy-Chain Gene Locus. *Cell* 147, 332–343. <https://doi.org/10.1016/j.cell.2011.08.049>

Guttman, M., Rinn, J.L., 2012. Modular regulatory principles of large non-coding RNAs. *Nature* 482, 339–346. <https://doi.org/10.1038/nature10887>

- H -

Haarhuis, J.H.I., van der Weide, R.H., Blomen, V.A., Yáñez-Cuna, J.O., Amendola, M., van Ruiten, M.S., Krijger, P.H.L., Teunissen, H., Medema, R.H., van Steensel, B., Brummelkamp, T.R., de Wit, E., Rowland, B.D., 2017. The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* 169, 693–707.e14. <https://doi.org/10.1016/j.cell.2017.04.013>

Hagman, J., Rudin, C.M., Haasch, D., Chaplin, D., Storb, U., 1990. A novel enhancer in the immunoglobulin lambda locus is duplicated and functionally independent of NF kappa B. *Genes Dev.* 4, 978–992. <https://doi.org/10.1101/gad.4.6.978>

Hakes, D.J., Berezney, R., 1991. DNA binding properties of the nuclear matrix and individual nuclear matrix proteins. Evidence for salt-resistant DNA binding sites. *J. Biol. Chem.* 266, 11131–11140.

Halverson, R., Torres, R.M., Pelanda, R., 2004. Receptor editing is the main mechanism of B cell tolerance toward membrane antigens. *Nat. Immunol.* 5, 645–650. <https://doi.org/10.1038/ni1076>

Han, L., Masani, S., Yu, K., 2011. Overlapping activation-induced cytidine deaminase hotspot motifs in Ig class-switch recombination. *Proc. Natl. Acad. Sci.* 108, 11584–11589. <https://doi.org/10.1073/pnas.1018726108>

Haniuda, K., Kitamura, D., 2019. Induced Germinal Center B Cell Culture System. *BIO-Protoc.* 9. <https://doi.org/10.21769/BioProtoc.3163>

Hansen, A.S., Pustova, I., Cattoglio, C., Tjian, R., Darzacq, X., 2017. CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* 6, e25776. <https://doi.org/10.7554/eLife.25776>

Hardy, R.R., Kincade, P.W., Dorshkind, K., 2007. The protean nature of cells in the B lymphocyte lineage. *Immunity* 26, 703–714. <https://doi.org/10.1016/j.immuni.2007.05.013>

Hayakawa, K., Hardy, R.R., Herzenberg, L.A., Herzenberg, L.A., 1985. Progenitors for Ly-1 B cells are distinct from progenitors for other B cells. *J. Exp. Med.* 161, 1554–1568. <https://doi.org/10.1084/jem.161.6.1554>

Heng, H.H.Q., Goetze, S., Ye, C.J., Liu, G., Stevens, J.B., Bremer, S.W., Wykes, S.M., Bode, J., Krawetz, S.A., 2004. Chromatin loops are selectively anchored using scaffold/matrix-attachment regions. *J. Cell Sci.* 117, 999–1008. <https://doi.org/10.1242/jcs.00976>

Henseler, R.A., Romer, E.J., Sulentic, C.E.W., 2009. Diverse chemicals including aryl hydrocarbon receptor ligands modulate transcriptional activity of the 3'immunoglobulin heavy chain regulatory region. *Toxicology* 261, 9–18. <https://doi.org/10.1016/j.tox.2009.03.015>

Higashi, T.L., Eickhoff, P., Sousa, J.S., Locke, J., Nans, A., Flynn, H.R., Snijders, A.P., Papageorgiou, G., O'Reilly, N., Chen, Z.A., O'Reilly, F.J., Rappsilber, J., Costa, A., Uhlmann, F., 2020. A Structure-Based Mechanism for DNA Entry into the Cohesin Ring. *Mol. Cell* 79, 917–933.e9. <https://doi.org/10.1016/j.molcel.2020.07.013>

Hill, L., Ebert, A., Jaritz, M., Wutz, G., Nagasaka, K., Tagoh, H., Kostanova-Poliakova, D., Schindler, K., Sun, Q., Bönelt, P., Fischer, M., Peters, J.-M., Busslinger, M., 2020. Wapl repression by Pax5 promotes V gene recombination by Igh loop extrusion. *Nature* 584, 142–147. <https://doi.org/10.1038/s41586-020-2454-y>

Hnisz, D., Abraham, B.J., Lee, T.I., Lau, A., Saint-André, V., Sigova, A.A., Hoke, H.A., Young, R.A., 2013. Super-enhancers in the control of cell identity and disease. *Cell* 155, 934–947. <https://doi.org/10.1016/j.cell.2013.09.053>

Holley, D.W., Groh, B.S., Wozniak, G., Donohoe, D.R., Sun, W., Godfrey, V., Bultman, S.J., 2014. The BRG1 chromatin remodeler regulates widespread changes in gene expression and cell proliferation during B cell activation. *J. Cell. Physiol.* 229, 44–52. <https://doi.org/10.1002/jcp.24414>

Holwerda, S.J.B., van de Werken, H.J.G., Ribeiro de Almeida, C., Bergen, I.M., de Bruijn, M.J.W., Verstegen, M.J.A.M., Simonis, M., Splinter, E., Wijchers, P.J., Hendriks, R.W., de Laat, W., 2013. Allelic exclusion of the immunoglobulin heavy chain locus is independent of its nuclear localization in mature B cells. *Nucleic Acids Res.* 41, 6905–6916. <https://doi.org/10.1093/nar/gkt491>

Hu, J., Meyers, R.M., Dong, J., Panchakshari, R.A., Alt, F.W., Frock, R.L., 2016. Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat. Protoc.* 11, 853–871. <https://doi.org/10.1038/nprot.2016.043>

Hwang, J.K., Alt, F.W., Yeap, L.-S., 2015. Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination. *Microbiol. Spectr.* 3, MDNA3-0037–2014. <https://doi.org/10.1128/microbiolspec.MDNA3-0037-2014>

- I -

Inlay, M., Alt, F.W., Baltimore, D., Xu, Y., 2002. Essential roles of the kappa light chain intronic enhancer and 3' enhancer in kappa rearrangement and demethylation. *Nat. Immunol.* 3, 463–468. <https://doi.org/10.1038/ni790>

Iqbal, J., Greiner, T.C., Patel, K., Dave, B.J., Smith, L., Ji, J., Wright, G., Sanger, W.G., Pickering, D.L., Jain, S., Horsman, D.E., Shen, Y., Fu, K., Weisenburger, D.D., Hans, C.P., Campo, E., Gascoyne, R.D., Rosenwald, A., Jaffe, E.S., Delabie, J., Rimsza, L., Ott, G., Müller-Hermelink, H.K., Connors, J.M., Vose, J.M., McKeithan, T., Staudt, L.M., Chan, W.C., for the Leukemia/Lymphoma Molecular Profiling Project (LLMPP), 2007. Distinctive patterns of BCL6 molecular alterations and their functional consequences in different subgroups of diffuse large B-cell lymphoma. *Leukemia* 21, 2332–2343. <https://doi.org/10.1038/sj.leu.2404856>

Ito, S., Nagaoka, H., Shinkura, R., Begum, N., Muramatsu, M., Nakata, M., Honjo, T., 2004. Activation-induced cytidine deaminase shuttles between nucleus and cytoplasm like apolipoprotein B mRNA editing catalytic polypeptide 1. *Proc. Natl. Acad. Sci. U. S. A.* 101, 1975–1980. <https://doi.org/10.1073/pnas.0307335101>

- J -

Johnston, C.M., Wood, A.L., Bolland, D.J., Corcoran, A.E., 2006. Complete sequence assembly and characterization of the C57BL/6 mouse Ig heavy chain V region. *J. Immunol. Baltim. Md 1950* 176, 4221–4234. <https://doi.org/10.4049/jimmunol.176.7.4221>

Jung, D., Giallourakis, C., Mostoslavsky, R., Alt, F.W., 2006. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu. Rev. Immunol.* 24, 541–570. <https://doi.org/10.1146/annurev.immunol.23.021704.115830>

- K -

Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., Taatjes, D.J., Dekker, J., Young, R.A., 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467, 430–435. <https://doi.org/10.1038/nature09380>

Kanehiro, Y., Todo, K., Negishi, M., Fukuoka, J., Gan, W., Hikasa, T., Kaga, Y., Takemoto, M., Magari, M., Li, X., Manley, J.L., Ohmori, H., Kanayama, N., 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.* 109, 1216–1221. <https://doi.org/10.1073/pnas.1120368109>

Kato, L., Begum, N.A., Burroughs, A.M., Doi, T., Kawai, J., Daub, C.O., Kawaguchi, T., Matsuda, F., Hayashizaki, Y., Honjo, T., 2012. Nonimmunoglobulin target loci of activation-induced cytidine deaminase (AID) share unique features with immunoglobulin genes. *Proc. Natl. Acad. Sci. U. S. A.* 109, 2479–2484. <https://doi.org/10.1073/pnas.1120791109>

Kawaguchi, Y., Nariki, H., Kawamoto, N., Kanehiro, Y., Miyazaki, S., Suzuki, M., Magari, M., Tokumitsu, H., Kanayama, N., 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. <https://doi.org/10.1016/j.bbrc.2017.02.097>

- Khalil, A.M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., Thomas, K., Presser, A., Bernstein, B.E., van Oudenaarden, A., Regev, A., Lander, E.S., Rinn, J.L., 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 106, 11667–11672. <https://doi.org/10.1073/pnas.0904715106>
- Khodabakhshi, A.H., Morin, R.D., Fejes, A.P., Mungall, A.J., Mungall, K.L., Bolger-Munro, M., Johnson, N.A., Connors, J.M., Gascoyne, R.D., Marra, M.A., Birol, I., Jones, S.J.M., 2012. Recurrent targets of aberrant somatic hypermutation in lymphoma. *Oncotarget* 3, 1308–1319. <https://doi.org/10.18632/oncotarget.653>
- Kim, Y.W., Lee, S., Yun, J., Kim, A., 2015. Chromatin looping and eRNA transcription precede the transcriptional activation of gene in the  $\beta$ -globin locus. *Biosci. Rep.* 35, e00179. <https://doi.org/10.1042/BSR20140126>
- Klymenko, T., Bloehdorn, J., Bahlo, J., Robrecht, S., Akylzhanova, G., Cox, K., Estenfelder, S., Wang, J., Edelmann, J., Strefford, J.C., Wojdacz, T.K., Fischer, K., Hallek, M., Stilgenbauer, S., Cragg, M., Gribben, J., Braun, A., 2018. Lamin B1 regulates somatic mutations and progression of B-cell malignancies. *Leukemia* 32, 364–375. <https://doi.org/10.1038/leu.2017.255>
- Komori, A., Xu, Z., Wu, X., Zan, H., Casali, P., 2006. Biased dA/dT somatic hypermutation as regulated by the heavy chain intronic iEmu enhancer and 3' Ealpha enhancers in human lymphoblastoid B cells. *Mol. Immunol.* 43, 1817–1826. <https://doi.org/10.1016/j.molimm.2005.10.018>
- Kosak, S.T., Skok, J.A., Medina, K.L., Riblet, R., Le Beau, M.M., Fisher, A.G., Singh, H., 2002. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* 296, 158–162. <https://doi.org/10.1126/science.1068768>
- Kothapalli, N.R., Collura, K.M., Norton, D.D., Fugmann, S.D., 2011. Separation of mutational and transcriptional enhancers in Ig genes. *J. Immunol. Baltim. Md* 1950 187, 3247–3255. <https://doi.org/10.4049/jimmunol.1101568>
- Kottmann, A.H., Zevnik, B., Welte, M., Nielsen, P.J., Köhler, G., 1994. A second promoter and enhancer element within the immunoglobulin heavy chain locus. *Eur. J. Immunol.* 24, 817–821. <https://doi.org/10.1002/eji.1830240407>
- Kozmik, Z., Wang, S., Dörfler, P., Adams, B., Busslinger, M., 1992. The promoter of the CD19 gene is a target for the B-cell-specific transcription factor BSAP. *Mol. Cell. Biol.* 12, 2662–2672. <https://doi.org/10.1128/mcb.12.6.2662-2672.1992>
- Krokan, H.E., Bjoras, M., 2013. Base Excision Repair. *Cold Spring Harb. Perspect. Biol.* 5, a012583–a012583. <https://doi.org/10.1101/cshperspect.a012583>
- Kueng, S., Hegemann, B., Peters, B.H., Lipp, J.J., Schleiffer, A., Mechtler, K., Peters, J.-M., 2006. Wapl Controls the Dynamic Association of Cohesin with Chromatin. *Cell* 127, 955–967. <https://doi.org/10.1016/j.cell.2006.09.040>
- Kumar Singh, A., Tamrakar, A., Jaiswal, A., Kanayama, N., Agarwal, A., Tripathi, P., Kodgire, P., 2019. Splicing regulator SRSF1-3 that controls somatic hypermutation of IgV genes interacts with topoisomerase 1 and AID. *Mol. Immunol.* 116, 63–72. <https://doi.org/10.1016/j.molimm.2019.10.002>
- Kurosaki, T., Kometani, K., Ise, W., 2015. Memory B cells. *Nat. Rev. Immunol.* 15, 149–159. <https://doi.org/10.1038/nri3802>

- L -

- Laffleur, B., Basu, U., Lim, J., 2017. RNA Exosome and Non-coding RNA-Coupled Mechanisms in AID-Mediated Genomic Alterations. *J. Mol. Biol.* 429, 3230–3241. <https://doi.org/10.1016/j.jmb.2016.12.021>
- Laffleur, B., Lim, J., Zhang, W., Chen, Y., Pefanis, E., Bizarro, J., Batista, C.R., Wu, L., Economides, A.N., Wang, J., Basu, U., 2021. Noncoding RNA processing by DIS3 regulates chromosomal architecture and somatic hypermutation in B cells. *Nat. Genet.* 53, 230–242. <https://doi.org/10.1038/s41588-020-00772-0>
- Lam, T., Thomas, L.M., White, C.A., Li, G., Pone, E.J., Xu, Z., Casali, P., 2017. Correction: Scaffold Functions of 14-3-3 Adaptors in B Cell Immunoglobulin Class Switch DNA Recombination. *PLOS ONE* 12, e0174195. <https://doi.org/10.1371/journal.pone.0174195>
- Langerak, P., Nygren, A.O.H., Krijger, P.H.L., van den Berk, P.C.M., Jacobs, H., 2007. A/T mutagenesis in hypermutated immunoglobulin genes strongly depends on PCNAK164 modification. *J. Exp. Med.* 204, 1989–1998. <https://doi.org/10.1084/jem.20070902>

- Larijani, M., Martin, A., 2007. Single-Stranded DNA Structure and Positional Context of the Target Cytidine Determine the Enzymatic Efficiency of AID. *Mol. Cell. Biol.* 27, 8038–8048. <https://doi.org/10.1128/MCB.01046-07>
- Le Noir, S., Laffleur, B., Carrion, C., Garot, A., Lecardeur, S., Pinaud, E., Denizot, Y., Skok, J., Cogné, M., 2017. The IgH locus 3' cis-regulatory super-enhancer co-opts AID for allelic transvection. *Oncotarget* 8, 12929–12940. <https://doi.org/10.18632/oncotarget.14585>
- Lee, J.-H., Paull, T.T., 2007. Activation and regulation of ATM kinase activity in response to DNA double-strand breaks. *Oncogene* 26, 7741–7748. <https://doi.org/10.1038/sj.onc.1210872>
- Lemaître, C., Grabarz, A., Tsouroula, K., Andronov, L., Furst, A., Pankotai, T., Heyer, V., Rogier, M., Attwood, K.M., Kessler, P., Dellaire, G., Klaholz, B., Reina-San-Martin, B., Soutoglou, E., 2014. Nuclear position dictates DNA repair pathway choice. *Genes Dev.* 28, 2450–2463. <https://doi.org/10.1101/gad.248369.114>
- Lengronne, A., Katou, Y., Mori, S., Yokobayashi, S., Kelly, G.P., Itoh, T., Watanabe, Y., Shirahige, K., Uhlmann, F., 2004. Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* 430, 573–578. <https://doi.org/10.1038/nature02742>
- Lennon, G.G., Perry, R.P., 1985. C mu-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5'-nontranslatable exon. *Nature* 318, 475–478. <https://doi.org/10.1038/318475a0>
- Le Noir, S., Boyer, F., Lecardeur, S., Brousse, M., Oruc, Z., Cook-Moreau, J., Denizot, Y., Cogné, M., 2017. Functional anatomy of the immunoglobulin heavy chain 3' super-enhancer needs not only core enhancer elements but also their unique DNA context. *Nucleic Acids Res.* 45, 5829–5837. <https://doi.org/10.1093/nar/gkx203>
- Li, F., Yan, Y., Pieretti, J., Feldman, D.A., Eckhardt, L.A., 2010. Comparison of identical and functional Igh alleles reveals a nonessential role for Eμ in somatic hypermutation and class-switch recombination. *J. Immunol. Baltim. Md* 1950 185, 6049–6057. <https://doi.org/10.4049/jimmunol.0902992>
- Li, Z., Otevrel, T., Gao, Y., Cheng, H.-L., Seed, B., Stamato, T.D., Taccioli, G.E., Alt, F.W., 1995. The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. *Cell* 83, 1079–1089. [https://doi.org/10.1016/0092-8674\(95\)90135-3](https://doi.org/10.1016/0092-8674(95)90135-3)
- Liang, Z., Zhao, L., Ye, A.Y., Lin, S.G., Zhang, Y., Guo, C., Dai, H.-Q., Ba, Z., Alt, F.W., 2023. Contribution of the IGCR1 regulatory element and the 3' *Igh* CTCF-binding elements to regulation of *Igh* V(D)J recombination. *Proc. Natl. Acad. Sci.* 120, e2306564120. <https://doi.org/10.1073/pnas.2306564120>
- Lieber, M.R., 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* 79, 181–211. <https://doi.org/10.1146/annurev.biochem.052308.093131>
- Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., Sandstrom, R., Bernstein, B., Bender, M.A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L.A., Lander, E.S., Dekker, J., 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293. <https://doi.org/10.1126/science.1181369>
- Lieberman, R., Giannini, S.L., Birshtein, B.K., Eckhardt, L.A., 1991. An enhancer at the 3' end of the mouse immunoglobulin heavy chain locus. *Nucleic Acids Res.* 19, 933–937. <https://doi.org/10.1093/nar/19.4.933>
- Lin, H., Grosschedl, R., 1995. Failure of B-cell differentiation in mice lacking the transcription factor EBF. *Nature* 376, 263–267. <https://doi.org/10.1038/376263a0>
- Liu, M., Duke, J.L., Richter, D.J., Vinuesa, C.G., Goodnow, C.C., Kleinstein, S.H., Schatz, D.G., 2008. Two levels of protection for the B cell genome during somatic hypermutation. *Nature* 451, 841–845. <https://doi.org/10.1038/nature06547>
- Liu, Z.-M., George-Raizen, J.B., Li, S., Meyers, K.C., Chang, M.Y., Garrard, W.T., 2002. Chromatin structural analyses of the mouse Igkappa gene locus reveal new hypersensitive sites specifying a transcriptional silencer and enhancer. *J. Biol. Chem.* 277, 32640–32649. <https://doi.org/10.1074/jbc.M204065200>
- Liubchenko, G.A., Appleberry, H.C., Holers, V.M., Banda, N.K., Willis, V.C., Lyubchenko, T., 2012. Potentially autoreactive naturally occurring transitional T3 B lymphocytes exhibit a unique signaling profile. *J. Autoimmun.* 38, 293–303. <https://doi.org/10.1016/j.jaut.2011.12.005>

Loder, F., Mutschler, B., Ray, R.J., Paige, C.J., Sideras, P., Torres, R., Lamers, M.C., Carsetti, R., 1999. B cell development in the spleen takes place in discrete steps and is determined by the quality of B cell receptor-derived signals. *J. Exp. Med.* 190, 75–89. <https://doi.org/10.1084/jem.190.1.75>

Lovely, G.A., Brewster, R.C., Schatz, D.G., Baltimore, D., Phillips, R., 2015. Single-molecule analysis of RAG-mediated V(D)J DNA cleavage. *Proc. Natl. Acad. Sci.* 112. <https://doi.org/10.1073/pnas.1503477112>

Low, M.S.Y., Brodie, E.J., Fedele, P.L., Liao, Y., Grigoriadis, G., Strasser, A., Kallies, A., Willis, S.N., Tellier, J., Shi, W., Gabriel, S., O'Donnell, K., Pitt, C., Nutt, S.L., Tarlinton, D., 2019. IRF4 Activity Is Required in Established Plasma Cells to Regulate Gene Transcription and Mitochondrial Homeostasis. *Cell Rep.* 29, 2634-2645.e5. <https://doi.org/10.1016/j.celrep.2019.10.097>

- M -

Ma, Y., Pannicke, U., Schwarz, K., Lieber, M.R., 2002. Hairpin Opening and Overhang Processing by an Artemis/DNA-Dependent Protein Kinase Complex in Nonhomologous End Joining and V(D)J Recombination. *Cell* 108, 781–794. [https://doi.org/10.1016/S0092-8674\(02\)00671-2](https://doi.org/10.1016/S0092-8674(02)00671-2)

Malu, S., De Ioannes, P., Kozlov, M., Greene, M., Francis, D., Hanna, M., Pena, J., Escalante, C.R., Kurosawa, A., Erdjument-Bromage, H., Tempst, P., Adachi, N., Vezzone, P., Villa, A., Aggarwal, A.K., Cortes, P., 2012. Artemis C-terminal region facilitates V(D)J recombination through its interactions with DNA Ligase IV and DNA-PKcs. *J. Exp. Med.* 209, 955–963. <https://doi.org/10.1084/jem.20111437>

Marchal, C., Sima, J., Gilbert, D.M., 2019. Control of DNA replication timing in the 3D genome. *Nat. Rev. Mol. Cell Biol.* 20, 721–737. <https://doi.org/10.1038/s41580-019-0162-y>

Marquet, M., Garot, A., Bender, S., Carrion, C., Rouaud, P., Lecardeur, S., Denizot, Y., Cogné, M., Pinaud, E., 2014. The E $\mu$  Enhancer Region Influences H Chain Expression and B Cell Fate without Impacting IgVH Repertoire and Immune Response In Vivo. *J. Immunol.* 193, 1171–1183. <https://doi.org/10.4049/jimmunol.1302868>

Mårtensson, I.L., Ceredig, R., 2000. Review article: role of the surrogate light chain and the pre-B-cell receptor in mouse B-cell development. *Immunology* 101, 435–441. <https://doi.org/10.1046/j.1365-2567.2000.00151.x>

Martin, F., Oliver, A.M., Kearney, J.F., 2001. Marginal Zone and B1 B Cells Unite in the Early Response against T-Independent Blood-Borne Particulate Antigens. *Immunity* 14, 617–629. [https://doi.org/10.1016/S1074-7613\(01\)00129-7](https://doi.org/10.1016/S1074-7613(01)00129-7)

Martin, O.A., Thomas, M., Marquet, M., Bruzeau, C., Garot, A., Brousse, M., Bender, S., Carrion, C., Choi, J.E., Vuong, B.Q., Gearhart, P.J., Maul, R.W., Le Noir, S., Pinaud, E., 2023. The IgH E $\mu$ -MAR regions promote UNG-dependent error-prone repair to optimize somatic hypermutation. *Front. Immunol.* 14, 1030813. <https://doi.org/10.3389/fimmu.2023.1030813>

Matthias, P., Baltimore, D., 1993. The immunoglobulin heavy chain locus contains another B-cell-specific 3' enhancer close to the alpha constant region. *Mol. Cell. Biol.* 13, 1547–1553. <https://doi.org/10.1128/mcb.13.3.1547-1553.1993>

Maul, R.W., Cao, Z., Venkataraman, L., Giorgetti, C.A., Press, J.L., Denizot, Y., Du, H., Sen, R., Gearhart, P.J., 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. <https://doi.org/10.1084/jem.20131512>

Maul, R.W., Saribasak, H., Cao, Z., Gearhart, P.J., 2015. Topoisomerase I deficiency causes RNA polymerase II accumulation and increases AID abundance in immunoglobulin variable genes. *DNA Repair* 30, 46–52. <https://doi.org/10.1016/j.dnarep.2015.03.004>

Mayer, R., Brero, A., von Hase, J., Schroeder, T., Cremer, T., Dietzel, S., 2005. Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC Cell Biol.* 6, 44. <https://doi.org/10.1186/1471-2121-6-44>

McBlane, J.F., van Gent, D.C., Ramsden, D.A., Romeo, C., Cuomo, C.A., Gellert, M., Oettinger, M.A., 1995. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell* 83, 387–395. [https://doi.org/10.1016/0092-8674\(95\)90116-7](https://doi.org/10.1016/0092-8674(95)90116-7)

Medina, K.L., Pongubala, J.M.R., Reddy, K.L., Lancki, D.W., Dekoter, R., Kieslinger, M., Grosschedl, R., Singh, H., 2004. Assembling a gene regulatory network for specification of the B cell fate. *Dev. Cell* 7, 607–617. <https://doi.org/10.1016/j.devcel.2004.08.006>

Medvedovic, J., Ebert, A., Tagoh, H., Tamir, I.M., Schwickert, T.A., Novatchkova, M., Sun, Q., Huis in 't Veld, P.J., Guo, C., Yoon, H.S., Denizot, Y., Holwerda, S.J.B., de Laat, W., Cogné, M., Shi, Y., Alt, F.W., Busslinger, M., 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. <https://doi.org/10.1016/j.immuni.2013.08.011>

Melchers, F., Karasuyama, H., Haasner, D., Bauer, S., Kudo, A., Sakaguchi, N., Jameson, B., Rolink, A., 1993. The surrogate light chain in B-cell development. *Immunol. Today* 14, 60–68. [https://doi.org/10.1016/0167-5699\(93\)90060-X](https://doi.org/10.1016/0167-5699(93)90060-X)

Merrell, K.T., Benschop, R.J., Gauld, S.B., Aviszus, K., Decote-Ricardo, D., Wysocki, L.J., Cambier, J.C., 2006. Identification of Anergic B Cells within a Wild-Type Repertoire. *Immunity* 25, 953–962. <https://doi.org/10.1016/j.immuni.2006.10.017>

Michaelson, J.S., Giannini, S.L., Birshtein, B.K., 1995. Identification of 3' alpha-hs4, a novel Ig heavy chain enhancer element regulated at multiple stages of B cell differentiation. *Nucleic Acids Res.* 23, 975–981. <https://doi.org/10.1093/nar/23.6.975>

Mikkola, I., Heavey, B., Horcher, M., Busslinger, M., 2002. Reversion of B cell commitment upon loss of Pax5 expression. *Science* 297, 110–113. <https://doi.org/10.1126/science.1067518>

Monroe, J.G., Dorshkind, K., 2007. Fate decisions regulating bone marrow and peripheral B lymphocyte development. *Adv. Immunol.* 95, 1–50. [https://doi.org/10.1016/S0065-2776\(07\)95001-4](https://doi.org/10.1016/S0065-2776(07)95001-4)

Morgado, M.G., Cam, P., Gris-Liebe, C., Cazenave, P.A., Jouvin-Marche, E., 1989. Further evidence that BALB/c and C57BL/6 gamma 2a genes originate from two distinct isotypes. *EMBO J.* 8, 3245–3251. <https://doi.org/10.1002/j.1460-2075.1989.tb08484.x>

Morshead, K.B., Ciccone, D.N., Taverna, S.D., Allis, C.D., Oettinger, M.A., 2003. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. *Proc. Natl. Acad. Sci. U. S. A.* 100, 11577–11582. <https://doi.org/10.1073/pnas.1932643100>

Muramatsu, M., Kinoshita, K., Fagarasan, S., Yamada, S., Shinkai, Y., Honjo, T., 2000. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell* 102, 553–563. [https://doi.org/10.1016/S0092-8674\(00\)00078-7](https://doi.org/10.1016/S0092-8674(00)00078-7)

- N -

Nadel, B., Marculescu, R., Le, T., Rudnicki, M., Böcskő, S., Jäger, U., 2001. Novel insights into the mechanism of t(14;18)(q32;q21) translocation in follicular lymphoma. *Leuk. Lymphoma* 42, 1181–1194. <https://doi.org/10.3109/10428190109097743>

Nambu, Y., Sugai, M., Gonda, H., Lee, C.-G., Katakai, T., Agata, Y., Yokota, Y., Shimizu, A., 2003. Transcription-coupled events associating with immunoglobulin switch region chromatin. *Science* 302, 2137–2140. <https://doi.org/10.1126/science.1092481>

Nanni, L., Ceri, S., Logie, C., 2020. Spatial patterns of CTCF sites define the anatomy of TADs and their boundaries. *Genome Biol.* 21, 197. <https://doi.org/10.1186/s13059-020-02108-x>

Nera, K.-P., Lassila, O., 2006. Pax5 – a Critical Inhibitor of Plasma Cell Fate. *Scand. J. Immunol.* 64, 190–199. <https://doi.org/10.1111/j.1365-3083.2006.01809.x>

Neuberger, M.S., Noia, J.M.D., Beale, R.C.L., Williams, G.T., Yang, Z., Rada, C., 2005. Somatic hypermutation at A·T pairs: polymerase error versus dUTP incorporation. *Nat. Rev. Immunol.* 5, 171–178. <https://doi.org/10.1038/nri1553>

Ng, H.H., Robert, F., Young, R.A., Struhl, K., 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* 11, 709–719. [https://doi.org/10.1016/S1097-2765\(03\)00092-3](https://doi.org/10.1016/S1097-2765(03)00092-3)

Nichogiannopoulou, A., Trevisan, M., Neben, S., Friedrich, C., Georgopoulos, K., 1999. Defects in hemopoietic stem cell activity in Ikaros mutant mice. *J. Exp. Med.* 190, 1201–1214. <https://doi.org/10.1084/jem.190.9.1201>

Niirö, H., Clark, E.A., 2002. Regulation of B-cell fate by antigen-receptor signals. *Nat. Rev. Immunol.* 2, 945–956. <https://doi.org/10.1038/nri955>



Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., Bruneau, B.G., 2017. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* 169, 930-944.e22. <https://doi.org/10.1016/j.cell.2017.05.004>

Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Blüthgen, N., Dekker, J., Heard, E., 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485, 381–385. <https://doi.org/10.1038/nature11049>

Norvell, A., Mandik, L., Monroe, J.G., 1995. Engagement of the antigen-receptor on immature murine B lymphocytes results in death by apoptosis. *J. Immunol. Baltim. Md* 1950 154, 4404–4413.

Nussenzweig, A., Chen, C., da Costa Soares, V., Sanchez, M., Sokol, K., Nussenzweig, M.C., Li, G.C., 1996. Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature* 382, 551–555. <https://doi.org/10.1038/382551a0>

Nussenzweig, A., Nussenzweig, M.C., 2010. Origin of chromosomal translocations in lymphoid cancer. *Cell* 141, 27–38. <https://doi.org/10.1016/j.cell.2010.03.016>

Nutt, S.L., Heavey, B., Rolink, A.G., Busslinger, M., 1999. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* 401, 556–562. <https://doi.org/10.1038/44076>

Nutt, S.L., Hodgkin, P.D., Tarlinton, D.M., Corcoran, L.M., 2015. The generation of antibody-secreting plasma cells. *Nat. Rev. Immunol.* 15, 160–171. <https://doi.org/10.1038/nri3795>

- O -

Oettinger, M.A., Schatz, D.G., Gorka, C., Baltimore, D., 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 248, 1517–1523. <https://doi.org/10.1126/science.2360047>

Oracki, S.A., Walker, J.A., Hibbs, M.L., Corcoran, L.M., Tarlinton, D.M., 2010. Plasma cell development and survival: Plasma cell development and survival. *Immunol. Rev.* 237, 140–159. <https://doi.org/10.1111/j.1600-065X.2010.00940.x>

Osborne, C.S., Chakalova, L., Mitchell, J.A., Horton, A., Wood, A.L., Bolland, D.J., Corcoran, A.E., Fraser, P., 2007. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biol.* 5, e192. <https://doi.org/10.1371/journal.pbio.0050192>

Osipovich, O.A., Subrahmanyam, R., Pierce, S., Sen, R., Oltz, E.M., 2009. Cutting edge: SWI/SNF mediates antisense Igh transcription and locus-wide accessibility in B cell precursors. *J. Immunol. Baltim. Md* 1950 183, 1509–1513. <https://doi.org/10.4049/jimmunol.0900896>

- P -

Padlan, E.A., 1994. Anatomy of the antibody molecule. *Mol. Immunol.* 31, 169–217. [https://doi.org/10.1016/0161-5890\(94\)90001-9](https://doi.org/10.1016/0161-5890(94)90001-9)

Park, S.-K., Xiang, Y., Feng, X., Garrard, W.T., 2014. Pronounced cohabitation of active immunoglobulin genes from three different chromosomes in transcription factories during maximal antibody synthesis. *Genes Dev.* 28, 1159–1164. <https://doi.org/10.1101/gad.237479.114>

Parslow, T.G., Blair, D.L., Murphy, W.J., Granner, D.K., 1984. Structure of the 5' ends of immunoglobulin genes: a novel conserved sequence. *Proc. Natl. Acad. Sci. U. S. A.* 81, 2650–2654. <https://doi.org/10.1073/pnas.81.9.2650>

Pasqualucci, L., Dalla-Favera, R., 2018. Genetics of diffuse large B-cell lymphoma. *Blood* 131, 2307–2319. <https://doi.org/10.1182/blood-2017-11-764332>

Pasqualucci, L., Migliazza, A., Basso, K., Houldsworth, J., Chaganti, R.S.K., Dalla-Favera, R., 2003. Mutations of the BCL6 proto-oncogene disrupt its negative autoregulation in diffuse large B-cell lymphoma. *Blood* 101, 2914–2923. <https://doi.org/10.1182/blood-2002-11-3387>

Pasqualucci, L., Migliazza, A., Fracchiolla, N., William, C., Neri, A., Baldini, L., Chaganti, R.S.K., Klein, U., Küppers, R., Rajewsky, K., Dalla-Favera, R., 1998. BCL-6 mutations in normal germinal center B cells: Evidence of somatic hypermutation acting outside Ig loci. *Proc. Natl. Acad. Sci.* 95, 11816–11821. <https://doi.org/10.1073/pnas.95.20.11816>

- Pasqualucci, L., Neumeister, P., Goossens, T., Nanjangud, G., Chaganti, R.S.K., Küppers, R., Dalla-Favera, R., 2001. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* 412, 341–346. <https://doi.org/10.1038/35085588>
- Pavlicek, J.W., Lyubchenko, Y.L., Chang, Y., 2008. Quantitative Analyses of RAG–RSS Interactions and Conformations Revealed by Atomic Force Microscopy. *Biochemistry* 47, 11204–11211. <https://doi.org/10.1021/bi801426x>
- Pavri, R., Gazumyan, A., Jankovic, M., Di Virgilio, M., Klein, I., Ansarah-Sobrinho, C., Resch, W., Yamane, A., Reina San-Martin, B., Barreto, V., Nieland, T.J., Root, D.E., Casellas, R., Nussenzweig, M.C., 2010. Activation-induced cytidine deaminase targets DNA at sites of RNA polymerase II stalling by interaction with Spt5. *Cell* 143, 122–133. <https://doi.org/10.1016/j.cell.2010.09.017>
- Pavri, R., Nussenzweig, M.C., 2011. AID targeting in antibody diversity. *Adv. Immunol.* 110, 1–26. <https://doi.org/10.1016/B978-0-12-387663-8.00005-3>
- Pawlitzy, I., Angeles, C.V., Siegel, A.M., Stanton, M.L., Riblet, R., Brodeur, P.H., 2006. Identification of a Candidate Regulatory Element within the 5' Flanking Region of the Mouse *Igh* Locus Defined by Pro-B Cell-Specific Hypersensitivity Associated with Binding of PU.1, Pax5, and E2A. *J. Immunol.* 176, 6839–6851. <https://doi.org/10.4049/jimmunol.176.11.6839>
- Pefanis, E., Basu, U., 2015. RNA Exosome Regulates AID DNA Mutator Activity in the B Cell Genome. *Adv. Immunol.* 127, 257–308. <https://doi.org/10.1016/bs.ai.2015.04.002>
- Pefanis, E., Wang, J., Rothschild, G., Lim, J., Chao, J., Rabadan, R., Economides, A.N., Basu, U., 2014. Noncoding RNA transcription targets AID to divergently transcribed loci in B cells. *Nature* 514, 389–393. <https://doi.org/10.1038/nature13580>
- Peled, J.U., Kuang, F.L., Iglesias-Ussel, M.D., Roa, S., Kalis, S.L., Goodman, M.F., Scharff, M.D., 2008. The biochemistry of somatic hypermutation. *Annu. Rev. Immunol.* 26, 481–511. <https://doi.org/10.1146/annurev.immunol.26.021607.090236>
- Perlot, T., Alt, F.W., Bassing, C.H., Suh, H., Pinaud, E., 2005. Elucidation of IgH intronic enhancer functions via germ-line deletion. *Proc. Natl. Acad. Sci.* 102, 14362–14367. <https://doi.org/10.1073/pnas.0507090102>
- Perlot, T., Li, G., Alt, F.W., 2008. Antisense transcripts from immunoglobulin heavy-chain locus V(D)J and switch regions. *Proc. Natl. Acad. Sci.* 105, 3843–3848. <https://doi.org/10.1073/pnas.0712291105>
- Petro, J.B., Gerstein, R.M., Lowe, J., Carter, R.S., Shinnars, N., Khan, W.N., 2002. Transitional type 1 and 2 B lymphocyte subsets are differentially responsive to antigen receptor signaling. *J. Biol. Chem.* 277, 48009–48019. <https://doi.org/10.1074/jbc.M200305200>
- Pike, B.L., Boyd, A.W., Nossal, G.J., 1982. Clonal anergy: the universally anergic B lymphocyte. *Proc. Natl. Acad. Sci. U. S. A.* 79, 2013–2017. <https://doi.org/10.1073/pnas.79.6.2013>
- Pillai, S., Cariappa, A., 2009. The follicular versus marginal zone B lymphocyte cell fate decision. *Nat. Rev. Immunol.* 9, 767–777. <https://doi.org/10.1038/nri2656>
- Pillai, S., Cariappa, A., Moran, S.T., 2005. MARGINAL ZONE B CELLS. *Annu. Rev. Immunol.* 23, 161–196. <https://doi.org/10.1146/annurev.immunol.23.021704.115728>
- Pinaud, E., Khamlichi, A.A., Le Morvan, C., Drouet, M., Nalesso, V., Le Bert, M., Cogné, M., 2001. Localization of the 3' IgH Locus Elements that Effect Long-Distance Regulation of Class Switch Recombination. *Immunity* 15, 187–199. [https://doi.org/10.1016/S1074-7613\(01\)00181-9](https://doi.org/10.1016/S1074-7613(01)00181-9)
- Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K., Thurman, R.E., Cheng, Y., Gülsoy, G., Dennis, J.H., Snyder, M.P., Stamatoyannopoulos, J.A., Taylor, J., Hardison, R.C., Kahveci, T., Ren, B., Gilbert, D.M., 2014. Topologically associating domains are stable units of replication-timing regulation. *Nature* 515, 402–405. <https://doi.org/10.1038/nature13986>
- Predeus, A.V., Gopalakrishnan, S., Huang, Y., Tang, J., Feeney, A.J., Oltz, E.M., Artyomov, M.N., 2014. Targeted Chromatin Profiling Reveals Novel Enhancers in Ig H and Ig L Chain Loci. *J. Immunol.* 192, 1064–1070. <https://doi.org/10.4049/jimmunol.1302800>
- Pucella, J.N., Chaudhuri, J., 2017. AID Invited to the G4 Summit. *Mol. Cell* 67, 355–357. <https://doi.org/10.1016/j.molcel.2017.07.020>

- Q -

Qiao, Q., Wang, L., Meng, F.-L., Hwang, J.K., Alt, F.W., Wu, H., 2017. AID Recognizes Structured DNA for Class Switch Recombination. *Mol. Cell* 67, 361-373.e4. <https://doi.org/10.1016/j.molcel.2017.06.034>

- R -

Rada, C., Di Noia, J.M., Neuberger, M.S., 2004. Mismatch Recognition and Uracil Excision Provide Complementary Paths to Both Ig Switching and the A/T-Focused Phase of Somatic Mutation. *Mol. Cell* 16, 163–171. <https://doi.org/10.1016/j.molcel.2004.10.011>

Rada, C., Ehrenstein, M.R., Neuberger, M.S., Milstein, C., 1998. Hot Spot Focusing of Somatic Hypermutation in MSH2-Deficient Mice Suggests Two Stages of Mutational Targeting. *Immunity* 9, 135–141. [https://doi.org/10.1016/S1074-7613\(00\)80595-6](https://doi.org/10.1016/S1074-7613(00)80595-6)

Rada, C., Jarvis, J.M., Milstein, C., 2002. AID-GFP chimeric protein increases hypermutation of Ig genes with no evidence of nuclear localization. *Proc. Natl. Acad. Sci. U. S. A.* 99, 7003–7008. <https://doi.org/10.1073/pnas.092160999>

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., Aiden, E.L., 2014. A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* 159, 1665–1680. <https://doi.org/10.1016/j.cell.2014.11.021>

Roa, S., Avdievich, E., Peled, J.U., MacCarthy, T., Werling, U., Kuang, F.L., Kan, R., Zhao, C., Bergman, A., Cohen, P.E., Edelman, W., Scharff, M.D., 2008. Ubiquitylated PCNA plays a role in somatic hypermutation and class-switch recombination and is required for meiotic progression. *Proc. Natl. Acad. Sci.* 105, 16248–16253. <https://doi.org/10.1073/pnas.0808182105>

Robertson, A.B., Klungland, A., Rognes, T., Leiros, I., 2009. DNA Repair in Mammalian Cells: Base excision repair: the long and short of it. *Cell. Mol. Life Sci.* 66, 981–993. <https://doi.org/10.1007/s00018-009-8736-z>

Rocha, P.P., Micsinai, M., Kim, J.R., Hewitt, S.L., Souza, P.P., Trimarchi, T., Strino, F., Parisi, F., Kluger, Y., Skok, J.A., 2012. Close Proximity to Igh Is a Contributing Factor to AID-Mediated Translocations. *Mol. Cell* 47, 873–885. <https://doi.org/10.1016/j.molcel.2012.06.036>

Rogozin, I.B., Kolchanov, N.A., 1992. Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. *Biochim. Biophys. Acta* 1171, 11–18. [https://doi.org/10.1016/0167-4781\(92\)90134-l](https://doi.org/10.1016/0167-4781(92)90134-l)

Ronai, D., Iglesias-Ussel, M.D., Fan, M., Shulman, M.J., Scharff, M.D., 2005. Complex regulation of somatic hypermutation by cis-acting sequences in the endogenous IgH gene in hybridoma cells. *Proc. Natl. Acad. Sci. U. S. A.* 102, 11829–11834. <https://doi.org/10.1073/pnas.0505449102>

Rothschild, G., Basu, U., 2017. Lingering Questions about Enhancer RNA and Enhancer Transcription-Coupled Genomic Instability. *Trends Genet.* 33, 143–154. <https://doi.org/10.1016/j.tig.2016.12.002>

Rouaud, P., Vincent-Fabert, C., Saintamand, A., Fiancette, R., Marquet, M., Robert, I., Reina-San-Martin, B., Pinaud, E., Cogné, M., Denizot, Y., 2013. The IgH 3' regulatory region controls somatic hypermutation in germinal center B cells. *J. Exp. Med.* 210, 1501–1507. <https://doi.org/10.1084/jem.20130072>

Rowley, M.J., Corces, V.G., 2018. Organizational principles of 3D genome architecture. *Nat. Rev. Genet.* 19, 789–800. <https://doi.org/10.1038/s41576-018-0060-8>

Roy, A.L., Sen, R., Roeder, R.G., 2011. Enhancer-promoter communication and transcriptional regulation of Igh. *Trends Immunol.* 32, 532–539. <https://doi.org/10.1016/j.it.2011.06.012>

Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T.C., Robins, A.J., Dalton, S., Gilbert, D.M., 2010. Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. *Genome Res.* 20, 761–770. <https://doi.org/10.1101/gr.099655.109>

- S -

Saintamand, A., Rouaud, P., Saad, F., Rios, G., Cogné, M., Denizot, Y., 2015. Elucidation of IgH 3' region regulatory role during class switch recombination via germline deletion. *Nat. Commun.* 6, 7084. <https://doi.org/10.1038/ncomms8084>

- Saintamand, A., Vincent-Fabert, C., Garot, A., Rouaud, P., Oruc, Z., Magnone, V., Cogné, M., Denizot, Y., 2016. Deciphering the importance of the palindromic architecture of the immunoglobulin heavy-chain 3' regulatory region. *Nat. Commun.* 7, 10730. <https://doi.org/10.1038/ncomms10730>
- Sakai, E., Bottaro, A., Davidson, L., Sleckman, B.P., Alt, F.W., 1999. Recombination and transcription of the endogenous Ig heavy chain locus is effected by the Ig heavy chain intronic enhancer core region in the absence of the matrix attachment regions. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1526–1531. <https://doi.org/10.1073/pnas.96.4.1526>
- Salisbury, R.L., Sulentic, C.E.W., 2015. The AhR and NF-κB/Rel Proteins Mediate the Inhibitory Effect of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin on the 3' Immunoglobulin Heavy Chain Regulatory Region. *Toxicol. Sci. Off. J. Soc. Toxicol.* 148, 443–459. <https://doi.org/10.1093/toxsci/kfv193>
- Saribasak, H., Maul, R.W., Cao, Z., McClure, R.L., Yang, W., McNeill, D.R., Wilson, D.M., Gearhart, P.J., 2011. XRCC1 suppresses somatic hypermutation and promotes alternative nonhomologous end joining in Igh genes. *J. Exp. Med.* 208, 2209–2216. <https://doi.org/10.1084/jem.20111135>
- Schatz, D.G., Ji, Y., 2011. Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* 11, 251–263. <https://doi.org/10.1038/nri2941>
- Schatz, D.G., Oettinger, M.A., Baltimore, D., 1989. The V(D)J recombination activating gene, RAG-1. *Cell* 59, 1035–1048. [https://doi.org/10.1016/0092-8674\(89\)90760-5](https://doi.org/10.1016/0092-8674(89)90760-5)
- Schatz, D.G., Swanson, P.C., 2011. V(D)J recombination: mechanisms of initiation. *Annu. Rev. Genet.* 45, 167–202. <https://doi.org/10.1146/annurev-genet-110410-132552>
- Schlissel, M.S., 2004. Regulation of activation and recombination of the murine Igkappa locus. *Immunol. Rev.* 200, 215–223. <https://doi.org/10.1111/j.0105-2896.2004.00157.x>
- Schmiedel, D., Hezroni, H., Hamburg, A., Shulman, Z., 2021. Brg1 Supports B Cell Proliferation and Germinal Center Formation Through Enhancer Activation. *Front. Immunol.* 12, 705848. <https://doi.org/10.3389/fimmu.2021.705848>
- Schmitz, R., Young, R.M., Ceribelli, M., Jhavar, S., Xiao, W., Zhang, M., Wright, G., Shaffer, A.L., Hodson, D.J., Buras, E., Liu, X., Powell, J., Yang, Y., Xu, W., Zhao, H., Kohlhammer, H., Rosenwald, A., Kluin, P., Müller-Hermelink, H.K., Ott, G., Gascoyne, R.D., Connors, J.M., Rimsza, L.M., Campo, E., Jaffe, E.S., Delabie, J., Smeland, E.B., Olgwang, M.D., Reynolds, S.J., Fisher, R.I., Braziel, R.M., Tubbs, R.R., Cook, J.R., Weisenburger, D.D., Chan, W.C., Pittaluga, S., Wilson, W., Waldmann, T.A., Rowe, M., Mbulaiteye, S.M., Rickinson, A.B., Staudt, L.M., 2012. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature* 490, 116–120. <https://doi.org/10.1038/nature11378>
- Schroeder, H.W., Cavacini, L., 2010. Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* 125, S41–S52. <https://doi.org/10.1016/j.jaci.2009.09.046>
- Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C.H., Mirny, L., Spitz, F., 2017. Two independent modes of chromatin organization revealed by cohesin removal. *Nature* 551, 51–56. <https://doi.org/10.1038/nature24281>
- Scourzic, L., Salataj, E., Apostolou, E., 2021. Deciphering the Complexity of 3D Chromatin Organization Driving Lymphopoiesis and Lymphoid Malignancies. *Front. Immunol.* 12, 669881. <https://doi.org/10.3389/fimmu.2021.669881>
- Seet, C.S., Brumbaugh, R.L., Kee, B.L., 2004. Early B Cell Factor Promotes B Lymphopoiesis with Reduced Interleukin 7 Responsiveness in the Absence of E2A. *J. Exp. Med.* 199, 1689–1700. <https://doi.org/10.1084/jem.20032202>
- Seitan, V.C., Faure, A.J., Zhan, Y., McCord, R.P., Lajoie, B.R., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A.G., Flicek, P., Dekker, J., Merkenschlager, M., 2013. Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res.* 23, 2066–2077. <https://doi.org/10.1101/gr.161620.113>
- Sen, R., Oltz, E., 2006. Genetic and epigenetic regulation of IgH gene assembly. *Curr. Opin. Immunol.* 18, 237–242. <https://doi.org/10.1016/j.coi.2006.03.008>

- Senigl, F., Maman, Y., Dinesh, R.K., Alinikula, J., Seth, R.B., Pecnova, L., Omer, A.D., Rao, S.S.P., Weisz, D., Buerstedde, J.-M., Aiden, E.L., Casellas, R., Hejnar, J., Schatz, D.G., 2019a. Topologically Associated Domains Delineate Susceptibility to Somatic Hypermutation. *Cell Rep.* 29, 3902-3915.e8. <https://doi.org/10.1016/j.celrep.2019.11.039>
- Senigl, F., Maman, Y., Dinesh, R.K., Alinikula, J., Seth, R.B., Pecnova, L., Omer, A.D., Rao, S.S.P., Weisz, D., Buerstedde, J.-M., Aiden, E.L., Casellas, R., Hejnar, J., Schatz, D.G., 2019b. Topologically Associated Domains Delineate Susceptibility to Somatic Hypermutation. *Cell Rep.* 29, 3902-3915.e8. <https://doi.org/10.1016/j.celrep.2019.11.039>
- Sepulveda, M.A., Garrett, F.E., Price-Whelan, A., Birshstein, B.K., 2005. Comparative analysis of human and mouse 3' Igh regulatory regions identifies distinctive structural features. *Mol. Immunol.* 42, 605-615. <https://doi.org/10.1016/j.molimm.2004.09.006>
- Serwe, M., Sablitzky, F., 1993. V(D)J recombination in B cells is impaired but not blocked by targeted deletion of the immunoglobulin heavy chain intron enhancer. *EMBO J.* 12, 2321-2327. <https://doi.org/10.1002/j.1460-2075.1993.tb05886.x>
- Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., Cavalli, G., 2012. Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. *Cell* 148, 458-472. <https://doi.org/10.1016/j.cell.2012.01.010>
- Shaffer, A.L., Lin, K.I., Kuo, T.C., Yu, X., Hurt, E.M., Rosenwald, A., Giltmane, J.M., Yang, L., Zhao, H., Calame, K., Staudt, L.M., 2002. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity* 17, 51-62. [https://doi.org/10.1016/s1074-7613\(02\)00335-7](https://doi.org/10.1016/s1074-7613(02)00335-7)
- Shaffer, A.L., Yu, X., He, Y., Boldrick, J., Chan, E.P., Staudt, L.M., 2000. BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. *Immunity* 13, 199-212. [https://doi.org/10.1016/s1074-7613\(00\)00020-0](https://doi.org/10.1016/s1074-7613(00)00020-0)
- Shen, H.M., Peters, A., Baron, B., Zhu, X., Storb, U., 1998. Mutation of *BCL-6* Gene in Normal B Cells by the Process of Somatic Hypermutation of Ig Genes. *Science* 280, 1750-1752. <https://doi.org/10.1126/science.280.5370.1750>
- Shen, H.M., Wuerffel, R., Cantillo, J.F., Priyadarshi, S., Lei, X., Liang, J., Wu, Y.L., Kenter, A.L., 2021. Loop extrusion promotes an alternate pathway for isotype switching. *Cell Rep.* 37, 110059. <https://doi.org/10.1016/j.celrep.2021.110059>
- Shen, Y., Iqbal, J., Huang, J.Z., Zhou, G., Chan, W.C., 2004. BCL2 protein expression parallels its mRNA level in normal and malignant B cells. *Blood* 104, 2936-2939. <https://doi.org/10.1182/blood-2004-01-0243>
- Sheppard, E.C., Morrish, R.B., Dillon, M.J., Leyland, R., Chahwan, R., 2018. Epigenomic Modifications Mediating Antibody Maturation. *Front. Immunol.* 9, 355. <https://doi.org/10.3389/fimmu.2018.00355>
- Shi, Z., Gao, H., Bai, X.-C., Yu, H., 2020. Cryo-EM structure of the human cohesin-NIPBL-DNA complex. *Science* 368, 1454-1459. <https://doi.org/10.1126/science.abb0981>
- Shinkai, Y., Rathbun, G., Lam, K.P., Oltz, E.M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A.M., 1992. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* 68, 855-867. [https://doi.org/10.1016/0092-8674\(92\)90029-c](https://doi.org/10.1016/0092-8674(92)90029-c)
- Shinkura, R., Ito, S., Begum, N.A., Nagaoka, H., Muramatsu, M., Kinoshita, K., Sakakibara, Y., Hijikata, H., Honjo, T., 2004. Separate domains of AID are required for somatic hypermutation and class-switch recombination. *Nat. Immunol.* 5, 707-712. <https://doi.org/10.1038/ni1086>
- Sitnicka, E., Bryder, D., Theilgaard-Mönch, K., Buza-Vidas, N., Adolfsson, J., Jacobsen, S.E.W., 2002. Key Role of flt3 Ligand in Regulation of the Common Lymphoid Progenitor but Not in Maintenance of the Hematopoietic Stem Cell Pool. *Immunity* 17, 463-472. [https://doi.org/10.1016/S1074-7613\(02\)00419-3](https://doi.org/10.1016/S1074-7613(02)00419-3)
- Skok, J.A., Brown, K.E., Azuara, V., Caparros, M.-L., Baxter, J., Takacs, K., Dillon, N., Gray, D., Perry, R.P., Merkenschlager, M., Fisher, A.G., 2001. Nonequivalent nuclear location of immunoglobulin alleles in B lymphocytes. *Nat. Immunol.* 2, 848-854. <https://doi.org/10.1038/ni0901-848>
- Smith, K.G.C., Hewitson, T.D., Nossal, G.J.V., Tarlinton, D.M., 1996. The phenotype and fate of the antibody-forming cells of the splenic foci. *Eur. J. Immunol.* 26, 444-448. <https://doi.org/10.1002/eji.1830260226>

- Smith, K.S., Liu, L.L., Ganesan, S., Michor, F., De, S., 2017. Nuclear topology modulates the mutational landscapes of cancer genomes. *Nat. Struct. Mol. Biol.* 24, 1000–1006. <https://doi.org/10.1038/nsmb.3474>
- Sofueva, S., Yaffe, E., Chan, W.-C., Georgopoulou, D., Vietri Rudan, M., Mira-Bontenbal, H., Pollard, S.M., Schroth, G.P., Tanay, A., Hadjir, S., 2013. Cohesin-mediated interactions organize chromosomal domain architecture: Functional role for cohesin in chromosome domain structure. *EMBO J.* 32, 3119–3129. <https://doi.org/10.1038/emboj.2013.237>
- Song, H., Cerny, J., 2003. Functional Heterogeneity of Marginal Zone B Cells Revealed by Their Ability to Generate Both Early Antibody-forming Cells and Germinal Centers with Hypermutation and Memory in Response to a T-dependent Antigen. *J. Exp. Med.* 198, 1923–1935. <https://doi.org/10.1084/jem.20031498>
- Souabni, A., Cobaleda, C., Schebesta, M., Busslinger, M., 2002. Pax5 Promotes B Lymphopoiesis and Blocks T Cell Development by Repressing Notch1. *Immunity* 17, 781–793. [https://doi.org/10.1016/S1074-7613\(02\)00472-7](https://doi.org/10.1016/S1074-7613(02)00472-7)
- Spilianakis, C.G., Lalioti, M.D., Town, T., Lee, G.R., Flavell, R.A., 2005. Interchromosomal associations between alternatively expressed loci. *Nature* 435, 637–645. <https://doi.org/10.1038/nature03574>
- Stamatopoulos, K., Kosmas, C., Belessi, C., Kyriazopoulos, P., Papadaki, T., Anagnostou, D., Loukopoulos, D., 1999. Molecular analysis of bcl-1/IgH junctional sequences in mantle cell lymphoma: potential mechanism of the t(11;14) chromosomal translocation. *Br. J. Haematol.* 105, 190–197. <https://doi.org/10.1111/j.1365-2141.1999.01314.x>
- Stavnezer, J., Linehan, E.K., Thompson, M.R., Habboub, G., Ucher, A.J., Kadungure, T., Tsuchimoto, D., Nakabeppu, Y., Schrader, C.E., 2014. Differential expression of APE1 and APE2 in germinal centers promotes error-prone repair and A:T mutations during somatic hypermutation. *Proc. Natl. Acad. Sci.* 111, 9217–9222. <https://doi.org/10.1073/pnas.1405590111>
- Stavnezer, J., Schrader, C.E., 2014. IgH chain class switch recombination: mechanism and regulation. *J. Immunol. Baltim. Md 1950* 193, 5370–5378. <https://doi.org/10.4049/jimmunol.1401849>
- Stevens, T.J., Lando, D., Basu, S., Atkinson, L.P., Cao, Y., Lee, S.F., Leeb, M., Wohlfahrt, K.J., Boucher, W., O’Shaughnessy-Kirwan, A., Cramard, J., Faure, A.J., Ralser, M., Blanco, E., Morey, L., Sansó, M., Palayret, M.G.S., Lehner, B., Di Croce, L., Wutz, A., Hendrich, B., Klenerman, D., Laue, E.D., 2017. 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* 544, 59–64. <https://doi.org/10.1038/nature21429>
- Storb, U., 2014. Why does somatic hypermutation by AID require transcription of its target genes? *Adv. Immunol.* 122, 253–277. <https://doi.org/10.1016/B978-0-12-800267-4.00007-9>
- Storb, U., Haasch, D., Arp, B., Sanchez, P., Cazenave, P.A., Miller, J., 1989. Physical linkage of mouse lambda genes by pulsed-field gel electrophoresis suggests that the rearrangement process favors proximate target sequences. *Mol. Cell. Biol.* 9, 711–718. <https://doi.org/10.1128/mcb.9.2.711-718.1989>
- Su, T.T., Rawlings, D.J., 2002. Transitional B lymphocyte subsets operate as distinct checkpoints in murine splenic B cell development. *J. Immunol. Baltim. Md 1950* 168, 2101–2110. <https://doi.org/10.4049/jimmunol.168.5.2101>
- Swerdlow, S.H., Campo, E., Pileri, S.A., Harris, N.L., Stein, H., Siebert, R., Advani, R., Ghielmini, M., Salles, G.A., Zelenetz, A.D., Jaffe, E.S., 2016. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood* 127, 2375–2390. <https://doi.org/10.1182/blood-2016-01-643569>
- Szczepińska, T., Rusek, A.M., Plewczynski, D., 2019. Intermingling of chromosome territories. *Genes. Chromosomes Cancer* 58, 500–506. <https://doi.org/10.1002/gcc.22736>

- T -

- Takizawa, T., Meaburn, K.J., Misteli, T., 2008. The meaning of gene positioning. *Cell* 135, 9–13. <https://doi.org/10.1016/j.cell.2008.09.026>
- Tangye, S.G., 2011. Staying alive: regulation of plasma cell survival. *Trends Immunol.* 32, 595–602. <https://doi.org/10.1016/j.it.2011.09.001>
- Tantin, D., Tussie-Luna, M.I., Roy, A.L., Sharp, P.A., 2004. Regulation of Immunoglobulin Promoter Activity by TFII-I Class Transcription Factors. *J. Biol. Chem.* 279, 5460–5469. <https://doi.org/10.1074/jbc.M311177200>

- Taylor, J.J., Pape, K.A., Jenkins, M.K., 2012. A germinal center-independent pathway generates unswitched memory B cells early in the primary response. *J. Exp. Med.* 209, 597–606. <https://doi.org/10.1084/jem.20111696>
- Teague, B.N., Pan, Y., Mudd, P.A., Nakken, B., Zhang, Q., Szodoray, P., Kim-Howard, X., Wilson, P.C., Farris, A.D., 2007. Cutting Edge: Transitional T3 B Cells Do Not Give Rise to Mature B Cells, Have Undergone Selection, and Are Reduced in Murine Lupus. *J. Immunol.* 178, 7511–7515. <https://doi.org/10.4049/jimmunol.178.12.7511>
- Thomas, M., Bruzeau, C., Martin, O.A., Pollet, J., Bender, S., Carrion, C., Le Noir, S., Pinaud, E., 2023. A dual function for the chromatin organizer Special A-T rich Binding Protein 1 in B-lineage cells. *Cell. Mol. Immunol.* <https://doi.org/10.1038/s41423-023-01069-y>
- Tinguely, A., Chemin, G., Péron, S., Sirac, C., Reynaud, S., Cogné, M., Delpy, L., 2012. Cross talk between immunoglobulin heavy-chain transcription and RNA surveillance during B cell development. *Mol. Cell. Biol.* 32, 107–117. <https://doi.org/10.1128/MCB.06138-11>
- Truffinet, V., Pinaud, E., Cogné, N., Petit, B., Guglielmi, L., Cogné, M., Denizot, Y., 2007. The 3' IgH locus control region is sufficient to deregulate a c-myc transgene and promote mature B cell malignancies with a predominant Burkitt-like phenotype. *J. Immunol. Baltim. Md* 1950 179, 6033–6042. <https://doi.org/10.4049/jimmunol.179.9.6033>
- Tsujimoto, Y., Gorham, J., Cossman, J., Jaffe, E., Croce, C.M., 1985. The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. *Science* 229, 1390–1393. <https://doi.org/10.1126/science.3929382>
- Tsukada, S., Sugiyama, H., Oka, Y., Kishimoto, S., 1990. Estimation of D segment usage in initial D to JH joinings in a murine immature B cell line. Preferential usage of DFL16.1, the most 5' D segment and DQ52, the most JH-proximal D segment. *J. Immunol. Baltim. Md* 1950 144, 4053–4059.
- Tunyaplin, C., Shaffer, A.L., Angelin-Duclos, C.D., Yu, X., Staudt, L.M., Calame, K.L., 2004. Direct Repression of *prdm1* by Bcl-6 Inhibits Plasmacytic Differentiation. *J. Immunol.* 173, 1158–1165. <https://doi.org/10.4049/jimmunol.173.2.1158>

- V -

- van Steensel, B., Belmont, A.S., 2017. Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* 169, 780–791. <https://doi.org/10.1016/j.cell.2017.04.022>
- Victoria, G.D., Schwickert, T.A., Fooksman, D.R., Kamphorst, A.O., Meyer-Hermann, M., Dustin, M.L., Nussenzweig, M.C., 2010. Germinal Center Dynamics Revealed by Multiphoton Microscopy with a Photoactivatable Fluorescent Reporter. *Cell* 143, 592–605. <https://doi.org/10.1016/j.cell.2010.10.032>
- Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A., Hadjur, S., 2015. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.* 10, 1297–1309. <https://doi.org/10.1016/j.celrep.2015.02.004>
- Vincent-Fabert, C., Fiancette, R., Cogné, M., Pinaud, E., Denizot, Y., 2010a. The IgH 3' regulatory region and its implication in lymphomagenesis. *Eur. J. Immunol.* 40, 3306–3311. <https://doi.org/10.1002/eji.201040778>
- Vincent-Fabert, C., Fiancette, R., Pinaud, E., Denizot, Y., 2010b. Genomic deletion of the whole IgH 3J regulatory region (hs3a, hs1,2, hs3b, and hs4) dramatically affects class switch recombination and Ig secretion to all isotypes 116, 5.
- Volpi, S.A., Verma-Gaur, J., Hassan, R., Ju, Z., Roa, S., Chatterjee, S., Werling, U., Hou, H., Will, B., Steidl, U., Scharff, M., Edelman, W., Feeney, A.J., Birshtein, B.K., 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. *J. Immunol.* 188, 2556–2566. <https://doi.org/10.4049/jimmunol.1102763>

- W -

- Wang, J., Boxer, L.M., 2005. Regulatory elements in the immunoglobulin heavy chain gene 3'-enhancers induce c-myc deregulation and lymphomagenesis in murine B cells. *J. Biol. Chem.* 280, 12766–12773. <https://doi.org/10.1074/jbc.M412446200>
- Wang, J.H., Gostissa, M., Yan, C.T., Goff, P., Hickernell, T., Hansen, E., Difilippantonio, S., Wesemann, D.R., Zarrin, A.A., Rajewsky, K., Nussenzweig, A., Alt, F.W., 2009. Mechanisms promoting translocations in editing and switching peripheral B cells. *Nature* 460, 231–236. <https://doi.org/10.1038/nature08159>

Wang, J.H., Gostissa, M., Yan, C.T., Goff, P., Hickernell, T., Hansen, E., Difilippantonio, S., Wesemann, D.R., Zarrin, A.A., Rajewsky, K., Nussenzweig, A., Alt, F.W., 2009. Mechanisms promoting translocations in editing and switching peripheral B cells. *Nature* 460, 231–236. <https://doi.org/10.1038/nature08159>

Wang, T.-Y., Han, Z.-M., Chai, Y.-R., Zhang, J.-H., 2010. A mini review of MAR-binding proteins. *Mol. Biol. Rep.* 37, 3553–3560. <https://doi.org/10.1007/s11033-010-0003-8>

Welzel, N., Le, T., Marculescu, R., Mitterbauer, G., Chott, A., Pott, C., Kneba, M., Du, M.Q., Kusec, R., Drach, J., Raderer, M., Mannhalter, C., Lechner, K., Nadel, B., Jaeger, U., 2001. Templated nucleotide addition and immunoglobulin JH-gene utilization in t(11;14) junctions: implications for the mechanism of translocation and the origin of mantle cell lymphoma. *Cancer Res.* 61, 1629–1636.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., Yahata, K., Imamoto, F., Aburatani, H., Nakao, M., Imamoto, N., Maeshima, K., Shirahige, K., Peters, J.-M., 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* 451, 796–801. <https://doi.org/10.1038/nature06634>

Williams, A., Spilianakis, C.G., Flavell, R.A., 2010. Interchromosomal association and gene regulation in trans. *Trends Genet.* TIG 26, 188–197. <https://doi.org/10.1016/j.tig.2010.01.007>

Woo, C.J., Martin, A., Scharff, M.D., 2003. Induction of somatic hypermutation is associated with modifications in immunoglobulin variable region chromatin. *Immunity* 19, 479–489. [https://doi.org/10.1016/s1074-7613\(03\)00261-9](https://doi.org/10.1016/s1074-7613(03)00261-9)

Wu, H., Deng, Y., Feng, Y., Long, D., Ma, K., Wang, X., Zhao, M., Lu, L., Lu, Q., 2018. Epigenetic regulation in B-cell maturation and its dysregulation in autoimmunity. *Cell. Mol. Immunol.* 15, 676–684. <https://doi.org/10.1038/cmi.2017.133>

Wuerffel, R., Wang, L., Grigera, F., Manis, J., Selsing, E., Perlot, T., Alt, F.W., Cogne, M., Pinaud, E., Kenter, A.L., 2007. S-S Synapsis during Class Switch Recombination Is Promoted by Distantly Located Transcriptional Elements and Activation-Induced Deaminase. *Immunity* 27, 711–722. <https://doi.org/10.1016/j.immuni.2007.09.007>

Wutz, G., Várnai, C., Nagasaka, K., Cisneros, D.A., Stocsits, R.R., Tang, W., Schoenfelder, S., Jessberger, G., Muhar, M., Hossain, M.J., Walther, N., Koch, B., Kueblbeck, M., Ellenberg, J., Zuber, J., Fraser, P., Peters, J.-M., 2017. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *EMBO J.* 36, 3573–3599. <https://doi.org/10.15252/embj.201798004>

- X -

Xiang, Y., Park, S.-K., Garrard, W.T., 2013. V $\kappa$  Gene Repertoire and Locus Contraction Are Specified by Critical DNase I Hypersensitive Sites within the V $\kappa$ -J $\kappa$  Intervening Region. *J. Immunol.* 190, 1819–1826. <https://doi.org/10.4049/jimmunol.1203127>

Xiang, Y., Zhou, X., Hewitt, S.L., Skok, J.A., Garrard, W.T., 2011. A Multifunctional Element in the Mouse *Igk* Locus That Specifies Repertoire and *Ig* Loci Subnuclear Location. *J. Immunol.* 186, 5356–5366. <https://doi.org/10.4049/jimmunol.1003794>

Xu, Y., Davidson, L., Alt, F.W., Baltimore, D., 1996. Deletion of the Ig $\kappa$  Light Chain Intronic Enhancer/Matrix Attachment Region Impairs but Does Not Abolish V $\kappa$ J $\kappa$  Rearrangement. *Immunity* 4, 377–385. [https://doi.org/10.1016/S1074-7613\(00\)80251-4](https://doi.org/10.1016/S1074-7613(00)80251-4)

Xu, Z., Fulop, Z., Wu, G., Pone, E.J., Zhang, J., Mai, T., Thomas, L.M., Al-Qahtani, A., White, C.A., Park, S.-R., Steinacker, P., Li, Z., Yates, J., Herron, B., Otto, M., Zan, H., Fu, H., Casali, P., 2010. 14-3-3 adaptor proteins recruit AID to 5'-AGCT-3'-rich switch regions for class switch recombination. *Nat. Struct. Mol. Biol.* 17, 1124–1135. <https://doi.org/10.1038/nsmb.1884>

Xu, Z., Zan, H., Pone, E.J., Mai, T., Casali, P., 2012. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat. Rev. Immunol.* 12, 517–531. <https://doi.org/10.1038/nri3216>

- Y -

Yang, Y., Ghosn, E.E.B., Cole, L.E., Obukhanych, T.V., Sadate-Ngatchou, P., Vogel, S.N., Herzenberg, Leonard A., Herzenberg, Leonore A., 2012. Antigen-specific memory in B-1a and its relationship to natural immunity. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5388–5393. <https://doi.org/10.1073/pnas.1121627109>



Yewdell, W.T., Kim, Y., Chowdhury, P., Lau, C.M., Smolkin, R.M., Belcheva, K.T., Fernandez, K.C., Cols, M., Yen, W.-F., Vaidyanathan, B., Angeletti, D., McDermott, A.B., Yewdell, J.W., Sun, J.C., Chaudhuri, J., 2020. A Hyper-IgM Syndrome Mutation in Activation-Induced Cytidine Deaminase Disrupts G-Quadruplex Binding and Genome-wide Chromatin Localization. *Immunity* 53, 952-970.e11. <https://doi.org/10.1016/j.immuni.2020.10.003>

Yokota, T., Kanakura, Y., 2014. Role of tissue-specific AT-rich DNA sequence-binding proteins in lymphocyte differentiation. *Int. J. Hematol.* 100, 238–245. <https://doi.org/10.1007/s12185-014-1602-2>

Yoshida, T., Yao-Ming Ng, S., Zuniga-Pflucker, J.C., Georgopoulos, K., 2006. Early hematopoietic lineage restrictions directed by Ikaros. *Nat. Immunol.* 7, 382–391. <https://doi.org/10.1038/ni1314>

Yu, K., Chedin, F., Hsieh, C.-L., Wilson, T.E., Lieber, M.R., 2003. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat. Immunol.* 4, 442–451. <https://doi.org/10.1038/ni919>

- Z -

Zhang, X., Yoon, H.S., Chapdelaine-Williams, A.M., Kyritsis, N., Alt, F.W., 2021. Physiological role of the 3'IgH CBEs super-anchor in antibody class switching. *Proc. Natl. Acad. Sci.* 118, e2024392118. <https://doi.org/10.1073/pnas.2024392118>

Zhang, X., Zhang, Y., Ba, Z., Kyritsis, N., Casellas, R., Alt, F.W., 2019. Fundamental roles of chromatin loop extrusion in antibody class switching. *Nature* 575, 385–389. <https://doi.org/10.1038/s41586-019-1723-0>

Zhang, Y., McCord, R.P., Ho, Y.-J., Lajoie, B.R., Hildebrand, D.G., Simon, A.C., Becker, M.S., Alt, F.W., Dekker, J., 2012. Spatial organization of the mouse genome and its role in recurrent chromosomal translocations. *Cell* 148, 908–921. <https://doi.org/10.1016/j.cell.2012.02.002>

Zhang, Y., Zhang, X., Ba, Z., Liang, Z., Dring, E.W., Hu, H., Lou, J., Kyritsis, N., Zurita, J., Shamim, M.S., Presser Aiden, A., Lieberman Aiden, E., Alt, F.W., 2019. The fundamental role of chromatin loop extrusion in physiological V(D)J recombination. *Nature* 573, 600–604. <https://doi.org/10.1038/s41586-019-1547-y>

Zhang, Z.Z., Pannunzio, N.R., Han, L., Hsieh, C.-L., Yu, K., Lieber, M.R., 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. <https://doi.org/10.1016/j.celrep.2014.06.021>

Zhuang, Y., Soriano, P., Weintraub, H., 1994. The helix-loop-helix gene E2A is required for B cell formation. *Cell* 79, 875–884. [https://doi.org/10.1016/0092-8674\(94\)90076-0](https://doi.org/10.1016/0092-8674(94)90076-0)

Zuin, J., Dixon, J.R., van der Reijden, M.I.J.A., Ye, Z., Kolovos, P., Brouwer, R.W.W., van de Corput, M.P.C., van de Werken, H.J.G., Knoch, T.A., van IJcken, W.F.J., Grosveld, F.G., Ren, B., Wendt, K.S., 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci.* 111, 996–1001. <https://doi.org/10.1073/pnas.1317788111>

## *Annexes*

---

## Annexes

---

Annexe 1. The IgH E $\mu$ -MAR regions promote UNG-dependent error-prone repair to optimize somatic hypermutation.....	125
Annexe 2. A dual function for the chromatin organizer Special A-T rich Binding Protein 1 in B-lineage cells .....	126

**Annexe 1. The IgH E $\mu$ -MAR regions promote UNG-dependent error-prone repair to optimize somatic hypermutation**

Ophélie Alyssa Martin, Morgane Thomas, Marie Marquet, Charlotte Bruzeau, Armand Garot, Mylène Brousse, Sébastien Bender, Claire Carrion, Jee Eun Choi, Bao Q. Vuong, Patricia J. Gearhart, Robert W. Maul, Sandrine Le Noir and Eric Pinaud

Front. Immunol. 14:1030813.doi: 10.3389/fimmu.2023.1030813



## OPEN ACCESS

## EDITED BY

Paolo Casali,  
The University of Texas Health Science  
Center at San Antonio, United States

## REVIEWED BY

Uttiya Basu,  
Columbia University, United States  
Jayanta Chaudhuri,  
Memorial Sloan Kettering Cancer Center,  
United States  
Frederick Alt,  
Harvard Medical School, United States

## \*CORRESPONDENCE

Eric Pinaud

✉ [eric.pinaud@unilim.fr](mailto:eric.pinaud@unilim.fr)

Sandrine Le Noir

✉ [sandrine.le-noir@unilim.fr](mailto:sandrine.le-noir@unilim.fr)

## †PRESENT ADDRESS

Ophélie A. Martin,  
Genome Damage and Stability Centre,  
School of Life Sciences, University of  
Sussex, Brighton, United Kingdom

†These authors have contributed equally to  
this work

## SPECIALTY SECTION

This article was submitted to  
B Cell Biology,  
a section of the journal  
Frontiers in Immunology

RECEIVED 29 August 2022

ACCEPTED 13 January 2023

PUBLISHED 14 February 2023

## CITATION

Martin OA, Thomas M, Marquet M,  
Bruzeau C, Garot A, Brousse M, Bender S,  
Carrion C, Choi JE, Vuong BQ,  
Gearhart PJ, Maul RW, Le Noir S and  
Pinaud E (2023) The IgH  $E\mu$ -MAR regions  
promote UNG-dependent error-prone  
repair to optimize somatic hypermutation.  
*Front. Immunol.* 14:1030813.  
doi: 10.3389/fimmu.2023.1030813

## COPYRIGHT

© 2023 Martin, Thomas, Marquet, Bruzeau,  
Garot, Brousse, Bender, Carrion, Choi,  
Vuong, Gearhart, Maul, Le Noir and Pinaud.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# The IgH $E\mu$ -MAR regions promote UNG-dependent error-prone repair to optimize somatic hypermutation

Ophélie A. Martin<sup>1†</sup>, Morgane Thomas<sup>1†</sup>, Marie Marquet<sup>1†</sup>,  
Charlotte Bruzeau<sup>1</sup>, Armand Garot<sup>1</sup>, Mylène Brousse<sup>1</sup>,  
Sébastien Bender<sup>1,2</sup>, Claire Carrion<sup>1</sup>, Jee Eun Choi<sup>3</sup>,  
Bao Q. Vuong<sup>3</sup>, Patricia J. Gearhart<sup>4</sup>, Robert W. Maul<sup>4</sup>,  
Sandrine Le Noir<sup>1\*</sup> and Eric Pinaud<sup>1\*</sup>

<sup>1</sup>Laboratoire Contrôle de la Réponse Immune B et des Lymphoproliférations (CRIBL), Université de Limoges, CNRS Unité Mixte de Recherche 7276, Inserm Unité 1262, Limoges, France, <sup>2</sup>Centre Hospitalier Universitaire Dupuytren, Service d'Immunopathologie, Limoges, France, <sup>3</sup>The Graduate Center, The City University of New York, New York, NY, United States, <sup>4</sup>Laboratory of Molecular Biology and Immunology, National Institute on Aging, National Institutes of Health, Baltimore, MD, United States

**Introduction:** Two scaffold/matrix attachment regions (5'- and 3'- $MAR_{E\mu}$ ) flank the intronic core enhancer ( $cE\mu$ ) within the immunoglobulin heavy chain locus ( $IgH$ ). Besides their conservation in mice and humans, the physiological role of  $MAR_{E\mu}$  is still unclear and their involvement in somatic hypermutation (SHM) has never been deeply evaluated.

**Methods:** Our study analyzed SHM and its transcriptional control in a mouse model devoid of  $MAR_{E\mu}$ , further combined to relevant models deficient for base excision repair and mismatch repair.

**Results:** We observed an inverted substitution pattern in of  $MAR_{E\mu}$ -deficient animals: SHM being decreased upstream from  $cE\mu$  and increased downstream of it. Strikingly, the SHM defect induced by  $MAR_{E\mu}$ -deletion was accompanied by an increase of sense transcription of the IgH V region, excluding a direct transcription-coupled effect. Interestingly, by breeding to DNA repair-deficient backgrounds, we showed that the SHM defect, observed upstream from  $cE\mu$  in this model, was not due to a decrease in AID deamination but rather the consequence of a defect in base excision repair-associated unfaithful repair process.

**Discussion:** Our study pointed out an unexpected "fence" function of  $MAR_{E\mu}$  regions in limiting the error-prone repair machinery to the variable region of Ig gene loci.

## KEYWORDS

B cell, immunoglobulin gene, MARs region, somatic hypermutation (SHM), UNG

## Introduction

The *IgH* locus, encoding the immunoglobulin heavy chain, is among the most complex in mammals, with multiple *cis*-regulatory elements controlling stepwise DNA accessibility to recombination and mutation through mechanisms that mainly rely on transcription (1). Current studies of the dynamic processes that regulate chromatin conformation changes and subnuclear location have renewed interest in *cis*-regulatory regions that delimit differentially regulated chromosomal domains. Among such DNA regulatory regions, nuclear Scaffold/Matrix Attachment Regions (MARs) have been implicated in the structural and functional organization of these domains. The juxtaposition of MARs to intronic enhancer elements in both *IgH* and *IgL* loci and their conservation in humans, mice and rabbits (2) suggest that such regions serve physiological functions. They participate in the regulation of gene expression notably by increasing enhancer function and facilitating their action over large distances. Several proteins found to bind MARs are expressed ubiquitously or in a tissue-specific manner, respectively defining constitutive or facultative MARs (3). Once attached to the nuclear matrix in a tissue specific fashion, facultative MARs could form topological barriers that could isolate or fasten chromatin regions (3). Such barriers could induce DNA torsional strain with positive and negative DNA supercoiling, respectively, upstream and downstream from the RNA pol II-induced transcription bubble (4). The supercoils are then released by the action of dedicated topoisomerases (5).

The *IgH E $\mu$*  enhancer region is a combination of both the core *E $\mu$*  (*cE $\mu$* ) enhancer element (220 bp) and two 310–350-bp flanking MARs (*MARs<sub>E $\mu$</sub>* ) that were first defined by *in vitro* matrix-binding assays (6). This region, especially *cE $\mu$* , controls early VDJ recombination events (7, 8) and is also involved in Ig  $\mu$  chain expression in pre-B cells (9). However, its role in SHM remains unclear. An elegant model of deletion in the endogenous *E $\mu$*  region of hybridoma cells, enforced for human AID expression, suggested the requirement of *cE $\mu$*  and a substantial function of *MARs<sub>E $\mu$</sub>*  for SHM (10). Similarly, when added to transgenes, *cE $\mu$*  and its flanking MARs contribute to Ig  $\mu$  chain expression and high levels of SHM (11–16). In contrast, knock out (KO) models underlined the complexity of its physiological regulation. In a mouse model carrying the pre-rearranged VB1-8i region, *E $\mu$*  deletion still resulted in a high level of SHM in Peyer's patch germinal center (GC) B cells, arguing for a non-essential role of the enhancer (17). More clearly, deletion of *cE $\mu$*  in the mouse germline did not reduce SHM frequency but only slightly increased the proportion of unmutated alleles; this minor effect was likely due to the reduced inflow of peripheral and, consequently, GC B cells in this model (7). Strikingly, the role of *MARs<sub>E $\mu$</sub>*  was also elusive and somewhat controversial. Whereas their endogenous deletion, analysed in mouse chimeras by the RAG-2 complementation assay, demonstrated that *MARs<sub>E $\mu$</sub>*  are dispensable for VDJ recombination and IgH expression (18), the ambiguous function of *MARs<sub>E $\mu$</sub>*  was sustained by the discrepancy between their ability to either bind negative regulatory factors (19, 20), improve *cE $\mu$*  enhancer efficiency (21), or substitute for *cE $\mu$*  to maintain IgH expression (22). At the  $\kappa$  light chain locus (*Ig $\kappa$* ), the intronic enhancer *Ei $\kappa$*  region also contains an upstream MAR. The implication of *MAR<sub>Ei $\kappa$</sub>*  as an enhancer of SHM was first suggested in transgenic studies (23) and then tested in a

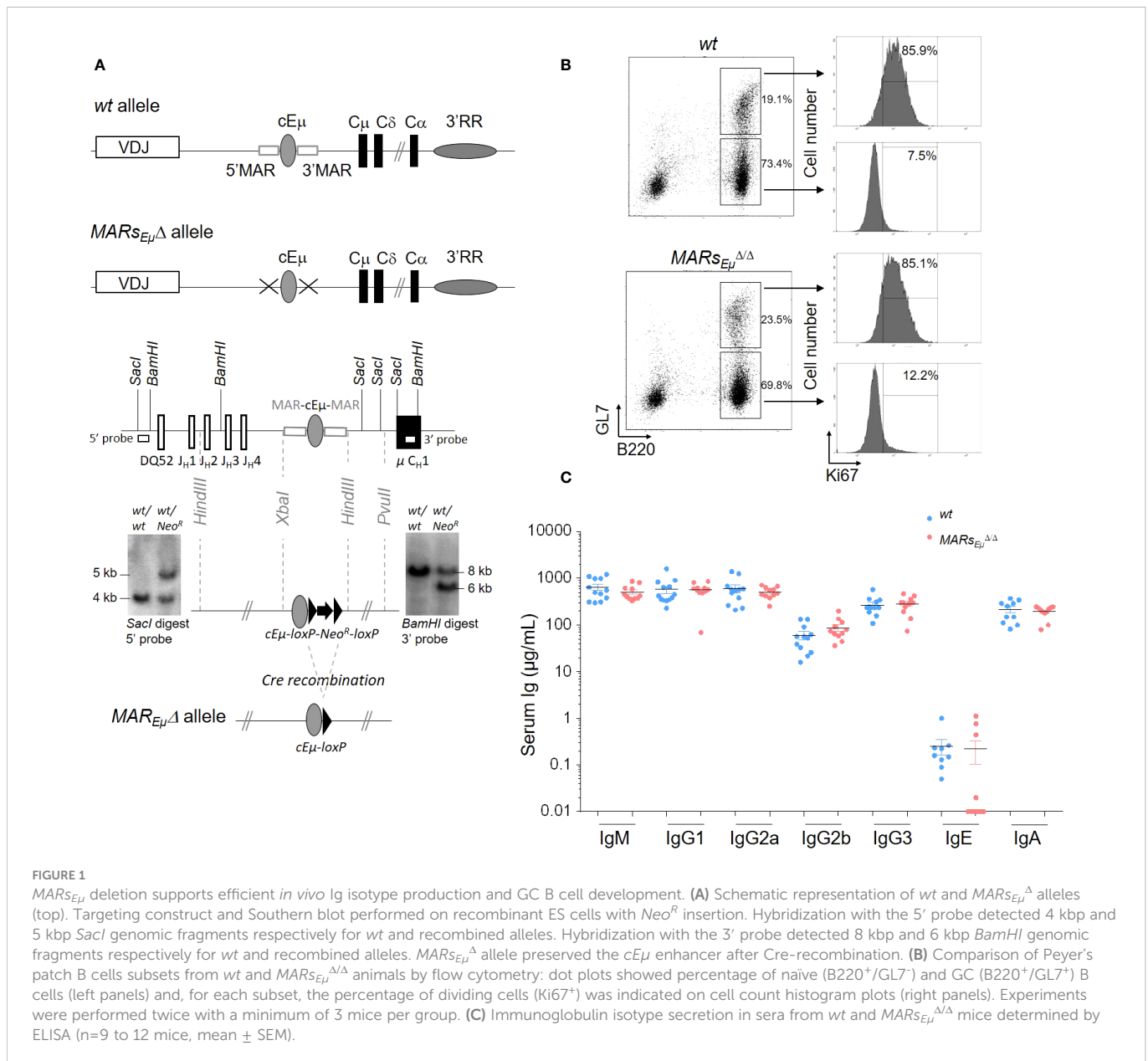
KO mouse model that accumulated premature light chain rearrangements with a mild SHM defect (24), an effect comparable to one observed at *IgH* locus in hybridoma cells devoid of *MARs<sub>E $\mu$</sub>*  (10). At that time, while these studies instigated a variety of hypotheses accounting for MARs in modulating SHM (25), these were contradicted by a study comparing 3'*E $\kappa$* - and *MAR<sub>Ei $\kappa$</sub>* - function in mouse KO models (26).

To address the controversy over the role of the scaffold in SHM, we generated a mouse model carrying a germline deletion of *MARs<sub>E $\mu$</sub>*  and bred it into DNA repair-deficient backgrounds. In our models devoid of *MARs<sub>E $\mu$</sub>*  and their *wt* counterparts, we proceeded to side by side comparison of total SHM, transcription patterns, AID targeting and error prone repair events leading to SHM, in regions located upstream and downstream from the intronic enhancer. Our study showed that the absence of *MARs<sub>E $\mu$</sub>*  allows some of the error-prone repair machinery to get access to the region downstream from the *E $\mu$*  enhancer. We propose that *MARs<sub>E $\mu$</sub>*  act as physiological barriers for error-prone repair in activated B cells. As a rational hypothesis, our study suggests that the conservation of nuclear matrix attachment regions in Ig genes serves to optimize SHM events upstream from the intronic enhancer region.

## Results

### Normal B cell development and Ig production in the absence of *MARs<sub>E $\mu$</sub>*

We generated a mouse mutant line carrying an endogenous deletion of both the 5' and 3' *IgH* matrix attachment regions that flank the *J<sub>H</sub>-C<sub>H</sub>* intronic *cE $\mu$*  enhancer. Although generated with slightly different targeting vector backbone and homology arms, the resulting *IgH* allele, so-called *MARs<sub>E $\mu$</sub> <sup>Δ</sup>* (Figure 1A), is similar to that generated by Sakai et al. (18). Bone marrow subsets of B cell precursors were analysed in *wt* and homozygous *MARs<sub>E $\mu$</sub> <sup>Δ/Δ</sup>* deficient mice. When compared to age-matched *wt* animals, *MARs<sub>E $\mu$</sub> <sup>Δ/Δ</sup>* mice exhibited normal proportions and numbers of pre-proB, pro-B and pre-B cell precursors (Supplementary Table 1). Unlike endogenous deletion of the entire *E $\mu$*  region (9), *MARs<sub>E $\mu$</sub>*  deletion did not modify Ig  $\mu$  heavy chain expression in early B lineage cells since proportions of IgM-expressing bone marrow B cell populations (immature, transitional and mature recirculating B cell subsets) were comparable to those of *wt* (Supplementary Table S1 and Supplementary Figure S1A). Mature B cell subsets were also similar to *wt* in the spleen and peritoneal cavity of homozygous *MARs<sub>E $\mu$</sub> <sup>Δ/Δ</sup>* mutants (Supplementary Table S1 and Supplementary Figures S1B, C). In agreement with the normal inflow of mature B cells in *MARs<sub>E $\mu$</sub> <sup>Δ/Δ</sup>* animals, Peyer's patches were efficiently colonized by naive and GC B cells. Numbers and proportions of GC B cells were even significantly increased in homozygous mutants (Figure 1B left panels and Supplementary Table S1). The similar proportion of proliferating KI67<sup>+</sup> GC B cells in *wt* and *MARs<sub>E $\mu$</sub> <sup>Δ/Δ</sup>* mice implied that this increase was not due to over-proliferation of Peyer's patch B cells (Figure 1B right panels). Finally, levels of serum Ig isotypes were unaffected in *MARs<sub>E $\mu$</sub> <sup>Δ/Δ</sup>* animals (Figure 1C). This normal B cell homeostasis in homozygous mutants confirmed that *MARs<sub>E $\mu$</sub>*  are dispensable for B cell ontogeny and antibody production. This



statement is in agreement with previous studies of an analogous MAR region in the *Ig $\kappa$*  locus (24).

### *MARSE $\mu$* deletion inverts SHM distribution on both sides of the *E $\mu$* enhancer region

To assess whether *MARSE $\mu$*  deletion could affect *IgH* somatic hypermutation, we first quantified mutations within the 500-bp regions downstream from the variable exons rearranged to *J<sub>H</sub>3* and *J<sub>H</sub>4* segments (Figure 2A) in Peyer's patch GC B cells sorted from *wt* and *MARSE $\mu$  $\Delta/\Delta$*  (animals overall data reported in Figure 2 left, data from individual animals reported in Supplementary Figure S2 and Supplementary Tables S2A, B). For this we used two complementary sequencing methods: the first one, based on classical Sanger approach and GS junior technology, allowed to discriminate and exclude unmutated and clonally related sequences from the calculation of

SHM frequency, as initially described (27). The second method used Ion proton deep sequencing coupled to DeMinEr filtering. Since using DNA templates including non-mutated alleles, this second approach underestimated the SHM frequency; but since including AID-deficient control samples as a reference, the method provided highly reproducible and reliable quantification of SHM in a DNA sample extracted from GC B cells (28). Interestingly, by using Sanger approach, *MARSE $\mu$  $\Delta/\Delta$*  GC B cells displayed significant differences in the distribution of mutations: an increased proportion of unmutated sequences (less than 10% in *wt* compared to 38% in *MARSE $\mu$  $\Delta/\Delta$* ) (from 30.6 to 45.5%, overall data collected from several mice, data from independent mice in Supplementary Figure S2A) Another effect of *MARSE $\mu$*  deletion on *IgH* SHM targeting was the strong decrease in highly mutated sequences (>10bp per sequence). In *wt*, the proportion of highly mutated alleles reached ~24% (from 18.9 to 28.3%) while in mutants they were barely present (~2%) (Figure 2B left and Supplementary Fig S2A). When comparing only the mutated

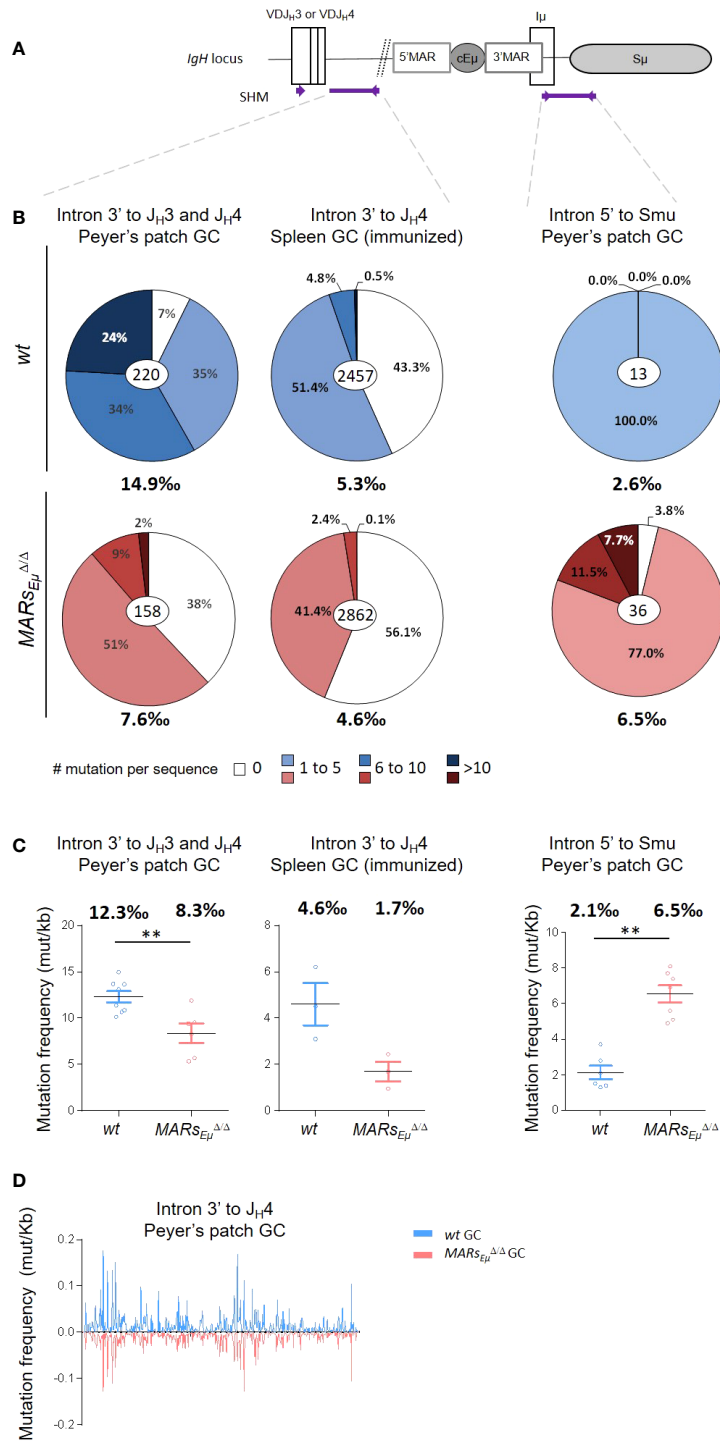


FIGURE 2

MARs<sub>Eμ</sub> deletion impairs the overall SHM frequency and distribution within the *IgH* J-C intronic region. **(A)** Location of *IgH* regions (thick purple lines) tested for SHM, arrows represent primers used for PCR amplification. **(B)** Pie charts represent distribution of mutated sequences (proportional to the area in each slice, data obtained by Sanger and GS Junior sequencing method) quantified in wt and MARs<sub>Eμ</sub><sup>ΔΔ</sup> mice in individually recombined *IgH* alleles. For each genotype number of individual clones is indicated in the center (after removal of clonally related sequences based on VDJ junction) and overall mutation frequencies (mutation per 1000 bp in mutated clones) are indicated below. Left: SHM downstream from J<sub>H3</sub> and J<sub>H4</sub> segments in Peyer's patch sorted GC B cells, data obtained after cloning and sequencing by classical Sanger method. Middle: SHM downstream from J<sub>H4</sub> segments in spleen GC B cells sorted from SRBC-immunized mice, data obtained by NGS (GS Junior). Right: SHM downstream from cEμ region from Peyer's patch GC sorted B cells, data obtained by classical Sanger method. **(C)** Graphical representation of SHM frequency in wt and MARs<sub>Eμ</sub><sup>ΔΔ</sup> mice, quantified by NGS (Ion Proton) submitted to DeMinEr filtering, a pipeline that identifies substitution frequency at each nucleotide based on an *Aicda*<sup>ΔΔ</sup> control sample (28). Since no indication in sequence distribution is available using this method, data were represented as scattered plots, each point refers to a mutation frequency from one individual mice, mean mutation frequencies are indicated above. p-value was determined with two tailed Mann Whitney test; significant differences are indicated by: \*\*P < 0.01 and error bars represent SEM of two independent experiments. **(D)** Mutation distribution along the J<sub>H4</sub> intron in wt (top) and in MARs<sub>Eμ</sub><sup>ΔΔ</sup> (bottom).



sequences, mutation frequency was decreased at least by two fold in  $MARs_{E\mu}^{\Delta/\Delta}$  mutants, with 7.6 mutations per 1000 bp (in average) compared to 14.9 in *wt* (in average) (Figure 2B left and Supplementary Figure S2A). By using next generation sequencing (NGS), the decreased SHM frequency was also highly significant (Figure 2C, 12.3‰ vs 8.3‰,  $p=0.008$ , individual mice in Supplementary Table S2A).

To monitor SHM upon antigen challenge, we analyzed mutations by Sanger and NGS in a large number of GC B cells sorted from spleen of SRBC-immunized  $MARs_{E\mu}^{\Delta/\Delta}$  and *wt* mice. In parallel; we verified on spleen frozen section that efficient GC formation was indeed comparable between  $MARs_{E\mu}^{\Delta/\Delta}$  and *wt* mice (Supplementary Figure S2B). In the intronic region downstream from the  $J_H4$  segment, SHM frequency dropped from 5.3 (Sanger method, excluding unmutated clones) or 4.6 (NGS bulk method) mutations per 1000bp in *wt* cells to respectively 4.6 or 1.7 in  $MARs_{E\mu}^{\Delta/\Delta}$  cells (Figure 2B middle, and 2C, Supplementary Figure S2C, Supplementary Table S2B). Similarly to what was observed in spontaneous GC B cells from Peyer's patches, NGS assays performed on few immunized animals also showed that SHM was decreased upon  $MARs_{E\mu}$ -deletion only eight days after antigen challenge (Figure 2C middle). Immunization assays also showed an increased proportion of unmutated or poorly mutated sequences in  $MARs_{E\mu}$ -deficient B cells (Figure 2B and Supplementary Figure S2C). This confirmed that the intronic region was less efficiently targeted by SHM in  $MARs_{E\mu}$  deficient mice. An identical SHM defect was also observed in mice harbouring deletion of the entire  $E\mu$  region (core enhancer and flanking MARs; (9)) (Supplementary Figure S2A top right); this data indicated that the SHM failure was the only consequence of MARs deletion. This hypothesis is completely consistent with a previous study showing that SHM efficiency was not affected by the endogenous deletion of the  $cE\mu$  enhancer alone (7). The comparison between those three models is certainly relevant since all knock outs were created in murine germlines with similar mixed genetic backgrounds.

To reinforce the unbiased SHM evaluation within  $J_H$  introns, we sought to evaluate the impact of  $MARs_{E\mu}$  deletion on acquired diversity of the VDJ exon by quantifying SHM events taking place in the out of frame (passenger) IgH alleles. This parallel unbiased analysis was made possible in a region physiologically highly targeted by SHM (overlapping parts of the FR3 and CDR3 domains) thanks to the IMGT/HighV-QUEST reference tool (29). Although focused on a restricted region of out of frame VDJ exons, this study performed on GC cells sorted from Peyer's patches still highlighted a decreased SHM frequency in the  $MARs_{E\mu}^{\Delta/\Delta}$  model (~27‰ in mutants compared to ~32‰ in *wt*, Supplementary Table S3). This strongly suggests that the SHM defect observed within  $J_H$  introns in the absence of  $MARs_{E\mu}$  extends to their respective upstream VDJ exons.

Since our  $MARs_{E\mu}$  deletion includes the 3'*Hinfl-XbaI* genomic region that contains transcription start sites and part of the  $I\mu$  exon (6, 30), we also quantified SHM immediately downstream from this exon in a 629bp region described as mutated in GC B cells (Figure 2A) (31, 32). Unlike the intronic regions downstream from the rearranged VDJ exon, the overall mutation frequency downstream from  $I\mu$  was strongly increased in GC B cells devoid of  $MARs_{E\mu}$  region and reached 6.5 (in average) mutations per 1000 bp compared to 2.6 (in average) in *wt* cells (Figure 2B right). This suggests that the region downstream

from the  $cE\mu$  was more efficiently targeted in the absence of its MARs. This was supported by the very low proportion of unmutated sequences in *wt* (Figure 2B right and Supplementary Figure S3A) and the increased proportion of highly mutated sequences in  $MARs_{E\mu}$ -deficient GC B cells (Figure 2B right and Supplementary Figure S3A). This data was efficiently confirmed by NGS analysis that estimated that mutation frequency was increased by 3.5 fold in the  $MARs_{E\mu}$ -deficient GC B compared to *wt* mice (2.1‰ vs 6.5‰) (Figure 2C right and Supplementary Table S2C). In our mouse model, 3'MARs deletion brought  $S\mu$  region closer to the  $cE\mu$  enhancer element. Given this, we also calculated SHM frequency in a 473bp region located at the same distance from the  $cE\mu$  element in *wt* and  $MARs_{E\mu}$ -deficient mice. Similarly to what was reported in Figure 2, we also observed, at the same distance of  $cE\mu$ , an increase in SHM frequency upon  $MARs_{E\mu}$ -deletion (Supplementary Figure S3B).

Analysis of mutation distribution in *wt* and in  $MARs_{E\mu}^{\Delta/\Delta}$  mice did not show any difference between models (Figure 2D), indicating that, while affecting SHM efficiency, the absence of  $MARs_{E\mu}$  region did not influence DNA sequence hotspot or preferences for SHM within the  $J_H4$  intron.

Importantly, our mouse model clearly assigns a specific function for endogenous  $MARs_{E\mu}$  on SHM at the *IgH* locus, in accord with the requirement of similar regions for efficient SHM previously pointed out in the endogenous *Igk* Kappa light chain locus. This pioneer study, describing the specific deletion of a 420bp MAR region upstream from the intronic Kappa enhancer (*Ei* $\kappa$ ), highlighted a modest decrease in SHM by quantifying mutations downstream from the  $J_{\kappa5}$  segment in GC B cells from Peyer's patches (24). While our study suggests that  $MARs_{E\mu}$  optimizes SHM upstream from the  $cE\mu$  enhancer; the presence of such regulatory regions does not prevent the SHM machinery to get access to downstream regions as reported in a recent study (33). This hypothesis is mostly supported by the increased SHM frequency downstream from the  $cE\mu$  enhancer in the absence of  $MARs_{E\mu}$ , a finding consistent with previous works describing increased  $S\mu$  internal deletions in hybridomas devoid of MARs regions (34). We could speculate that one physiological function of  $MARs_{E\mu}$  regions in GC B cells is to tightly isolate the VDJ transcription unit by, at least temporarily, attaching the  $E\mu$  region to the nuclear matrix. Such a "locked" target conformation could provide an optimal environment for somatic mutations by trapping the transcription machinery and its co-factors including AID and error-prone repair factors. This topological barrier could, at the same time, partially protect downstream constant regions from SHM; although this configuration should be brief since regions downstream from  $E\mu$  are also efficiently targeted by AID in GC B cells (35).

## $MARs_{E\mu}$ deletion modifies transcription patterns on both sides of the $E\mu$ enhancer region

It is well established that SHM in Ig V segments is coupled to transcription initiated at V promoters (36). To investigate transcription-related events in SHM-targeted regions upstream and downstream from  $E\mu$ , we precisely quantified the total amounts of steady state total *IgH* primary transcripts by using multiple q-PCR

probes located respectively downstream from  $J_{H4}$  and  $J_{H3}$ : the previously described probe A (37) (Figure 3A and Supplementary Figure S4) complemented by probes A' and C (Supplementary Figures S4, S5). The use of cDNA templates conducted with random hexamers showed that the amount of total IgH primary transcripts running upstream from  $E\mu$  did not display significant variations between *wt* and  $MARs_{E\mu}$ -deficient cells in both GC and *in vitro*-stimulated samples by using probe A (Figure 3B) as well as with probes A' and C, although an upward trend could be noticed in LPS-activated samples (Supplementary Figure S5A left). The intriguing discrepancy between the mutation phenotype observed in  $MARs_{E\mu}$ -deficient GC B cells and the silent effect on global transcription motivated a more complete study of transcription events occurring upstream from  $E\mu$ , particularly sense and antisense transcription since the latter has been found in cells undergoing SHM (38). To proceed, we generated cDNA templates with sense transcripts, initiated at the promoter of the rearranged *VDJ* segment, with three primers located downstream from the  $J_{H4}$  segment (S1 and S2) and within the *cE\mu* enhancer (S3) (Figure 3C and Supplementary Figure S4, S5A). Reciprocally, we generated cDNA templates with antisense transcripts, initiated in the intronic regions upstream from  $E\mu$  as described by Perlot et al. (38), with four primers respectively located downstream from  $J_{H2}$  (AS0),  $J_{H3}$  (AS1) and  $J_{H4}$  (AS2 and AS3) (Figure 3E and Supplementary Figure S4 and S5B). For both sense and antisense, quantification of steady state transcripts was possible with the same probes A, A' and C. For strand-specific quantification assays with a given probe, the baseline level was either provided by a control reaction (P-) measuring endogenous priming since devoid of primer or by one strand-specific template that cannot be detected by the probe (T-) as reported previously (39, 40). To note, strand-specific transcripts were optimally detected when primers and probes were closer (sense transcripts with primer S1/probe A or antisense transcripts with primer AS2/probe A). Side by side comparison of *wt* and  $MARs_{E\mu}$ -deficient activated B cells samples revealed several interesting differences. When quantified with optimal primer S1/probe A tandem, sense transcripts were significantly increased in the absence of  $MARs_{E\mu}$  (Figure 3C). In GC B cells, a two fold increase was noticed by using S1 template (Figure 3C, left bar graph,  $p=0.019$ ). In *in vitro*-activated cells, an increase of sense transcription was also observed upon  $MARs_{E\mu}$ -deletion, this effect became significant for long transcripts that reach the *cE\mu* (Figure 3C, right bar graphs,  $p=0.004$  for S3/probe A). By using A' and C probes, sense transcripts were hardly detectable in GC samples (Supplementary Figure S5A; middle bar graphs); although a significant increase was noticed with S3/probe C ( $p=0.043$ ) and S2/probe C ( $p=0.02$ ) tandems in LPS-activated samples upon  $MARs_{E\mu}$ -deletion (Supplementary Figure S5A; right bar graphs). As a potential consequence of the increased transcription of the *VDJ* unit in observed upon  $MARs_{E\mu}$  deletion, we measured by flow cytometry the level of intracellular Ig $\mu$  chain in Peyer's patch GC B cells (B220<sup>+</sup>/GL7<sup>+</sup>) of *wt* and  $MARs_{E\mu}^{\Delta/\Delta}$  mice. The significant increase of Ig $\mu$  chain expression in the absence of  $MARs_{E\mu}$  region (Figure 3D,  $p=0.002$ ) corroborate our sense-transcription data. This indicated that the absence of  $MARs_{E\mu}$  certainly did not hamper RNA pol II machinery to progress 3' to the *VDJ* unit and might even facilitate this process in activated cells.

Globally less abundant than their sense counterparts, quantification of antisense transcripts running downstream from  $J_H$  segments showed quite different patterns (Figure 3E and Supplementary Figure S5B). While quite similar levels were detected in LPS-activated samples (Figure 3E right and Supplementary Figure S5B), intronic antisense transcripts were about 2 fold less abundant in  $MARs_{E\mu}$ -deficient GC B cells when detection was allowed by optimal probe/primer combination (Figure 3E, left,  $p=0.025$ , AS2/probe A and Supplementary Figure S5B left,  $p=0.041$  for AS3/probe A'). As a complementary experiment, sense and antisense transcription patterns were also determined in naive B cells in order to assess whether the influence of MARs was restricted to the activated stage. Data proved this was the case since, in naive cells sorted from Peyer's patches, transcription patterns as well as intracellular IgM levels were unchanged in *wt* and  $MARs_{E\mu}^{\Delta/\Delta}$  mice (Supplementary Figures S6A, B); identical transcription patterns were also observed in splenic resting B cells (Supplementary Figure S6C).

The obvious unbalanced sense/antisense transcription ratio could result from either weak transcription efficiency or instability of antisense products. Nevertheless, Perlot et al. identified by RACE assays, in normal GC B cells, multiple antisense-transcript initiation start sites downstream from every  $J_H$  region and raised the question of specific enhancers. Our current data refines this previous study by identifying  $MARs_{E\mu}$  as potential boosters of antisense transcripts that, given their proximity to the enhancer, could achieve some regulatory function like eRNA or PROMPT/uaRNA (41). Highlighting a correlation between mutation efficacy and strand-specific transcription pattern upstream from  $E\mu$ , our data support the idea that some level antisense transcription downstream from the *VDJ* exon could prepare to SHM (38). Seemingly transient, specific to cell subsets and occurring upstream from an enhancer, such antisense transcripts could be substrates for RNA exosome and lead to optimized SHM targeting as proposed by Basu and colleagues (42–44).

Since a strong increase of mutations was observed within the  $S\mu$  region in the absence of  $MARs_{E\mu}$ , we also sought to correlate SHM and transcription on the other side of *cE\mu* by quantifying total transcripts (probe B) running in this region (Figure 3A). In this case and according to what could be expected, transcription was significantly increased in  $MARs_{E\mu}$ -deficient GC B cells from Peyer's patches (Figure 3F left,  $p=0.04$ ); a similar trend, although not significant, was observed in LPS activated B cells (Figure 3B right  $p=0.32$ ). Accordingly, we also observed a modest but reproducible increase of CSR C $\gamma$ 3 and C $\gamma$ 1 in  $MARs_{E\mu}$ -deficient B cells stimulated *in vitro* respectively by LPS or by LPS + IL4 cocktail (Supplementary Figure S7). A similar modest CSR effect associated to  $S\mu$  internal deletions has been previously reported in hybridomas carrying the same  $MARs_{E\mu}$ -deletion (34). This indicated that the absence of  $MARs_{E\mu}$  leads to a global increase in transcription of the donor S region and consequently favours SHM targeting.

The significant changes in transcription patterns upstream and downstream from *cE\mu* observed in our models put forward the hypothesis that  $MARs_{E\mu}$  act as physiological barriers in activated B cells, limiting sense transcription of the *VDJ* unit up to the intronic enhancer. For transcription running through the  $S\mu$  region, our data is in agreement with a repressive function of  $MARs_{E\mu}$  in activated B

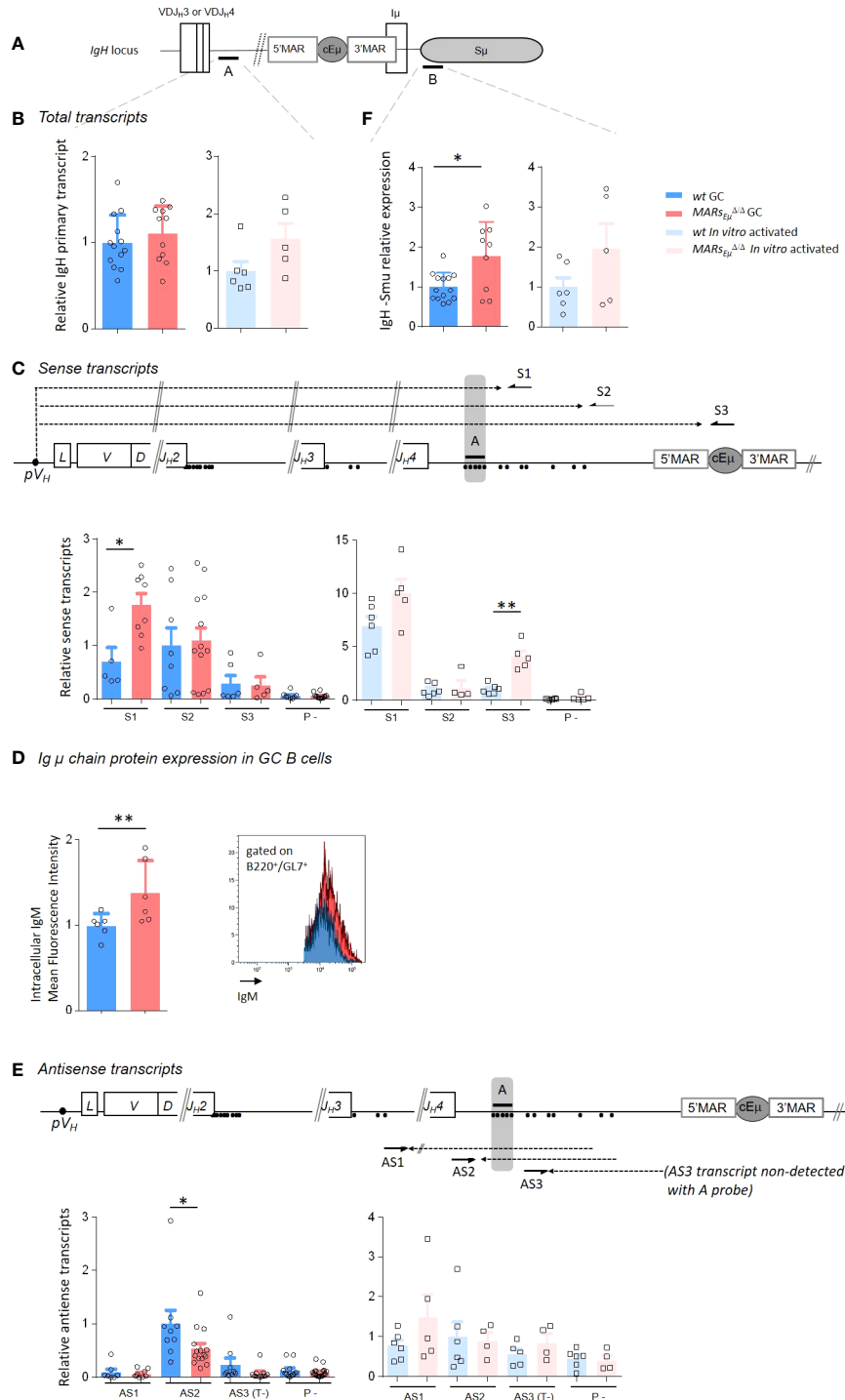


FIGURE 3

*MARS<sub>Eμ</sub>* deletion impairs strand-specific transcription upstream from *Eμ* region. **(A)** *IgH* locus specifying location of q-PCR probes **(A, B)** used for transcripts quantification. **(B)** Total primary transcripts quantified downstream from the *J<sub>H3</sub>* segment with probe A in Peyer's patch GC B cells (dark colors) and *in vitro*-activated B cells (light colors) from *wt* and *MARS<sub>Eμ</sub><sup>Δ/Δ</sup>* mice. **(C)** Detection of sense transcripts (dotted arrows) in murine *IgH* locus (not to scale). Arrows indicate primers (S1, S2, S3) downstream from *J<sub>H3</sub>* and *J<sub>H4</sub>* used for strand-specific reverse transcription. Primary sense transcripts were quantified with probe A (black bar) in Peyer's patch GC B cells and *in vitro*-activated B cells from *wt* and *MARS<sub>Eμ</sub><sup>Δ/Δ</sup>* mice. Dots indicate antisense transcript start sites according to Perlot et al. (38). Baseline levels were defined by using a RT template performed without primers (P-). Bar graphs show the relative quantity of sense transcripts obtained from template S1, S2 and S3 (mean ± SEM) from two to three independent experiments. **(D)** Intracellular IgM mean fluorescence intensity measured by flow cytometry in GC B cells from Peyer's patches of *wt* and *MARS<sub>Eμ</sub><sup>Δ/Δ</sup>* mice. Bar graphs indicate data from individual mice (*n*=6 mice in 2 independent experiments, mean ± SEM); a representative example of cell count overlay is shown. **(E)** Detection of antisense transcripts (dotted arrows) in murine *IgH* locus (not to scale). Arrows indicate primers (AS1, AS2, AS3) downstream from *J<sub>H3</sub>* and *J<sub>H4</sub>* used for strand-specific reverse transcription. Primary antisense transcripts were quantified with probe A (black bar) in Peyer's patch GC B (dark colors) cells and *in vitro*-activated (light colors) B cells from *wt* and *MARS<sub>Eμ</sub><sup>Δ/Δ</sup>* mice. Dots indicate antisense transcripts start sites according to the Alt study (38). Baseline levels were defined by using a RT template performed without primers (P-) or by using a strand-specific template that cannot be detected with A probe (T-). **(F)** Total primary transcripts quantified downstream from *cEμ* region with probe B in Peyer's patch GC B cells (dark colors) and *in vitro*-activated B cells (light colors) from *wt* and *MARS<sub>Eμ</sub><sup>Δ/Δ</sup>* mice. *p*-value was determined with two tailed Mann Whitney test; significant differences are indicated by: \**P* < 0.05; \*\**P* < 0.01 and error bars represent SEM of two to three independent experiments.

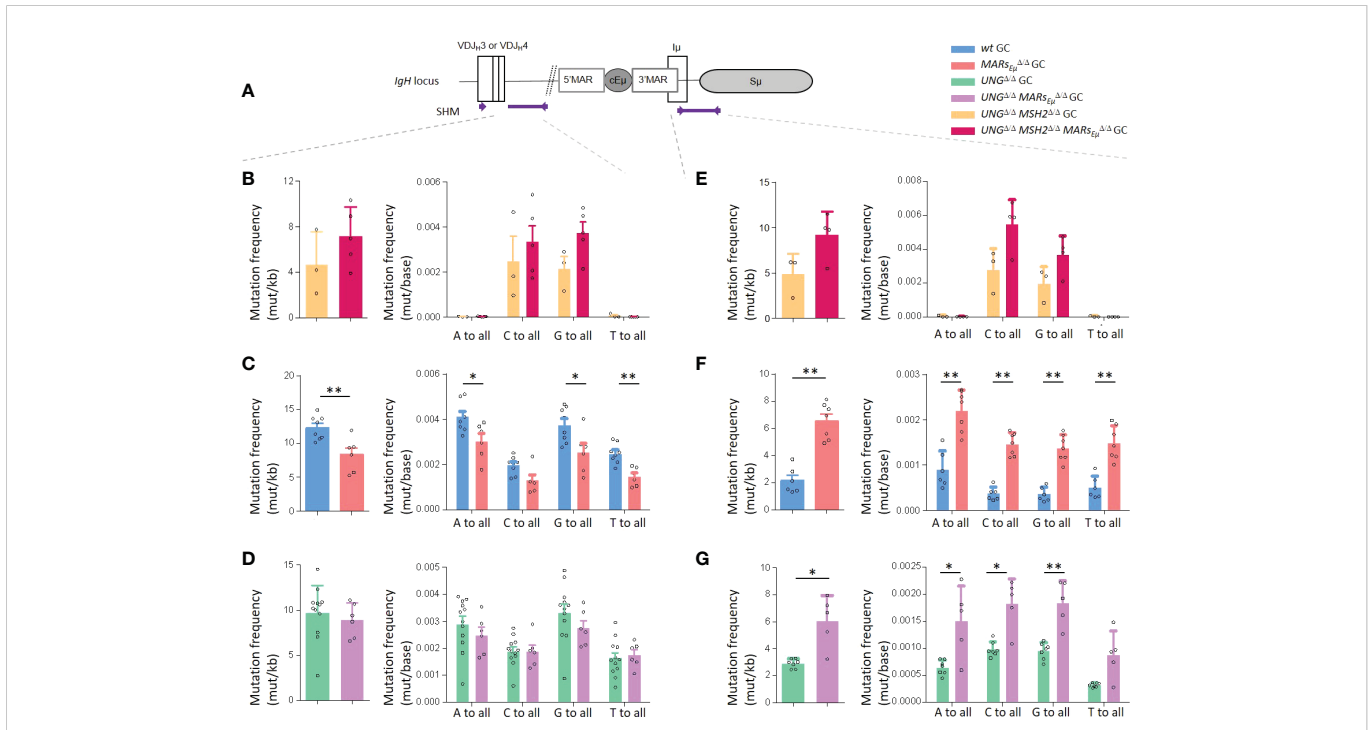
cells, in order to limit SHM targeting of this area. However, our data also suggest that *MARs<sub>Eμ</sub>* act as transcriptional repressors of the *VDJ* unit in both naïve and activated cells; a statement in contradiction with our hypothesis that *MARs<sub>Eμ</sub>* facilitates SHM upstream from *cEμ*. To settle such a discrepancy in our *MARs<sub>Eμ</sub>*-deficient B cells, we first questioned AID deamination efficiency and second error-prone repair pathways processing in SHM targeted regions: upstream and downstream from *cEμ*.

### *MARs<sub>Eμ</sub>* deletion impairs error-prone repair pathway upstream from the *Eμ* enhancer region

One critical experiment needed to challenge the function of *MARs<sub>Eμ</sub>* as physiological barrier for SHM machinery was to first assess whether IgH AID targeting could be impaired in the absence of *MARs<sub>Eμ</sub>*. To proceed, we bred our *MARs<sub>Eμ</sub>*-KO mice in a genetic background deficient for both base excision repair (*Ung<sup>Δ/Δ</sup>*) and mismatch repair (*Msh2<sup>Δ/Δ</sup>*) in order to evaluate, on an unbiased manner, the DNA footprint of AID deamination upstream and downstream from *cEμ* (Figure 4A). As expected and according to the literature (45–47), models deficient for both BER and MMR displayed only transitions at C/G pairs reflecting cytidine deamination on

respectively the template and non-template strands. By looking at deamination frequencies between control (*Ung<sup>Δ/Δ</sup> Msh2<sup>Δ/Δ</sup>*) and mutant animals (*Ung<sup>Δ/Δ</sup> Msh2<sup>Δ/Δ</sup> MARs<sub>Eμ</sub><sup>Δ/Δ</sup>*), our data showed that AID activity upstream from *cEμ* was not impeded upon *MARs<sub>Eμ</sub>*-deletion; while differences were not statistically significant (evaluated on n=3 to 4 mice of each genotype), cytidine deamination even tended to be increased in B cells devoid of *MARs<sub>Eμ</sub>*, on both sides of *cEμ* (Figure 4B and Table S4A). When compared to control animals (*Ung<sup>Δ/Δ</sup> Msh2<sup>Δ/Δ</sup>*), nucleotide substitution patterns were unchanged in the absence of *MARs<sub>Eμ</sub>* (Figure 4B and Supplementary Figure S8A), proving identical strand-specific cytidine deamination: roughly 2/3 on the non-template strand (C to T substitutions) and 1/3 on the template strand (G to A substitutions). Besides imbalanced transcription upstream from *Eμ*, this data indicates that *MARs<sub>Eμ</sub>*-deletion does not impact the choice of any DNA strand for AID targeting within intronic regions.

This notable increased AID deamination footprint prompted by *MARs<sub>Eμ</sub>*-deletion was in total agreement with the increased transcription observed in the corresponding regions of activated B cells. The obvious discrepancy between efficient C to U deamination events and the strong SHM targeting defect within the same *J<sub>H</sub>* intron region unravel the origin of the SHM defect in *MARs<sub>Eμ</sub>*-deficient mice as a default of the mutagenic process occurring downstream from the normally-introduced U-G mismatches in DNA.



**FIGURE 4** *MARs<sub>Eμ</sub>* deletion impedes error-prone repair pathways upstream from *Eμ* region. Comparison of *IgH* SHM events occurring on both sides of *cEμ* in Peyer’s patch GC B cells sorted from wt and *MARs<sub>Eμ</sub><sup>Δ/Δ</sup>* mice models, bred in genetic backgrounds deficient for base excision repair (*Ung* KO) and mismatch repair (*Msh2* KO). Data were obtained by NGS (Ion Proton) combined to DeMinEr filtering (28). In each region, analyzed and represented as a panel, bar graphs report overall mutation frequencies (left) and detailed mutation frequencies at all bases (right). (A) Location of *IgH* regions (thick purple lines) tested for SHM, arrows represent primers used for PCR amplification. (B) SHM downstream from *J<sub>H4</sub>* in double-deficient *Ung<sup>Δ/Δ</sup> Msh2<sup>Δ/Δ</sup>* background. (C) SHM downstream from *J<sub>H4</sub>* in DNA repair proficient (*Ung<sup>+/+</sup> Msh2<sup>+/+</sup>*) background. (D) SHM downstream from *J<sub>H4</sub>* in *Ung<sup>Δ/Δ</sup>* background. (E) SHM downstream from *cEμ* in double-deficient *Ung<sup>Δ/Δ</sup> Msh2<sup>Δ/Δ</sup>* background. (F) SHM downstream from *cEμ* in DNA repair proficient (*Ung<sup>+/+</sup> Msh2<sup>+/+</sup>*) background. (G) SHM downstream from *cEμ* in *Ung<sup>Δ/Δ</sup>* background. p-value was determined with two tailed Mann Whitney test; significant differences are indicated by: \*P < 0.05; \*\*P < 0.01 and error bars represent SEM of two to three independent experiments.



This prompted us to investigate whether  $MARs_{E\mu}$ -deletion could provoke skewed mutation patterns within SHM-targeted regions. In the intron region downstream from  $J_H4$ , mutation frequency at each of the four bases in *wt* and  $MARs_{E\mu}$ -deficient backgrounds (Figure 4C and Table S2A) revealed a global significant decrease of mutations at all bases except for substitutions occurring at C in the absence of  $MARs_{E\mu}$ . Similarly, beyond significant differences for C>A, G>C, T>A and T>G events, individual mutation patterns unveiled a global decrease that did not offer any clear hypothesis regarding the mechanism impeding SHM upon  $MARs_{E\mu}$ -deletion. Factually, the unchanged relative proportion of transitions and transversions at C:G base pairs in  $MARs_{E\mu}$  deficient cells did not suggest any defect in UNG activity (45, 48, 49) (Supplementary Figure S8B).

To solve this paradox, we bred our  $MARs_{E\mu}$ -KO mice into base excision repair deficient background ( $Ung^{\Delta/\Delta}$ ) and analysed SHM in the same region. Strikingly, in the absence of UNG, SHM frequency within the  $J_H4$  intron region was identical upon the presence ( $Ung^{\Delta/\Delta}$  control mice) or the absence of  $MARs_{E\mu}$  ( $Ung^{\Delta/\Delta} MARs_{E\mu}^{\Delta/\Delta}$  mice) (Figure 4D and Supplementary Table S4B). Beyond the expected increase of G/C transitions, a typical hallmark of UNG-deficient background, substitution frequencies at all four bases were also identical in  $Ung^{\Delta/\Delta} MARs_{E\mu}^{\Delta/\Delta}$  mice (Figure 4D) the same was true when looking at individual substitution events (Supplementary Figure S8C). The fact that the SHM deficiency induced by  $MARs_{E\mu}$  deletion was no more observed in UNG-deficient background (Figures 4C, D) strongly imply the involvement of BER pathway in the initial mutagenic defect. This same data also proved that SHM events occurring independently of UNG (e.g. altogether obtained by replication across U and/or processed by MMR pathway) took place normally within the  $J_H$  intron in the absence of  $MARs_{E\mu}$ . Given this, a rational hypothesis to explain the origin of the SHM defect in our model was that abasic sites generated by UNG upstream from  $cE\mu$  are processed differently upon the absence of  $MARs_{E\mu}$ . Our data suggests that U:G mismatches processed by UNG are accurately repaired in the absence of  $MARs_{E\mu}$  while these are normally subject to error-prone repair; sustaining for a specific function of  $MARs_{E\mu}$  in recruiting mutagenic factors associated to BER.

In contrast to what observed in the  $J_H$  intron, substitution frequencies and mutation patterns downstream from  $cE\mu$  evidenced a different function for such regulatory regions. In B cells capable of BER and MMR, the absence of  $MARs_{E\mu}$  significantly boosted mutations at all bases by at least two fold (Figure 4F), this was true for any kind of substitution (Supplementary Figure S8E). Substitution patterns collected in mutant animals devoid of BER and MMR highlighted a global “overtargeting” of the  $S\mu$  region induced by  $MARs_{E\mu}$  deletion (Figure 4E, Supplementary Figure S8D and Supplementary Table S4C). This was in line with the general increase in both  $S\mu$  germline transcription observed in this model. In models impaired for BER, our data showed that UNG-deficiency combined to deletion of  $MARs_{E\mu}$  ( $Ung^{\Delta/\Delta} MARs_{E\mu}^{\Delta/\Delta}$  mice) maintained the SHM burden downstream from  $cE\mu$  significantly higher than what observed for UNG alone ( $Ung^{\Delta/\Delta}$  control mice) (Figure 4G; Supplementary Figure S8F and Supplementary Table S4D). Such a comparison suggests that error-prone repair factors could more readily access to abasic sites generated in the S region when  $MARs_{E\mu}$  are missing. In this way, our data support the idea that  $MARs_{E\mu}$  act as physiological barrier that optimize SHM upstream from the  $E\mu$  region and rationalize the fact that  $MARs$  are

evolutionary conserved downstream from Ig gene V regions (24); and moreover conserved structures in mammals (2).

## Concluding remarks

One simplistic model would argue that the most important regulatory regions for  $IgH$  locus expression are conserved upon any reshaping event occurring in developing B lineage cells (VDJ recombination, CSR and SHM). Beyond the major enhancer regions, e.g.  $cE\mu$  and the 3'RR, our current study identifies  $MARs_{E\mu}$ , also conserved upon any rearrangement, as part of these most critical  $IgH$  elements. Since flanking  $cE\mu$ ,  $MARs_{E\mu}$  could potentially interfere with the function of this regulatory element. Our current study showed that  $MARs_{E\mu}$  deletion impact the transcription pattern in this region but also points out, when compared to deletion of  $cE\mu$  alone (7), that the core enhancer and  $MARs_{E\mu}$  act independently on SHM. Our data indicate that  $IgH$   $MARs_{E\mu}$  delimit, upstream from the enhancer region, some error-prone repair processes coupled to BER.

Several studies indeed proposed that  $J$ - $C$  intronic  $MARs$  help generate negative supercoiling and consequently increased ssDNA and potential other secondary structures that could promote accessibility to AID (50–52). The hypothesis that  $MARs_{E\mu}$  add again more DNA strain to the sense-transcribed VDJ transcription unit is relevant to the positive effect of topoisomerase depletion on AID targeting and SHM (51, 53, 54). In line with transcription dynamics, RNA pol II stalling (55, 56) and DIVAC regulatory regions are being proposed to facilitate AID targeting for SHM (57, 58). A model proposed by Alt and colleagues (59) would be that the optimal chromatin environment for AID-induced mutations would be provided by convergent transcription as the result of fine balance between sense and antisense events. As possible byproducts of RNA Pol II collision, antisense or regulatory transcripts in such regions remain transient and difficult to detect in a *wt* context; probably because processed by RNA exosome or other RNase activities (60, 61). In line with our data, a recent study from the Basu group emphasize the critical function of the RNA exosome for the fine tuning of sense transcription of the VDJ exon, leading to modifications on strand-specific AID targeting and SHM (44). While also modifying sense and antisense transcription pattern, our study shows that  $MARs_{E\mu}$ -deletion does not impede AID footprint but rather some of repair mechanisms acting downstream from the U-G mismatch. Since we hypothesize that UNG activity by itself is not necessarily impaired in our model, our data strongly suggest that only BER-dependent error-prone repair is impeded by  $MARs_{E\mu}$ -deletion. This statement supposes that efficient MMR-dependent error-prone repair in this region does not require such a barrier. The question of  $MARs$  binding factors and their respective dynamic association to such regulatory regions needs to be further investigated. The literature already suggest that some of them, like the Special AT-rich binding factor 1 (SATB1), could act as accessory factors in BER (62). Recent findings, showing that UNG2-interacting protein FAM72A promotes error prone processing of U-G mismatch in Ig genes (48, 49), raise the question of its specific recruitment to AID-targeted regions; our current study suggests that  $MARs_{E\mu}$  could potentially interact with error-prone factors acting downstream from UNG2. Another future challenge remains to define whether some components of the nuclear matrix, nuclear filaments or proteins anchored in the envelope, could be involved in the anchorage of SHM targets.

## Material and methods

### Mouse models

To generate  $MARs_{E\mu}$  KO model, gene targeting for matrix attachment regions flanking the *IgH E $\mu$*  enhancer element was performed by homologous recombination, in the murine E14 ES cell line, with a vector kindly provided by Dr. Frederick Alt that permitted replacement of the 995 pb region (including *cE $\mu$*  and its flanking *MARs*) by a 220 pb *Hinfl* genomic fragment that reintroduced the *cE $\mu$*  enhancer, a 256 bp fragment containing plasmid sequence and a “*loxP-pGK-Neo<sup>R</sup>-loxP*” cassette (18). Once introduced in the mouse germline, the selection cassette was deleted *in vivo* by *cre-loxP* recombination as previously described (9) to obtain the  $MARs_{E\mu}^{\Delta\Delta}$  *IgH* allele devoid of both 5' and 3'  $MARs_{E\mu}$  (respectively 344 pb *XbaI-Hinfl* and 426bp *Hinfl-XbaI* genomic fragments) (Figure 1A). Animal procedures were performed on 8 weeks old male and female mice. All models were created in mixed 129Sv;C57BL/6 background. *Wt*,  $MARs_{E\mu}^{\Delta\Delta}$ ,  $E\mu^{\Delta\Delta}$  (9), *Msh2<sup>ΔΔ</sup>*, *Ung<sup>ΔΔ</sup>* (a kind gift of Dr S. Storck) and *Aicda<sup>-/-</sup>* (a kind gift of Pr. T. Honjo) homozygous mice were used for our experiments and maintained at 21–23°C with a 12-h light/dark cycle. All experiments comparing *Wt*,  $MARs_{E\mu}^{\Delta\Delta}$ ,  $E\mu^{\Delta\Delta}$  in DNA repair proficient background used systematically control mice carrying two *IgH* alleles of the same haplotype: *IgHa* (from 129/Ola strain). Procedures were reviewed and approved by the Ministère de l'Éducation Nationale de l'Enseignement Supérieur et Recherche autorisation APAFIS#16639-2018090612249522v2.

### Southern blots and PCR analysis of *cre*-mediated $MARs_{E\mu}$ deletion

Genomic Southern blots were performed as follows: 20  $\mu$ g genomic DNA were digested by *SacI* or *BamHI* and submitted to electrophoresis on a 0.7% agarose gel. DNA was transferred to nylon membranes (MP Biomedicals) by capillarity. Blots were hybridized with [<sup>32</sup>P]-labeled probes generated by random priming. Hybridization with 5' probe (0.803 kpb *SacI-SphI* fragment) and 3' probe (0.8 kpb *XbaI-BamHI* fragment) located outside the homology arms were used to identify ES cell clones in which  $MARs_{E\mu}$  were replaced by the *loxP-pGK-Neo<sup>R</sup>-loxP* cassette (Figure 1A).

### Total serum Ig quantification by ELISA

Sera were collected at 8 weeks of age from non-immunized *wt* and  $MARs_{E\mu}^{\Delta\Delta}$  mice and analyzed for the presence of different Ig classes and subclasses by ELISA as previously described (63).

### SRBC immunisation

Mice were challenged by intraperitoneal injection with 200 $\mu$ L 50% sheep red blood cell suspension and sacrificed 8 days later to collect GC B cells ( $B220^+/GL7^+$ ) in the spleen. Efficient immunization was assessed by counting GC numbers on spleen frozen sections upon immunofluorescence labelling.

### Immunofluorescence labelling of germinal center structures

Spleens cryosections (10 $\mu$ m) were fixed and permeabilized for 20 minutes in cold acetone at -20°C. Sections were stained with the following primary Abs GL7-FITC (1/50),  $\alpha$ B220 APC (1/1000),  $\alpha$ IgD-AF594 (1/1000) diluted in PBS/FCS 2%/EDTA 2mM, overnight at 4°C. After 2 washes with PBS, DAPI was added. Images were acquired using an epifluorescent microscope (NIKON). GC structures were identified according to their content in  $B220+GL7+$  cells.

### Flow cytometry and cell sorting

Flow cytometry analysis was performed on LSR-Fortessa cell analyzer (BD Biosciences) on single-cell suspensions from fresh organs. Once washed with 2% fetal calf serum-PBS, lymphoid cells from bone marrow, spleen, peritoneal cavity and Peyer's patches were labeled with various conjugated Abs:  $\alpha$ B220-V450,  $\alpha$ CD117-PE,  $\alpha$ CD43-PE for bone marrow cells.  $\alpha$ B220-V450,  $\alpha$ CD21-PE,  $\alpha$ CD23-FITC,  $\alpha$ IgM total-PE,  $\alpha$ IgD-FITC and  $\alpha$ CD3e-FITC for splenocytes.  $\alpha$ B220-V450,  $\alpha$ IgM-PE,  $\alpha$ CD5-FITC for peritoneal cavity.  $\alpha$ B220-V450,  $\alpha$ B220-APC,  $\alpha$ IgA-FITC,  $\alpha$ IgM-PE,  $\alpha$ PNA-FITC,  $\alpha$ FAS-PE,  $\alpha$ Ki67-FITC and GL7-FITC for Peyer's Patches. (Southern Biotechnology Associates; eBioscience; Sigma and BD Biosciences). Flow cytometry cell sorting was performed on an ARIA 3 (BD Biosciences) apparatus on single-cell suspensions from spleens or Peyer's patches. Once washed with 2% fetal calf serum-PBS, cells were labeled with PNA, GL7,  $\alpha$ B220, and  $\alpha$ FAS reagents and sorted based on distinct gates defined as germinal center B cells ( $B220^+/GL7^+$  or  $B220^+/PNA^{\text{High}}/Fas^+$ ).

### Cell culture

Splenocytes were collected, after red blood cells lysis,  $CD43^+$  cells were depleted using anti- $CD43$  MicroBeads (Miltenyi Biotec).  $CD43^-$  splenic B cells were cultured for 3 days at a density of  $1 \times 10^6$  cells per mL in RPMI 1640 supplemented in 10% serum calf fetal, sodium pyruvate (Lonza), amino acid (NEAA 100x Lonza) and Penicillin-Streptomycin (Gibco) with 1 $\mu$ g/ml LPS (*In vivo*Gen) alone (for transcription assays) or plus 20ng/ml IL4 (PeproTech) (for CSR experiments).

### SHM assays

SHM analysis within the  $J_H4$  intron was either performed by cloning followed by classical Sanger method as described (64) or performed directly on PCR products by next generation sequencing using GS Junior (Roche) or Ion Proton system (Applied Biosystem). SHM within Smu intron was performed by using the following primer SmuF: 5'-AAGGGCTTCTAAGCCAGTCC-3' and SmuR: 5'-TAGCCTGGGTCCCTCCTTAC-3' and sequenced using GS junior sequencer. For GS Junior, sequencing libraries were prepared

according to the manufacturer's instructions, adaptor sequences were added to the previous amplification primer sequences in order to be compatible with the GS-Junior sequencing technology. Amplifications were performed with Phusion<sup>®</sup> High-Fidelity DNA Polymerase (New England Biolabs) according to the following program: DNA was denatured 40 s at 98°C and then submitted to 38 cycles consisting of 98°C for 10 s, 68°C for 30 s and 72°C for 30 s, and 1 cycle at 72°C for 10 min. PCR products were first purified using NucleoSpin kit (Macherey-Nagel) followed by Ampure bead purification (Beckman Coulter). PCR products were subjected to "PCR emulsion step" (GS Junior+ emPCR Kit (Lib-A), Roche) and sequenced using GS Junior sequencing kit XL+ (Roche) according to the manufacturer's instructions. Raw sequences were aligned against reference sequences of *IgHJ<sub>4</sub>*-downstream intron or *Smu* and only full length sequences were kept for mutation analysis. For *IgHJ<sub>4</sub>*, clonally related sequences were removed based on the sequence of VDJ junction (*i.e.* CDR3) similarity. No further filtering steps were implemented in our analysis workflow. Mutations were called on each sequence using pairwise alignment algorithm (from biopython package) and only base substitutions were reported. Mutation frequencies were computed as the ratio between the sum of mutated bases in all complete sequences over the total number of aligned bases. For Ion Proton, sequencing libraries were prepared according to the user guide Ion Xpress<sup>™</sup> Plus gDNA Fragment Library Preparation (Cat. no. 4471269, Life Technologies). Briefly, PCR products (100ng) were fragmented by enzymatic digestion (Ion Shear<sup>™</sup> Plus Reagents Kit, Cat. no. 4471248) and ligated to Barcodes and Adapters (Ion Plus Fragment Library Kit, Cat. no. 4471252). After 200 bp size selection step on E-Gel precast agarose electrophoresis system, final amplification was performed. Raw data were processed using DeMinEr tool as described (28). In the case of studies performed in BER or MMR-deficient backgrounds of mixed 129Sv;C57BL/6 backgrounds, all SHM analysis were done by excluding nucleotides that differ between strains.

## RT-PCR and q-PCR

Total RNA was prepared by using TRIzol reagent (Ambion) procedures. RNA samples were first treated with DNase I (Invitrogen) for 15 min at 25°C. RT was performed on 200 ng of total RNA with random hexamers or with specific primer (10μM) (sequence available in [Supplementary Figure S2.](#)) using superscript III enzyme (Invitrogen). As control, we performed a reverse transcription without primer to determine the threshold (referred as P<sup>−</sup> in bar graphs). Each real-time qPCR reaction was performed in duplicate on 10 ng of RNA equivalent, using TaqMan Universal (except for Sμ quantification we used SYBR green Mastermix (TAKARA)) on StepOnePlus system (Applied Biosystems). Primary transcription at *IgH* locus was quantified as previously described (37) and completed with both a set of primers and q-PCR probes close to *J<sub>H</sub>* segments (listed in [Supplementary Figure 2.](#)) and a set of primers located 5' to Sμ: Smu-Fw (5'-ACCCAGGCTAAGAAGGCAAT-3'), Smu-Rev (5'-CCTTCCTTCTGCGTATCCAT-3'). Relative mRNA levels were normalized to *Gapdh* transcripts with the appropriate TaqMan probe (Mm99999915\_g1, Applied Biosystem). Data were analyzed by comparing threshold cycle (*CT*) values according to the 2<sup>−(ΔΔ*CT*)</sup>

<sup>(ΔΔ*CT*)</sup> method. The *wt* mice templates used as calibrators were S2 for sense transcripts, AS0 or AS2 for antisense transcripts.

## Statistical analysis

If not specified in the figure legend, Mann Whitney two-tailed tests were used for statistical analysis using GraphPad Prism software (\*p<0.05, \*\*p<0.01, \*\*\*p<0.001, \*\*\*\*p<0.0001).

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ebi.ac.uk/ena>, PRJEB52221.

## Ethics statement

The animal study was reviewed and approved by Procédures were reviewed and approved by the Ministère de l'Éducation Nationale de l'Enseignement Supérieur et Recherche autorisation APAFIS#16639-2018090612249522v2.

## Author contributions

OAM, MT, MM, AG, MB, SB, CC, CB, JC, EP and SN performed experiments. EP and SN conceived and supervised the study. MM developed the experimental model. OAM, BV, PG, RM, EP and SN wrote the manuscript. All authors contributed to the article and approved the submitted version.

## Funding

OM and MT were supported by PhD fellowship of the french Ministère de l'Enseignement Supérieur, de la Recherche et de de l'Innovation. This work was supported by Région Nouvelle Aquitaine, La Ligue Contre le Cancer (comités 87, 23 to EP and SN); the Fondation ARC pour la recherche sur le cancer (PJA 20181207918 to EP and PhD continuation fellowship to MT), Institut CARNOT CALYM, INCa-Cancéropôle GSO Emergence (to EP).

## Acknowledgments

The authors are grateful to BISCEm unit (Univ. Limoges, UAR 2015 CNRS, US 42 Inserm, CHU Limoges) for technical support regarding DNA-RNA sequencing, microscopy, histology, cytometry, cell sorting and animal core facility. We are grateful to Drs. Fred Alt for providing *MARs<sub>Eμ</sub>* targeting construct, Zéliha Oruc-Ratinaud and Tiffany Marchiol regarding animal health status. We thank Dominik Schenten, Sébastien Storck, Christophe Sirac, Laurent Delpy, Brice Laffleur, Alexis Saintamand and Jeanne Moreau for discussions and helpful comments. We are grateful to Michel Cogné for critical review of the manuscript.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1030813/full#supplementary-material>

### SUPPLEMENTARY FIGURE 1

Gating strategy to test developing B cells in mice. **(A)** Bone marrow B cell populations in *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. Top row: B220+ CD117+ pre-pro B cells were stained with V450-anti-B220, PE-anti-CD117 Abs. Bottom row: B220+ CD43<sup>High</sup> pro-B and B220+CD43<sup>Low</sup> pre-B cell populations were stained with V450-anti-B220, FITC-anti-IgM, and PE-anti-CD43 Abs, gated on the IgM-negative population. **(B)** Splenic B cell subsets in *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. Top row: CD21<sup>High</sup> CD23<sup>Low</sup> marginal zone and CD21<sup>Low</sup> CD23<sup>High</sup> follicular B cell populations were stained with V450-anti-B220, PE-anti-CD21 and FITC-anti-CD23 Abs, gated on B220+ population. Bottom row: IgM<sup>+</sup>IgD<sup>+</sup> mature B cells were stained with V450-anti-B220, FITC-anti-IgD and PE-anti-IgM Abs. **(C)** Peritoneal cavity B cells in *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. IgM+ CD5+ -B1a and IgM+ CD5- -B1b subsets were stained with V450-anti-B220, PE-anti-IgM, FITC-anti-CD5 Abs, gated on the B220+ population.

### SUPPLEMENTARY FIGURE 2

SHM frequency within *J<sub>H</sub>4* intron. **(A)** SHM downstream from *J<sub>H</sub>3* and *J<sub>H</sub>4* segments in Peyer's patch GC B cells sorted from *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. For each genotype, pie charts represent distribution of mutated sequences (proportional to the area in each slice, data obtained by Sanger sequencing method) in individually recombined *IgH* alleles. Number of individual clones is reported in the center (after removal of clonally related sequences based on VDJ junction). Each pie chart represent SHM obtained from an individual experiment. Under each pie chart, SHM frequency, sequencing strategy and sample type (individual mice or pool) is indicated. Mean SHM frequency and p values are reported. **(B)** Bar graph displayed numbers of germinal center structures obtained in the spleen of *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice upon SRBC-immunization (left). Representative images of spleen section from immunized *wt* and *MARSEμ<sup>Δ/Δ</sup>* animals (right) **(C)** Equivalent data representation than reported in **A** for SHM downstream from *J<sub>H</sub>4* segments in splenic GC B cells sorted from SRBC-immunized *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. Mean SHM frequency and p values are reported.

### SUPPLEMENTARY FIGURE 3

SHM frequency downstream from *cEμ* region. **(A)** SHM downstream from *cEμ* region in Peyer's patch GC B cells sorted from *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice obtained by GS junior sequencing. For each genotype, pie charts represent distribution of mutated sequences (proportional to the area in each slice, data obtained by Sanger and GS Junior sequencing method) in individually recombined *IgH* alleles. Number of individual clones is reported in the center (after removal of clonally related sequences based on VDJ junction). Each pie chart represent SHM obtained from an individual experiment. Under each pie chart, SHM frequency, sequencing strategy and sample type (individual mice or pool) is

indicated. Mean SHM frequency and p values are reported. **(B)** Left: schematic representation of both *wt* and *MARSEμ<sup>Δ/Δ</sup>* alleles to illustrate the distance between *cEμ* and *Sμ* region. Blue and red bars indicate the 473bp-designed region to quantify SHM at the same distance from *cEμ* in the two models. Right: Bar graph showing SHM frequency in the corresponding regions of *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice, the analysis was performed on samples described in **A**.

### SUPPLEMENTARY FIGURE 4

Annotated nucleotide map of the *IgH-J<sub>H</sub>1* to *Eμ* germline region from of 129 *wt* mice. All *J<sub>H</sub>* exons as well as *coreEμ* element are indicated by bold characters. Start sites for antisense transcripts are reported as (\*) according to the Alt study (38). Location of primers used for strand-specific reverse transcription (S1, S2, S3, AS0, AS1, AS2, AS3) are indicated by underlines. TaqMan qPCR amplicons (C, A, A') are highlighted in grey.

### SUPPLEMENTARY FIGURE 5

Sense and antisense transcripts quantified with *IgH J<sub>H</sub>3* and *J<sub>H</sub>4* exons with additional TaqMan probes. **(A)** Murine *IgH* locus (not to scale) indicating location of primers (S1, S2, S3; black arrows) within introns downstream from *J<sub>H</sub>3* and *J<sub>H</sub>4* used for strand-specific reverse transcription to detect sense transcripts (dotted arrows). Black bars (A' and C) indicate location of q-PCR probes. Total primary transcripts and primary sense transcripts were quantified with A' and C probes in Peyer's patch GC B cells (dark colors) and *in vitro*-activated B cells (light colors) from *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. **(B)** Murine *IgH* locus (not to scale) indicating location of primers (AS0, AS1, AS2, AS3; black arrows) within introns downstream from *J<sub>H</sub>2*, *J<sub>H</sub>3* and *J<sub>H</sub>4* used for strand-specific reverse transcription to detect antisense transcripts (dotted arrows). Black bars (A' and C) indicate location of q-PCR probes. Primary antisense transcripts quantified with A' and C probes in Peyer's patch GC B cells (dark colors) and *in vitro*-activated B cells (light colors) from *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. Dots indicated antisense transcripts start sites according to the Alt study (38). Baseline level was either provided by using a RT template performed without primers (P-) or using one strand-specific template that cannot be detected with the current probe (T-). p-value was determined with two tailed Mann Whitney test; significant differences are indicated by: \*P<0.05; \*\*P<0.01 and error bars represent SEM of two to three independent experiments.

### SUPPLEMENTARY FIGURE 6

Sense and antisense transcription and expression in naive B cells. **(A)** IgH Sense (left bar graph) and antisense (right bar graph) transcripts quantified with A probe in naive B cells sorted from Peyer's patch in *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. **(A)** Intracellular IgM mean fluorescence intensities measured by flow cytometry in naive B cells from Peyer's patches of *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. **(C)** IgH sense (left bar graph) and antisense (right bar graph) transcripts quantified with A probe in resting splenic B cells sorted from *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. Bar graphs show mean ± SEM of two independent experiments.

### SUPPLEMENTARY FIGURE 7

Comparison of *in-vitro* Ig class switching in *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. Percentage of IgG3 and IgG1 positive cells measured by flow cytometry after respectively LPS or LPS + IL4 stimulation for 3 days of splenic B cells sorted from *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice. Bar graphs show mean ± SEM of one representative experiment for each condition, gating strategy is indicated below.

### SUPPLEMENTARY FIGURE 8

Base substitution patterns in BER- and MMR-deficient backgrounds. Comparison of SHM-related base substitution patterns, reported as frequencies, at *IgH* in Peyer's patch GC B cells sorted from *wt* and *MARSEμ<sup>Δ/Δ</sup>* mice models, bred in genetic backgrounds deficient for base excision repair (*Ung<sup>Δ/Δ</sup>*) and mismatch repair (*Msh2<sup>Δ/Δ</sup>*). Data were obtained by NGS (Ion Proton) combined to DeMinEr filtering (28). **(A)** Substitution pattern downstream from *J<sub>H</sub>4* in double-deficient *Ung<sup>Δ/Δ</sup> Msh2<sup>Δ/Δ</sup>* background. **(B)** Substitution pattern downstream from *J<sub>H</sub>4* in DNA repair proficient (*Ung<sup>+/+</sup> Msh2<sup>+/+</sup>*) background. Relative proportions of transitions and transversions at dC/dG pairs are reported in an additional bar graph. **(C)** Substitution pattern downstream from *J<sub>H</sub>4* in *Ung<sup>Δ/Δ</sup>* background. **(D)** Substitution pattern downstream from *Iμ* in double-deficient *Ung<sup>Δ/Δ</sup> Msh2<sup>Δ/Δ</sup>* background. **(E)** Substitution pattern downstream from *Iμ* in DNA repair proficient (*Ung<sup>+/+</sup> Msh2<sup>+/+</sup>*) background. **(F)** Substitution pattern downstream from *Iμ* in *Ung<sup>Δ/Δ</sup>* background. p-value was determined with two tailed Mann Whitney test; significant differences are indicated by: \*P<0.05; \*\*P<0.01 and error bars represent SEM of two to three independent experiments.



## SUPPLEMENTARY TABLE 1

MARs<sub>EU</sub> deletion led to normal B-lineage cell development. Bone Marrow and peripheral B cell subsets counts in *wt* and MARs<sub>EU</sub><sup>Δ/Δ</sup> mice. Absolute numbers are reported as mean ± SEM. Significance was assessed with Student T test. P value is indicated when difference is significant.

## SUPPLEMENTARY TABLE 2

SHM data (NGS) from individual mice in DNA repair proficient background. Total number of mutations, total number of bp analyzed and mutation frequencies for *wt* and MARs<sub>EU</sub><sup>Δ/Δ</sup> mice. (A) Data from intron 3' to J<sub>H</sub>4 in Peyer's patches GC B cells, (B) Data from spleen GC B cells from SRBC-immunized mice, (C) Data from intron 3' to I<sub>μ</sub> in Peyer's patches GC B cells.

## SUPPLEMENTARY TABLE 3

Unselected SHM within out of frame VDJ exons in GC B cells sorted from Peyer's patch. Unique VDJ junctions, rearranged to J<sub>H</sub>3 and J<sub>H</sub>4 segments, were obtained

after amplification, cloning and sequencing by classical Sanger method of DNA extracted from Peyer's patch sorted GC B cells (n=2 to 3 independent samples corresponding to individual mice or pools of 7 to 12 mice of each genotype). SHM data were calculated after processing individually recombined IgH alleles by IMGT V-Quest software ([https://www.imgt.org/IMGT\\_vquest](https://www.imgt.org/IMGT_vquest)); numbers of base substitutions identified as SHM by the software in the FR3 and CDR3 regions are reported in the table and were used to calculate SHM frequencies.

## SUPPLEMENTARY TABLE 4

SHM data (NGS) from individual mice in genetic backgrounds deficient for base excision repair (*Ung*-deficient) and mismatch repair (*Msh2*-deficient). Total number of mutations, total number of bp analyzed and mutation frequencies. (A) Data from intron 3' to J<sub>H</sub>4 in Peyer's patches GC B cells of *Ung*<sup>Δ/Δ</sup> *Msh2*<sup>Δ/Δ</sup> and *Ung*<sup>Δ/Δ</sup> *Msh2*<sup>Δ/Δ</sup> MARs<sub>EU</sub><sup>Δ/Δ</sup> mice, (B) Data from *Ung*<sup>Δ/Δ</sup> and *Ung*<sup>Δ/Δ</sup> MARs<sub>EU</sub><sup>Δ/Δ</sup> mice, (C) Data from intron 3' to I<sub>μ</sub> in Peyer's patches GC B cells of *Ung*<sup>Δ/Δ</sup> *Msh2*<sup>Δ/Δ</sup> and *Ung*<sup>Δ/Δ</sup> *Msh2*<sup>Δ/Δ</sup> MARs<sub>EU</sub><sup>Δ/Δ</sup> mice, (D) Data from *Ung*<sup>Δ/Δ</sup> and *Ung*<sup>Δ/Δ</sup> MARs<sub>EU</sub><sup>Δ/Δ</sup> mice.

## References

- Perlot T, Alt FW. Cis-regulatory elements and epigenetic changes control genomic rearrangements of the IgH locus. *Adv Immunol* (2008) 99:1–32. doi: 10.1016/S0065-2776(08)00601-9
- Scheuermann RH, Garrard WT. MARs of antigen receptor and co-receptor genes. *Crit Rev Eukaryot Gene Expr* (1999) 9:295–310. doi: 10.1615/CritRevEukaryotGeneExpr.v9i3-4.140
- Gluch A, Vidakovic M, Bode J. Scaffold/matrix attachment regions (S/MARs): relevance for disease and therapy. *Handb Exp Pharmacol* (2008), 186:67–103. doi: 10.1007/978-3-540-72843-6\_4
- Teves SS, Henikoff S. DNA Torsion as a feedback mediator of transcription and chromatin dynamics. *Nucleus* (2014) 5:211–8. doi: 10.4161/nucl.29086
- Pommier Y, Sun Y, Huang S-YN, Nitiss JL. Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nat Rev Mol Cell Biol* (2016) 17:703–21. doi: 10.1038/nrm.2016.111
- Cockerill PN, Yuen MH, Garrard WT. The enhancer of the immunoglobulin heavy chain locus is flanked by presumptive chromosomal loop anchorage elements. *J Biol Chem* (1987) 262:5394–7. doi: 10.1016/S0021-9258(18)61200-1
- Perlot T, Alt FW, Bassing CH, Suh H, Pinaud E. Elucidation of IgH intronic enhancer functions via germ-line deletion. *Proc Natl Acad Sci USA* (2005) 102:14362–7. doi: 10.1073/pnas.0507090102
- Afshar R, Pierce S, Bolland DJ, Corcoran A, Oltz EM. Regulation of IgH gene assembly: role of the intronic enhancer and 5'DQ52 region in targeting DHJH recombination. *J Immunol* (2006) 176:2439–47. doi: 10.4049/jimmunol.176.4.2439
- Marquet M, Garot A, Bender S, Carrion C, Rouaud P, Lecardeur S, et al. The  $\epsilon\mu$  enhancer region influences h chain expression and b cell fate without impacting IgVH repertoire and immune response *in vivo*. *J Immunol* (2014) 193:1171–83. doi: 10.4049/jimmunol.1302868
- Ronai D, Iglesias-Ussel MD, Fan M, Shulman MJ, Scharff MD. Complex regulation of somatic hypermutation by cis-acting sequences in the endogenous IgH gene in hybridoma cells. *Proc Natl Acad Sci USA* (2005) 102:11829–34. doi: 10.1073/pnas.0505449102
- Azuma T, Motoyama N, Fields LE, Loh DY. Mutations of the chloramphenicol acetyl transferase transgene driven by the immunoglobulin promoter and intron enhancer. *Int Immunol* (1993) 5:121–30. doi: 10.1093/intimm/5.2.121
- Giusti AM, Manser T. Hypermutation is observed only in antibody h chain V region transgenes that have recombined with endogenous immunoglobulin h DNA: implications for the location of cis-acting elements required for somatic mutation. *J Exp Med* (1993) 177:797–809. doi: 10.1084/jem.177.3.797
- Motoyama N, Miwa T, Suzuki Y, Okada H, Azuma T. Comparison of somatic mutation frequency among immunoglobulin genes. *J Exp Med* (1994) 179:395–403. doi: 10.1084/jem.179.2.395
- Bachl J, Wabl M. Enhancers of hypermutation. *Immunogenetics* (1996) 45:59–64. doi: 10.1007/s002510050167
- Lin MM, Green NS, Zhang W, Scharff MD. The effects of  $\epsilon\mu$ , 3'alpha (hs 1,2) and 3'kappa enhancers on mutation of an ig-VDJ-Cgamma2a ig heavy gene in cultured b cells. *Int Immunol* (1998) 10:1121–9. doi: 10.1093/intimm/10.8.1121
- Ronai D, Berru M, Shulman MJ. Variegated expression of the endogenous immunoglobulin heavy-chain gene in the absence of the intronic locus control region. *Mol Cell Biol* (1999) 19:7031–40. doi: 10.1128/MCB.19.10.7031
- Li F, Yan Y, Pieretti J, Feldman DA, Eckhardt LA. Comparison of identical and functional igh alleles reveals a nonessential role for  $\epsilon\mu$  in somatic hypermutation and class-switch recombination. *J Immunol* (2010) 185:6049–57. doi: 10.4049/jimmunol.0902992
- Sakai E, Bottaro A, Davidson L, Sleckman BP, Alt FW. Recombination and transcription of the endogenous ig heavy chain locus is effected by the ig heavy chain intronic enhancer core region in the absence of the matrix attachment regions. *Proc Natl Acad Sci USA* (1999) 96:1526–31. doi: 10.1073/pnas.96.4.1526
- Kohwi-Shigematsu T, Maass K, Bode J. A thymocyte factor SATB1 suppresses transcription of stably integrated matrix-attachment region-linked reporter genes. *Biochemistry* (1997) 36:12005–10. doi: 10.1021/bi971444j
- Wang Z, Goldstein A, Zong RT, Lin D, Neufeld EJ, Scheuermann RH, et al. Cux/CDP homeoprotein is a component of NF- $\mu$ NR and represses the immunoglobulin heavy chain intronic enhancer by antagonizing the bright transcription activator. *Mol Cell Biol* (1999) 19:284–95. doi: 10.1128/MCB.19.1.284
- Kaplan MH, Zong RT, Herrscher RF, Scheuermann RH, Tucker PW. Transcriptional activation by a matrix associating region-binding protein. Contextual requirements for the function of bright. *J Biol Chem* (2001) 276:21325–30. doi: 10.1074/jbc.M100836200
- Wiersma EJ, Ronai D, Berru M, Tsui FW, Shulman MJ. Role of the intronic elements in the endogenous immunoglobulin heavy chain locus. Either the matrix attachment regions or the core enhancer is sufficient to maintain expression. *J Biol Chem* (1999) 274:4858–62. doi: 10.1074/jbc.274.8.4858
- Goyenechea B, Kliks N, Yelamos J, Williams GT, Riddell A, Neuberger MS, et al. Cells strongly expressing ig(kappa) transgenes show clonal recruitment of hypermutation: a role for both MAR and the enhancers. *EMBO J* (1997) 16:3987–94. doi: 10.1093/emboj/16.13.3987
- Yi M, Wu P, Trevorrow KW, Clafin L, Garrard WT. Evidence that the igkappa gene MAR regulates the probability of premature V-J joining and somatic hypermutation. *J Immunol* (1999) 162:6029–39. doi: 10.4049/jimmunol.162.10.6029
- Franklin A, Blanden RV. Hypothesis: biological role for J-c intronic matrix attachment regions in the molecular mechanism of antigen-driven somatic hypermutation. *Immunol Cell Biol* (2005) 83:383–91. doi: 10.1111/j.1440-1711.2005.01327.x
- Inlay MA, Gao HH, Odegard VH, Lin T, Schatz DG, Xu Y. Roles of the ig kappa light chain intronic and 3' enhancers in igk somatic hypermutation. *J Immunol* (2006) 177:1146–51. doi: 10.4049/jimmunol.177.2.1146
- Rada C, Gupta SK, Gherardi E, Milstein C. Mutation and selection during the secondary response to 2-phenylloxazalone. *Proc Natl Acad Sci U.S.A.* (1991) 88:5508–12. doi: 10.1073/pnas.88.13.5508
- Martin OA, Garot A, Le Noir S, Aldigier J-C, Cogné M, Pinaud E, et al. Detecting rare AID-induced mutations in b-lineage oncogenes from high-throughput sequencing data using the detection of minor variants by error correction method. *J Immunol* (2018) 201:950–6. doi: 10.4049/jimmunol.1800203
- Alamyar E, Duroux P, Lefranc M-P, Giudicelli V. IMGT<sup>®</sup> tools for the nucleotide analysis of immunoglobulin (IG) and T cell receptor (TR) V-(D)-J repertoires, polymorphisms, and IG mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. *Methods Mol Biol* (2012) 882:569–604. doi: 10.1007/978-1-61779-842-9\_32
- Lennon GG, Perry RP. C mu-containing transcripts initiate heterogeneously within the IgH enhancer region and contain a novel 5'-nontranslatable exon. *Nature* (1985) 318:475–8. doi: 10.1038/318475a0
- Petersen S, Casellas R, Reina-San-Martin B, Chen HT, Difilippantonio MJ, Wilson PC, et al. AID is required to initiate Nbs1/gamma-H2AX focus formation and mutations at sites of class switching. *Nature* (2001) 414:660–5. doi: 10.1038/414660a
- Nagaoka H, Muramatsu M, Yamamura N, Kinoshita K, Honjo T. Activation-induced deaminase (AID)-directed hypermutation in the immunoglobulin smu region: Implication of AID involvement in a common step of class switch recombination and somatic hypermutation. *J Exp Med* (2002) 195:529–34. doi: 10.1084/jem.20012144
- Heltzel JHM, Maul RW, Yang W, Gearhart PJ. Promoter proximity defines mutation window for VH and VK genes rearranged to different J genes. *J Immunol* (2022), 208:2220–2226. doi: 10.4049/jimmunol.2101002
- Sakai E, Bottaro A, Alt FW. The ig heavy chain intronic enhancer core region is necessary and sufficient to promote efficient class switch recombination. *Int Immunol* (1999) 11:1709–13. doi: 10.1093/intimm/11.10.1709

35. Xue K, Rada C, Neuberger MS. The *in vivo* pattern of AID targeting to immunoglobulin switch regions deduced from mutation spectra in *msh2<sup>-/-</sup> ung<sup>-/-</sup>* mice. *J Exp Med* (2006) 203:2085–94. doi: 10.1084/jem.20061067
36. Fukita Y, Jacobs H, Rajewsky K. Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity* (1998) 9:105–14. doi: 10.1016/S1074-7613(00)80592-0
37. Tinguely A, Chemin G, Péron S, Sirac C, Reynaud S, Cogné M, et al. Cross talk between immunoglobulin heavy-chain transcription and RNA surveillance during b cell development. *Mol Cell Biol* (2012) 32:107–17. doi: 10.1128/MCB.06138-11
38. Perlot T, Li G, Alt FW. Antisense transcripts from immunoglobulin heavy-chain locus V(D)J and switch regions. *Proc Natl Acad Sci USA* (2008) 105:3843–8. doi: 10.1073/pnas.0712291105
39. Bolland DJ, Wood AL, Johnston CM, Bunting SF, Morgan G, Chakalova L, et al. Antisense intergenic transcription in V(D)J recombination. *Nat Immunol* (2004) 5:630–7. doi: 10.1038/ni1068
40. Zhao Y, Dunn-Walters DK, Barone F, Spencer J. Antisense transcripts of V(D)J rearrangements; artifacts caused by false priming? *Mol Immunol* (2009) 46:2357–62. doi: 10.1016/j.molimm.2009.03.020
41. Li W, Notani D, Rosenfeld MG. Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat Rev Genet* (2016) 17:207–23. doi: 10.1038/nrg.2016.4
42. Laffleur B, Basu U, Lim J. RNA Exosome and non-coding RNA-coupled mechanisms in AID-mediated genomic alterations. *J Mol Biol* (2017) 429:3230–3241. doi: 10.1016/j.jmb.2016.12.021
43. Lim J, Giri PK, Kazadi D, Laffleur B, Zhang W, Grinstein V, et al. Nuclear proximity of Mtr4 to RNA exosome restricts DNA mutational asymmetry. *Cell* (2017) 169:523–537.e15. doi: 10.1016/j.cell.2017.03.043
44. Laffleur B, Lim J, Zhang W, Chen Y, Pefanis E, Bizarro J, et al. Noncoding RNA processing by DIS3 regulates chromosomal architecture and somatic hypermutation in b cells. *Nat Genet* (2021) 53:230–42. doi: 10.1038/s41588-020-00772-0
45. Rada C, Di Noia JM, Neuberger MS. Mismatch recognition and uracil excision provide complementary paths to both ig switching and the A/T-focused phase of somatic mutation. *Mol Cell* (2004) 16:163–71. doi: 10.1016/j.molcel.2004.10.011
46. Shen HM, Tanaka A, Bozek G, Nicolae D, Storb U. Somatic hypermutation and class switch recombination in *Msh6<sup>-/-</sup>Ung<sup>-/-</sup>* double-knockout mice. *J Immunol* (2006) 177:5386–92. doi: 10.4049/jimmunol.177.8.5386
47. Liu M, Duke JL, Richter DJ, Vinuesa CG, Goodnow CC, Kleinstein SH, et al. Two levels of protection for the b cell genome during somatic hypermutation. *Nature* (2008) 451:841–5. doi: 10.1038/nature06547
48. Feng Y, Li C, Stewart JA, Barbulescu P, Seija Desivo N, Álvarez-Quilón A, et al. FAM72A antagonizes UNG2 to promote mutagenic repair during antibody maturation. *Nature* (2021) 600:324–8. doi: 10.1038/s41586-021-04144-4
49. Rogier M, Moritz J, Robert I, Lescale C, Heyer V, Abello A, et al. Fam72a enforces error-prone DNA repair during antibody diversification. *Nature* (2021) 600:329–33. doi: 10.1038/s41586-021-04093-y
50. Lebecque SG, Gearhart PJ. Boundaries of somatic mutation in rearranged immunoglobulin genes: 5' boundary is near the promoter, and 3' boundary is approximately 1 kb from V(D)J gene. *J Exp Med* (1990) 172:1717–27. doi: 10.1084/jem.172.6.1717
51. Shen HM, Storb U. Activation-induced cytidine deaminase (AID) can target both DNA strands when the DNA is supercoiled. *Proc Natl Acad Sci USA* (2004) 101:12997–3002. doi: 10.1073/pnas.0404974101
52. Wright BE, Schmidt KH, Davis N, Hunt AT, Minnick MF. II. correlations between secondary structure stability and mutation frequency during somatic hypermutation. *Mol Immunol* (2008) 45:3600–8. doi: 10.1016/j.molimm.2008.05.012
53. Kobayashi M, Sabouri Z, Sabouri S, Kitawaki Y, Pommier Y, Abe T, et al. Decrease in topoisomerase I is responsible for activation-induced cytidine deaminase (AID)-dependent somatic hypermutation. *Proc Natl Acad Sci USA* (2011) 108:19305–10. doi: 10.1073/pnas.1114522108
54. Maul RW, Saribasak H, Cao Z, Gearhart PJ. Topoisomerase I deficiency causes RNA polymerase II accumulation and increases AID abundance in immunoglobulin variable genes. *DNA Repair (Amst)* (2015) 30:46–52. doi: 10.1016/j.dnarep.2015.03.004
55. Kodgire P, Mukkavar P, Ratnam S, Martin TE, Storb U. Changes in RNA polymerase II progression influence somatic hypermutation of ig-related genes by AID. *J Exp Med* (2013) 210:1481–92. doi: 10.1084/jem.20121523
56. Maul RW, Cao Z, Venkataraman L, Giorgetti CA, Press JL, Denizot Y, et al. Spt5 accumulation at variable genes distinguishes somatic hypermutation in germinal center b cells from ex vivo-activated cells. *J Exp Med* (2014) 211:2297–306. doi: 10.1084/jem.20131512
57. Kohler KM, McDonald JJ, Duke JL, Arakawa H, Tan S, Kleinstein SH, et al. Identification of core DNA elements that target somatic hypermutation. *J Immunol* (2012) 189:5314–26. doi: 10.4049/jimmunol.1202082
58. Tarsalainen A, Maman Y, Meng F-L, Kyläniemi MK, Soikkeli A, Budzyńska P, et al. Ig enhancers increase RNA polymerase II stalling at somatic hypermutation target sequences. *J Immunol* (2022) 208:143–54. doi: 10.4049/jimmunol.2100923
59. Meng F-L, Du Z, Federation A, Hu J, Wang Q, Kieffer-Kwon K-R, et al. Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell* (2014) 159:1538–48. doi: 10.1016/j.cell.2014.11.014
60. Basu U, Meng F-L, Keim C, Grinstein V, Pefanis E, Eccleston J, et al. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. *Cell* (2011) 144:353–63. doi: 10.1016/j.cell.2011.01.001
61. Pefanis E, Wang J, Rothschild G, Lim J, Chao J, Rabadan R, et al. Noncoding RNA transcription targets AID to divergently transcribed loci in b cells. *Nature* (2014) 514:389–93. doi: 10.1038/nature13580
62. Kaur S, Coulombe Y, Ramdzan ZM, Leduy L, Masson J-Y, Nepveu A. Special AT-rich sequence-binding protein 1 (SATB1) functions as an accessory factor in base excision repair. *J Biol Chem* (2016) 291:22769–80. doi: 10.1074/jbc.M116.735696
63. Pinaud E, Khamlichi AA, Le Morvan C, Drouet M, Nalesso V, Le Bert M, et al. Localization of the 3' IgH locus elements that effect long-distance regulation of class switch recombination. *Immunity* (2001) 15:187–99. doi: 10.1016/S1074-7613(01)00181-9
64. Rouaud P, Vincent-Fabert C, Saintamand A, Fiancette R, Marquet M, Robert I, et al. The IgH 3' regulatory region controls somatic hypermutation in germinal center b cells. *J Exp Med* (2013) 210:1501–7. doi: 10.1084/jem.20130072

**Annexe 2. A dual function for the chromatin organizer Special A-T rich Binding Protein 1 in B-lineage cells**

Morgane Thomas, Charlotte Bruzeau, Ophélie Alyssa Martin, Justine Pollet, Sébastien Bender, Claire Carrion, Sandrine Le Noir and Eric Pinaud.

Cell. Mol. Immunol. <https://doi.org/10.1038/s41423-023-01069-y>

## ARTICLE



# A dual function for the chromatin organizer Special A-T rich Binding Protein 1 in B-lineage cells

Morgane Thomas<sup>1,4</sup>, Charlotte Bruzeau<sup>1</sup>, Ophélie Alyssa Martin<sup>1</sup>, Justine Pollet<sup>1</sup>, Sébastien Bender<sup>1,2,3</sup>, Claire Carrion<sup>1</sup>, Sandrine Le Noir<sup>1</sup> and Eric Pinaud<sup>1</sup>

© The Author(s), under exclusive licence to CSI and USTC 2023

SATB1 (Special A-T rich Binding protein 1) is a cell type-specific factor that regulates the genetic network in developing T cells and neurons. In T cells, SATB1 is required for lineage commitment, VDJ recombination, development and maturation. Considering that its expression varies during B-cell differentiation, the involvement of SATB1 needs to be clarified in this lineage. Using a KO mouse model in which SATB1 was deleted from the pro-B-cell stage, we examined the consequences of SATB1 deletion in naive and activated B-cell subsets. Our model indicates first, unlike its essential function in T cells, that SATB1 is dispensable for B-cell development and the establishment of a broad IgH repertoire. Second, we show that SATB1 exhibits an ambivalent function in mature B cells, acting sequentially as a positive and negative regulator of Ig gene transcription in naive and activated cells, respectively. Third, our study indicates that the negative regulatory function of SATB1 in B cells extends to the germinal center response, in which this factor limits somatic hypermutation of Ig genes.

**Keywords:** B cells; Nuclear factor; Ambivalent; Somatic hypermutation

*Cellular & Molecular Immunology*; <https://doi.org/10.1038/s41423-023-01069-y>

## INTRODUCTION

The transcription factor Special A-T rich binding protein 1 (SATB1) is a factor able to bind to Matrix Attachment Regions (MAR) in the nucleus [1]. This specific MAR-binding protein (MAR-BP) is indispensable for T lymphocyte development [2–4] through its involvement in properly organizing nuclear architecture, especially chromatin folding [5–7]. Murine SATB1, sharing more than 98% identity with its human homolog, can multimerize through a ubiquitin-like domain [8] and interacts with chromatin through its CUT-like and homeodomains, which are involved in DNA binding affinity and specificity, respectively [9, 10]. One historic feature of SATB1 is its ability to bind nuclear matrix proteins [11, 12]. Another essential feature, as a transcription factor, is its capacity to bind DNA for gene regulation. Strikingly, such interactions were found to occur in nucleosome-dense regions, preferentially at AT-rich sequences in the nucleosomal core [13], a feature that strongly supports its candidacy as a pioneer factor. Given its multivalent potency as a nuclear matrix scaffolding factor, a genome organizer involved in chromatin looping and a transcription factor, the exact mechanism linking this MAR-BP to gene regulation remains puzzling. The large body of literature related to SATB1 in T cells suggests that this factor is capable of flexible functions, either binding to euchromatin [5, 6] or nucleosome-dense regions [13], depending on the cell type and development stage assessed. Long-distance interactions between promoters and enhancers can

be promoted by SATB1 homotetramerization [8]. A negative regulatory function for SATB1 was attributed to its capacity to recruit chromatin modifiers such as histone deacetylase 1 (HDAC1) [14]. Indeed, SATB1 is subject to posttranslational modification: phosphorylation of this MAR-BP modifies interactions with chromatin corepressor and coactivator complexes, leading to a switch in its transcriptional activity [15, 16]; SATB1 acetylation disrupts its interaction with C-terminal binding protein 1 (CtBP1) [17]. Moreover, regulation of the *Satb1* gene is tightly regulated in T cells by interchanging promoter usage, resulting in fine tuning of protein expression [18, 19]. These multiple levels of regulation suggest the need for meticulous control of SATB1 expression.

SATB1 deletion in mice led to major alterations in neuronal, hematopoietic and immune systems [2, 20] and pointed out broad and critical functions for this protein in mouse development. SATB1 is a major player in early hematopoiesis since it promotes hematopoietic stem cell (HSC) self-renewal [21]. Based on its relative expression level in differentiating HSCs [22], SATB1 favors the lymphoid lineage [23]. In erythrocytes and myeloid cells, this protein also modulates epigenetic marks by association with CREB binding protein (CBP) to control  $\beta$ -globin and NADPH oxidase gene expression [24–26]. There is also growing evidence showing that SATB1 is involved in cancer and autoimmune diseases [6, 7, 27, 28], highlighting the necessity to decipher its regulation in immune cells.

<sup>1</sup>Laboratoire Contrôle de la Réponse Immune B et des Lymphoproliférations (CRIBL), Université de Limoges, CNRS Unité Mixte de Recherche 7276, Inserm Unité 1262, Limoges, France. <sup>2</sup>Centre Hospitalier Universitaire Dupuytren, Service d'Immunopathologie, Limoges, France. <sup>3</sup>Centre Hospitalier Universitaire de Limoges, Centre National de l'Amylose AL et Autres Maladies par Dépôt d'Immunoglobulines Monoclonales, Limoges, France. <sup>4</sup>Present address: Laboratoire Suivi des Thérapies Innovantes, Institut de Génétique Humaine, UMR 9002 CNRS-UM, Montpellier, France. email: sandrine.le-noir@unilim.fr; eric.pinaud@unilim.fr

Received: 10 November 2022 Accepted: 17 July 2023

Published online: 07 August 2023

Due to thymus alteration in *Satb1* KO mice [2], SATB1's functions were largely studied in T cells. Initially described as a repressor factor [12, 29], SATB1 proved to also be a transcriptional activator [6, 30]. As an example of ambivalence, SATB1 could repress *c-myc* in resting T cells while stimulating its expression in activated T cells [3]. It is now established that SATB1 is a major regulator of thymocyte development throughout all differentiation stages from progenitors to regulatory subsets [28, 31, 32]. One notable role of SATB1 in T cells is its activator function for *Rag* gene expression to promote VDJ recombination and shape the T-cell receptor repertoire [33, 34]. In this context, SATB1 has been shown to bridge *Rag1* and *Rag2* gene promoters and the distant antisilencer element (ASE) of this locus [33].

In contrast to the vast knowledge on SATB1 fine tuning and regulatory functions in T lymphocytes, little is known about the potency of SATB1 in regulating B-cell development beyond HSC fate decisions [22, 23]. Initial studies reported a discrete defect in B-cell numbers in *Satb1* KO mice [2]. Taking advantage of their elegant fluorescent reporter model, the Yokota group described fluctuations in *Satb1* expression throughout B-cell development with higher expression in naive B-cell subsets. In this study, the authors proposed that SATB1 is involved in BCR-mediated B-cell survival [35]. Many aspects of the literature point to a potential function for SATB1 during B-cell development. Given its regulatory function in *Rag* gene expression in T cells [33], the potential implication of SATB1 in V(D)J recombination of Ig genes in B cells needs to be examined. Moreover, since SATB1 was discovered by its ability to bind the *MAR<sub>Eu</sub>* region of the *IgH* locus [1], we postulated a role for SATB1 in Ig heavy chain expression.

In this study, we examined SATB1 involvement in B-lineage cell development and depicted its critical function in immunoglobulin production using a conditional KO mouse model allowing deletion of this factor in B-lineage cells. Although we found SATB1 to be nonessential for B-cell development, we showed that this factor displays an ambivalent function in late developing B-cell subsets, acting sequentially as a positive and then negative regulator of Ig gene transcription. The negative regulatory function of SATB1 extends to the germinal center reaction in which this factor limits Ig gene somatic hypermutation.

## MATERIALS AND METHODS

### Mice

*Satb1<sup>tm1a/wt</sup>* mice were obtained from the Mouse Clinical Institute (MCI) (IR00004167/P4167). First, crossings with 129S4/SvJae-*Gt(ROSA)26Sor<sup>tm2(FLP)</sup>/sor*/J mice (The Jackson Laboratory) removed the reporter and marker cassette placed between *Frt* sites. Second crossings with B6.C(CG)-*Cd79a<sup>tm1(cre)Reth</sup>/EhobJ* mice (The Jackson Laboratory) deleted *Satb1-exon4* due to CRE recombinase expression in B-lineage cells [36]. *Aicda<sup>-/-</sup>* homozygous mice (kindly provided by Pr. T. Honjo) were used to prepare control samples devoid of SHM, as reported in [37]. All experiments, except 3C-HTGTS, were performed on two-month-old mice. The primers used for genotyping are listed in Supplementary Table S1. All animal strains were bred and maintained in SPF conditions at 21–23 °C with a 12-h light/dark cycle. Procedures were reviewed and approved by the Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation APA-FIS#16639-2018090612249522v2.

### Western blotting

B cells from the spleen were sorted with an EASYSEP MOUSE B-cell isolation kit (Stem Cell Technologies) and lysed using RIPA buffer (Santa Cruz) supplemented with a protease inhibitor (Orthovanadate, PMSF, anti-protease cocktail). Proteins were quantified with a Pierce BCA Protein Assay kit (Thermo Scientific) and denatured for 5 min at 95 °C. TGX Stain-free FastCast 12% acrylamide gels (Bio-Rad Laboratories) were used to separate proteins that were transferred to Transblot Turbo polyvinylidene fluoride membranes (Bio-Rad Laboratories) with a Transblot Turbo Transfer System. After blocking incubation with phosphate-buffered saline (PBS)-5% skim milk, the membrane was incubated with SATB1 primary antibody overnight at 4 °C. The

membrane was incubated with secondary antibody in PBS-3% milk. Proteins were detected and quantified with PIERCE ECL Western blotting Substrate (Thermo Scientific) and ChemiDoc Touch Imaging System coupled to Image Lab J6.0 software (Bio-Rad Laboratories). The antibodies and concentrations used are listed in Supplementary Table S2. Unprocessed raw data are provided in Supplementary Fig. S1A.)

### Enzyme-linked immunosorbent assay (ELISA)

Sera were collected from the blood of two-month-old mice. Supernatants were obtained after sorting B cells from spleens and in vitro LPS activation for 4 days. Plates were coated overnight with 1 µg/ml of primary antibody goat anti-mouse unlabeled IgM, IgG3, IgG1 or IgA (Southern Biotech) diluted in sodium carbonate buffer. Plates were blocked with PBS-3% bovine serum albumin for 1 h at 37 °C. Supernatant or diluted sera were incubated for 2 h at 37 °C, and then secondary antibody conjugated with alkaline phosphatase (Southern Biotech) was incubated for 1 h at 37 °C. Enzymatic reactions were performed by adding SigmaFast™ p-nitrophenyl phosphate tablets. DO was measured at 405 nm on a Thermo Scientific MULTISKAN FC. The antibodies and concentrations used are listed in Supplementary Table S2.

### Flow cytometry

To analyze B-cell populations, spleen and bone marrow were collected, crushed and filtered through a 40 µm Clearline Strainer (Fisherbrand); lymphocyte numbers were quantified in a Cell-Dyn Emerald counter (Abbott). Cell suspensions were labeled in FACS buffer (PBS buffer (Eurobio), 10% BSA (Stem Cell Technologies) and 2 mM EDTA) containing antibodies for 20 min at 4 °C and analyzed on a BD LSRFortessa SORP flow cytometer (BD Bioscience). For intracellular staining, cells were first stained for surface markers. After washing, the cells were treated with the Intraprep fixation/permeabilization kit (Beckman Coulter, A07803) following the manufacturer's instructions. Mean fluorescence intensity values were normalized to WT values; antibodies used are provided in Supplementary Table S2. Gating strategies to evaluate the proportions and numbers of B-cell subsets are provided in Supplementary Fig. S1.

### RNA extraction

Pre-B cells, splenic resting B cells, in vitro-activated cells with LPS for 2 or 4 days, plasma cells and GC B cells were lysed with TRIZOL Reagent (Ambion Life Technologies), and RNA extraction was performed following the recommendations of the DirectZol RNA microprep kit (ZymoResearch).

### RT-qPCR

One microgram of RNA from B-cell subset samples was treated with DNase I Amplification Grade (Invitrogen), and reverse transcription was performed following the High Capacity cDNA Reverse Transcription kit protocol (Applied Biosystems). Quantitative PCR was performed on 20 ng cDNA using either SensiFAST Probe Hi-ROX or SensiFAST SYBR Hi-ROX kits (BioLine) on a Quant Studio III (Applied BioSystem). *IgH* and *Igk* primary transcripts were quantified as previously described [38]. *Satb1* transcripts were measured by TaqMan assays with Mm.PT.58.13287891 (IDT), *Rag 1* and *Rag 2* transcripts with Mm01270936\_m1 and Mm00501300\_m1 (ThermoFisher), respectively. Transcript quantification was carried out with normalization to *Hprt* (Mm.PT.58.32092191 - IDT) for resting and in vitro-stimulated cells and to *B220* (Mm.PT.58.32086646 - IDT) for GC B cells. Specific primers designed to amplify membrane and secreted *Igμ* transcripts, as well as *μ* transcripts, are listed in Supplementary Table S1.

### Repertoire

Pre-B cells from bone marrow were enriched using a mouse CD25 MicroBead Kit (Miltenyi Biotec). Library preparation was adapted from methods previously described [39, 40]. Transcripts were amplified by 5'RACE PCR using reverse primer hybridizing within the *μ* gene and Cap race primer carrying unique molecular identifiers (UMIs) [41]. ProtoScript® II (New England Biolabs, Ipswich, MA) was used for reverse transcription, and amplicons were obtained using Phusion® High Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA) according to the manufacturer's instructions. (Primers are listed in Supplementary Table S1). Illumina adapter and tag sequences were added by primer extension [42]. The resulting libraries were sequenced on an Illumina MiSeq sequencing system using MiSeq Reagent kit V3 600 cycles. Paired reads were merged as previously described [42] and UMIs were treated with MIGEC software.



Repertoire analysis was performed using the IMG2/HIGHV-QUEST online tool (<http://img2.org/>) [43].

### In vitro stimulation

Splenic B cells isolated with the EASYSEP MOUSE B-cell isolation kit (Stem Cell Technologies) were cultured at  $1 \times 10^6$  cells/mL in RPMI 1640 medium (Eurobio) supplemented with 10% fetal bovine serum (Dutscher), 2 mM glutamine (Eurobio), 1% Eagle's nonessential amino acids (Eurobio), 50 U/ml penicillin–streptomycin (Gibco), 1 mM sodium pyruvate (Eurobio), and 129  $\mu$ M 2- $\beta$ -mercaptoethanol (Sigma–Aldrich) in the presence of 1  $\mu$ g/ml lipopolysaccharide (LPS–Invivogen). An aliquot of cells was collected for analysis and sample preparation at Day 2 or 4 after stimulation. The same isolation protocol and cell culture reagents were used for anti-CD40 + IL4 in vitro stimulation by replacing LPS with 5  $\mu$ g/ml anti-CD40 antibody (R&D Systems, listed in Supplementary Table S2) and 20 ng/mL recombinant mouse IL4 (Peprtech). An aliquot of cells was collected for analysis and sample preparation at Days 3 and 4 after stimulation.

### In vitro ethynyl-deoxyuridine (EdU) incorporation

Splenic B cells were cultured as described above, and EdU was added 48 h later and incubated for 24 h. Cells were then processed by following the recommendations of the Click-it EdU Alexa Fluor 488 Flow Cytometry Assay Kit (Thermo Fisher Scientific). EdU incorporation was evaluated by flow cytometry on a BD LSRFortessa SORP flow cytometer (BD Bioscience).

### NP-CGG immunization

Nitrophenylacetyl–Chicken Gamma Globulin (NP-CGG) immunization was realized by injecting 100  $\mu$ g NP-CGG (N-5055C-5 Biosearch Technologies) precipitated with complete for the first intraperitoneal injection (Day 0) and incomplete Freund's adjuvant for the second one at Day 12. Preimmune sera were collected before Day 0, and immunized sera were collected at Days 8 and 17. Mice were sacrificed on Day 17, and splenic GC B cells were sorted to extract DNA for SHM quantification by HTS.

### Cell sorting

Plasma cells from spleen or plasmablasts from in vitro stimulation were sorted using B220 and CD138 surface markers on a BD FACS ARIA III (BD Bioscience). GC B cells were sorted using B220 and GL7 cell surface markers (Supplementary Table S2).

### Somatic hypermutation analysis

SHM analysis was performed on B220<sup>+</sup>/GL7<sup>+</sup> GC B cells sorted from Peyer's patches or immunized spleens from *wt*, *Satb1* cKO and *Aicda* KO mice. 5'*S* <sub>$\mu$</sub> , 3'*J*<sub>H</sub>4 and 3'*J*<sub>H</sub>5 intronic regions were amplified from 10,000 cells with specific primers (listed in Supplementary Table S1) using Phusion High-Fidelity DNA Polymerase (New England Biolabs) and according to the following program: denaturation (98 °C 10 s), hybridization (69 °C 30 s) and amplification (72 °C 1 min) for 38 cycles. Libraries were constructed with an Ion Xpress Plus gDNA Fragment Library kit (Cat. no. 4471269, Life Technologies) and sequenced on the Ion-Proton System S5. SHM frequencies were determined using raw data analyzed with the DeMinEr tool [37].

### RNA sequencing

RNA-Seq analysis was performed on splenic resting B cells sorted with an Easysep Mouse B-cell isolation kit (Stem Cell Technologies), in vitro-activated B cells with LPS (bulk culture, at Day 2) and in vitro LPS-differentiated plasmablasts (sorted at Day 4). Sample quality controls and library preparation were performed at the GeT-Santé facility (Inserm, Toulouse, France, [get.genotoul.fr](http://get.genotoul.fr)). i/Quality control: RNA concentration and purity were determined using an ND-2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, USA). RNA integrity was checked with a Fragment Analyzer (Agilent Technologies, Santa Clara, USA) using the RNA Standard Sensitivity Kit. 260/280 purity ratios were all  $\geq 1.8$ , and integrity indices were of good quality (8.3–10 RIN and  $>1.7$  28S/18S ratios). ii/Library preparation (GeT-Santé): RNA-seq paired-end libraries were prepared according to Illumina's protocol with some adjustments using the TruSeq Stranded Total RNA Gold Library Prep Kit (Illumina, San Diego, USA). Briefly, between 934–1000 ng of total RNA was first ribo-zero depleted using Illumina Ribo-Zero probes. Then, the remaining RNA was fragmented for 2 min and retrotranscribed to generate double-stranded

cDNA. Compatible adaptors were ligated, allowing the barcoding of the samples with unique dual indices. Libraries were amplified by 12 PCR cycles, and an additional final purification step resulted in 280–700 bp fragments. Library quality was assessed using the HS NGS kit on the Fragment Analyzer (Agilent Technologies, Santa Clara, USA). iii/Library quantification (GeT-PlaGe): quantification and sequencing were performed at the GeT-PlaGe core facility (INRAE, Toulouse, France). Libraries were quantified by qPCR using the KAPA Library Quantification Kit (Roche, Basel, Switzerland) to obtain accurate quantification. iiiii/Sequencing (GeT-PlaGe): libraries were equimolarly pooled, and RNA sequencing was then performed on one S4 lane of the Illumina NovaSeq 6000 instrument (Illumina, San Diego, USA) using the NovaSeq 6000 S4 v1.5 Reagent Kit (300 cycles), and a paired-end 2  $\times$  150 bp strategy.

### RNA-sequencing analysis

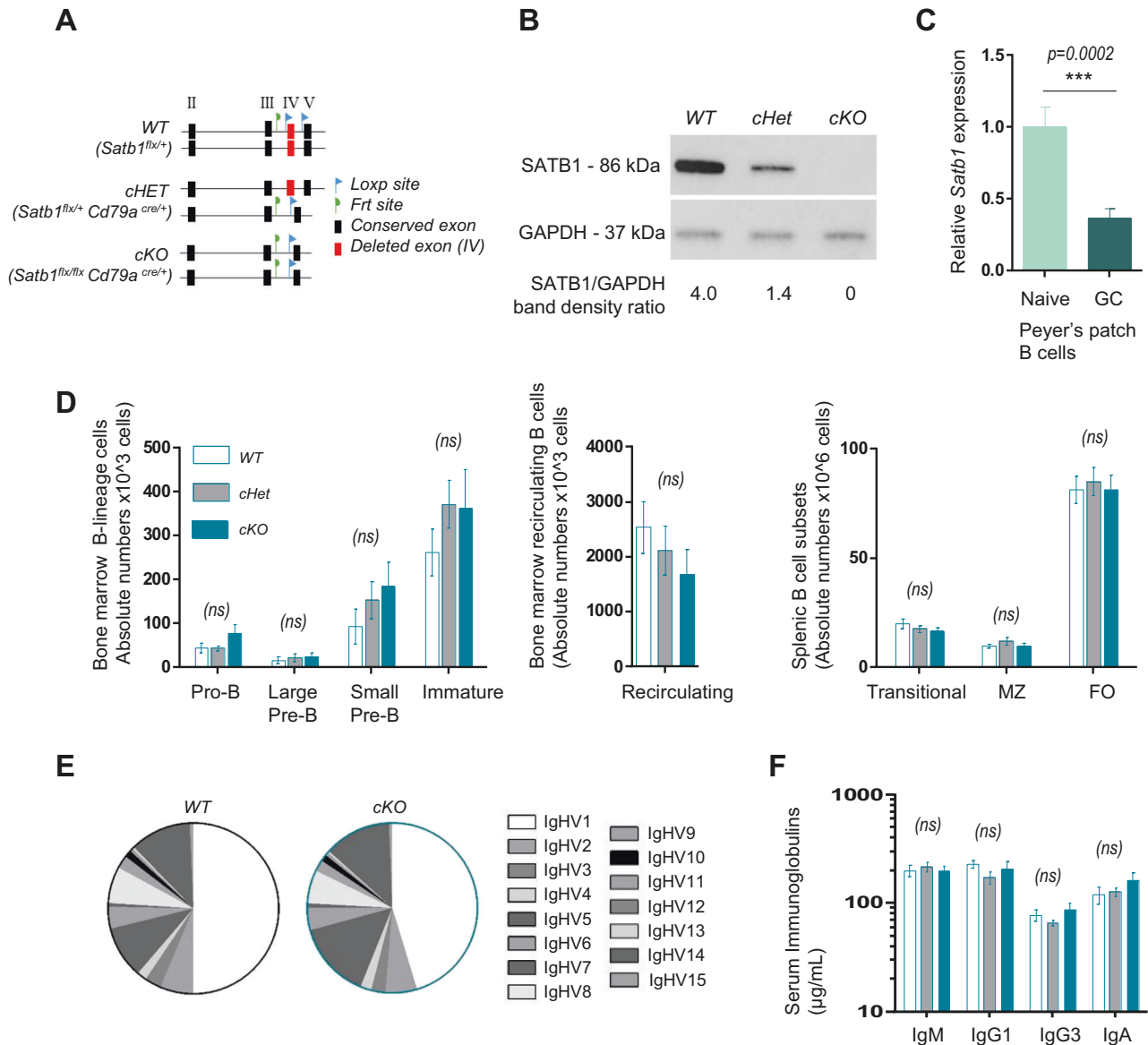
Paired-end reads were mapped on the GRCh38 mouse genome that was previously indexed with "Mus\_musculus.GRCh38.dna.primary\_assembly.fa" and "Mus\_musculus.GRCh38.102.chr\_patch\_hapl\_scaff.gtf" files from ENSEMBL release 102. Index and mapping steps were both performed with STAR v2.6.0c [44]. Then, featureCounts v2.0.1 [45] was used to count reads by gene. An R script named template\_script\_DE-Seq2\_CL.r of SARTools [46] was run first with all count data to retrieve a PCA and check if biological variability was the main source of variance in the data. Then, the same script was run for each desired differential analysis with count data from defined reference and interest conditions. Differentially regulated genes with an adjusted p value of 0.05 and a fold change  $\leq -1.5$  or  $\geq 1.5$  were selected for downstream analysis. Gene SYMBOLS were converted to ENTREZIDs with the bitr function of the R ClusterProfiler package [47]. The resulting ENTREZIDs and their associated log2fold change were then used to calculate enriched biological pathway profiles of different gene clusters (Down in cKO, Up in cKO, Down in WT and Up in WT) using the CompareCluster function of ClusterProfiler [47] with p value and q value thresholds set to 0.01 and 0.05, respectively. The resulting enriched functional profiles were filtered through a Gene Ontology list consisting of the hierarchical children of the following biological pathway terms: B-cell costimulation, B-cell selection, humoral immune response, immunoglobulin production, memory B-cell differentiation, regulation of the apoptotic process of mature B cells, and SHM (Supplementary Table S3). Terms that were enriched in all downregulated gene clusters or all upregulated clusters were discarded.

### 3C-HTGTS

3C-HTGTS was performed as previously described [48]. Briefly, 10 million resting and LPS-stimulated B cells (at Day 3) were crosslinked with 2% formaldehyde in 10% FCS PBS for 10 min at RT with rotation. Crosslinking was stopped by adding glycine at 0.1 M. Then, the cells were lysed in 50 mM Tris, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, and 1% TX-100 supplemented with protease inhibitor (ROCHE #11873580001). Nuclei were resuspended in 0.3% SDS for 1 h at 37 °C at 900 rpm and then neutralized with Triton TX-100 for 1 h. DNA restriction was performed using CviQ1 (Thermo Fisher ERO211) in B buffer (Thermo Fisher #BB5) overnight at 37 °C before heat inactivation for 25 min at 65 °C. Overnight ligation was performed at 16 °C and 300 rpm. Next, DNA was treated with proteinase K and RNase and cleaned with phenol/chloroform. After the 3C step, the LAM-HTGTS protocol was performed [49]. Briefly, 3C DNA was sonicated using the Bioruptor (Diagenode; two pulses at low intensity for 20 s), and 10  $\mu$ g was used for the LAM-HTGTS step.  $\mu$ m bait was used for primer elongation. These single-stranded DNA fragments were incubated with streptavidin beads (Dynabeads C1 streptavidin beads; Invitrogen) overnight at RT and washed with BW buffer (1 M NaCl, 5 mM Tris-HCl pH 7.4, 0.5 mM EDTA pH 8.0). Universal I7 adaptors were ligated before nested PCR was performed with  $\mu$ m nested and universal I7 reverse primers (all primers are listed in Supplementary Table S1). After the Tagged PCR with I7 and I5 Illumina primers, PCR products were cleaned using a PCR clean-up kit (Macherey-Nagel REF#740609) and validated after migration on a BioAnalyser (Agilent). 3C-HTGTS libraries were sequenced on a 300 bp paired-end MiSeq V3 with 20% PhiX.

### Analysis of 3C-HTGTS

Sequencing reads were aligned to the mm10 genome and processed as previously described [49]. Each 3C-HTGTS library plotted for comparison was normalized by randomly selecting a number of junctions equal to the total number of junctions present in the smallest library in the comparison set.



**Fig. 1** **A** Conditional deletion of *Satb1*-exon 4 by *Cd79a<sup>cre/+</sup>* recombinase in B cells. Exon 4, cre site, frt site, conditional allele (*Satb1<sup>flx</sup>*) and wild-type allele (*Satb1<sup>flx/+</sup>*) are indicated. **B** SATB1 protein quantification by Western blotting in splenic B cells from WT (*Satb1<sup>flx/+</sup>*), conditional heterozygous (*Satb1<sup>flx/+</sup> Cd79a<sup>cre/+</sup>*) and conditional homozygous (*Satb1<sup>flx/flx</sup> Cd79a<sup>cre/+</sup>*) mice. **C** *Satb1* transcript quantification by RT-qPCR (normalized to *Hprt*) in Peyer's patch naive and GC B cells sorted from WT mice ( $n = 8$ ) from 3 independent experiments. **D** Bar graphs displaying absolute numbers of B-cell subsets determined by flow cytometry from bone marrow (per femur,  $n = 7-10$ ) and spleen ( $n = 11-15$ ) from 4-5 independent experiments. **E** Pie chart representation of *IgHV* family gene usage, quantified by RACE-RepSeq, expressed by pre-B-cell-enriched bone marrow populations in WT ( $n = 5$ ) and cKO mice ( $n = 6$ ) from 2 independent experiments. **F** Immunoglobulin serum isotype levels quantified by ELISA in 8- to 10-week-old WT and cKO mice ( $n = 8-15$ ) from 4 independent experiments. Error bars represent SEM;  $p$  value was determined with two tailed Mann-Whitney test; nonsignificant (ns) and significant differences are indicated (\*\*\* $p < 0.001$ )

## RESULTS

### Conditional deletion of *Satb1* in murine B lineage cells

Since depletion of *Satb1* in ES cells and mouse embryos resulted in multiple dysfunctions and did not allow animals to survive beyond 1 month of age [2], we evaluated the SATB1 contribution to B-cell development in a conditional knockout model inducing SATB1 depletion in B lineage cells from the pro-B-cell stage. In our model, initially derived from the *Satb1<sup>tm1a</sup>* allele, the *Satb1* conditional allele (*Satb1<sup>flx</sup>*) contains exon 4 flanked with two *LoxP* sites that permit its specific deletion when coupled to one allele carrying a CRE recombinase insertion into the *Cd79a* gene (*Cd79a<sup>cre</sup>*) [36] (Fig. 1A). Mice carrying the homozygous SATB1 deletion in B-lineage cells (*Satb1<sup>flx/flx</sup> Cd79a<sup>cre/+</sup>*, referred as cKO conditional knockout) were

compared to heterozygous littermates (*Satb1<sup>flx/+</sup> Cd79a<sup>cre/+</sup>*, referred as cHet: conditional heterozygous) and to *wt* littermates carrying an identical conditional allele but devoid of CRE expression (*Satb1<sup>flx/+</sup> Cd79a<sup>+/+</sup>*, referred to as WT). Western blot experiments performed on sorted splenic B cells from these three genotypes confirmed that *Satb1*-exon4 deletion induced complete protein depletion (Fig. 1B, unprocessed raw data are provided in Supplementary Fig. S1).

### SATB1 depletion in B lineage cells allows normal B-cell development and IgH repertoire

By querying the ImmGen database [50], we examined *Satb1* expression in developing B-lineage cells and noted that all subtypes of B-lineage cells expressed *Satb1* transcripts

(Supplementary Fig. S2A). Common lymphoid progenitor and prepro B cells, also mentioned as Hardy fraction A, both displayed higher SATB1 expression. This was not surprising given that SATB1 has been reported to favor lymphocyte lineage differentiation from hematopoietic stem cells [23]. While a consistent drop in *Satb1* transcription was observed at pro- and pre-B stages, a second wave of *Satb1* gene expression occurred in transitional and mature resting B cells present in splenic follicles or the marginal zone as well as in splenic memory B cells. *Satb1* gene expression further decreased in antigen-activated cells such as germinal center centrocytes to reach its minimal level in proliferating cell subsets such as centroblasts and plasmablasts (Supplementary Fig. S2A). In B cells from WT mice, we experimentally confirmed a significant decrease in *Satb1* transcripts by RT-qPCR assays when comparing naive and GC B cells sorted from Peyer's patches (Fig. 1C). The same findings, corroborating expression profiles described in the ImmGen database, were very recently documented in a mouse model carrying a tomato-reporter transgene knocked into a *Satb1* allele [35]. This converging evidence for variable expression of *Satb1*, according to B-cell subsets, supposes a narrow regulation and suggests an accurate function of this factor in the B-cell lineage.

To determine the SATB1 contribution throughout B-cell development and maturation, we analyzed B-cell populations from WT, cHet, and cKO mice by flow cytometry (Supplementary Fig. S1). Absolute numbers of B-lineage cell subsets from bone marrow (Fig. 1D left and middle) and spleen (Fig. 1D right) were not affected by *Satb1* gene inactivation, suggesting, despite its contribution to lymphocyte lineage initiation [22], that SATB1 was not required for B-cell development. Our findings that early B-cell development was not impaired in our cKO mice excluded any function for SATB1 in B-lineage choice maintenance. These data were consistent with studies from Kanakura and Steidl groups [21, 22] that previously hypothesized that SATB1 function was restricted to stem cell renewal and fate of hematopoietic lineages.

Since SATB1 binding to the ASE region is known to modulate *Rag* gene expression in T cells and incidentally impact TCR rearrangements [33], the consistent expression of this factor in early committed and developing B cells could confer a similar function on Ig gene rearrangements. We investigated this point by first analyzing *Rag* gene expression by RT-qPCR and found that sorted bone marrow pre-B cells from WT and cKO animals expressed similar transcript levels (Supplementary Fig. S2B). The strikingly different impact of SATB1 deletion on *Rag* expression in B and T cells could be indirectly explored by comparing DNA accessibility in the region close to *Rag* gene regions from WT B- and T-lineage cell precursors. Such a detailed study was provided by the consortium of the Immunological Genome Project that performed ATAC-seq in B and T-cell precursors [51]. This study indicates that while the DNA region encompassing the SATB1-binding ASE regulatory element was accessible in T-lineage precursor cells, this was not the case in bone marrow cell subsets of the B-cell lineage (a comparison focused on *Rag* gene loci is provided in Supplementary Fig. S2C). Notably, B-lineage precursors display other accessible regulatory regions, such as the *Erag* element located close to *Rag* gene regions. It is then reasonable to suppose that the ASE region is not active in the B-cell lineage, which could potentially explain normal *Rag* gene transcription in SATB1-deficient B-cell precursors. Given that SATB1 acts as a chromatin loop organizer in T cells [5, 34, 52–54], we suspected a potential effect of its deletion on *IgH* V region accessibility in developing B cells. To assess this point, we examined VDJ recombination diversity by repertoire sequencing experiments on RNA samples (RepSeq) extracted from pre-B-cell-enriched bone marrow fractions. Our data displayed an equivalent broad distribution of each rearranged and expressed V gene (Fig. 1E and Supplementary Fig. S2D) in WT and cKO mice. A similar

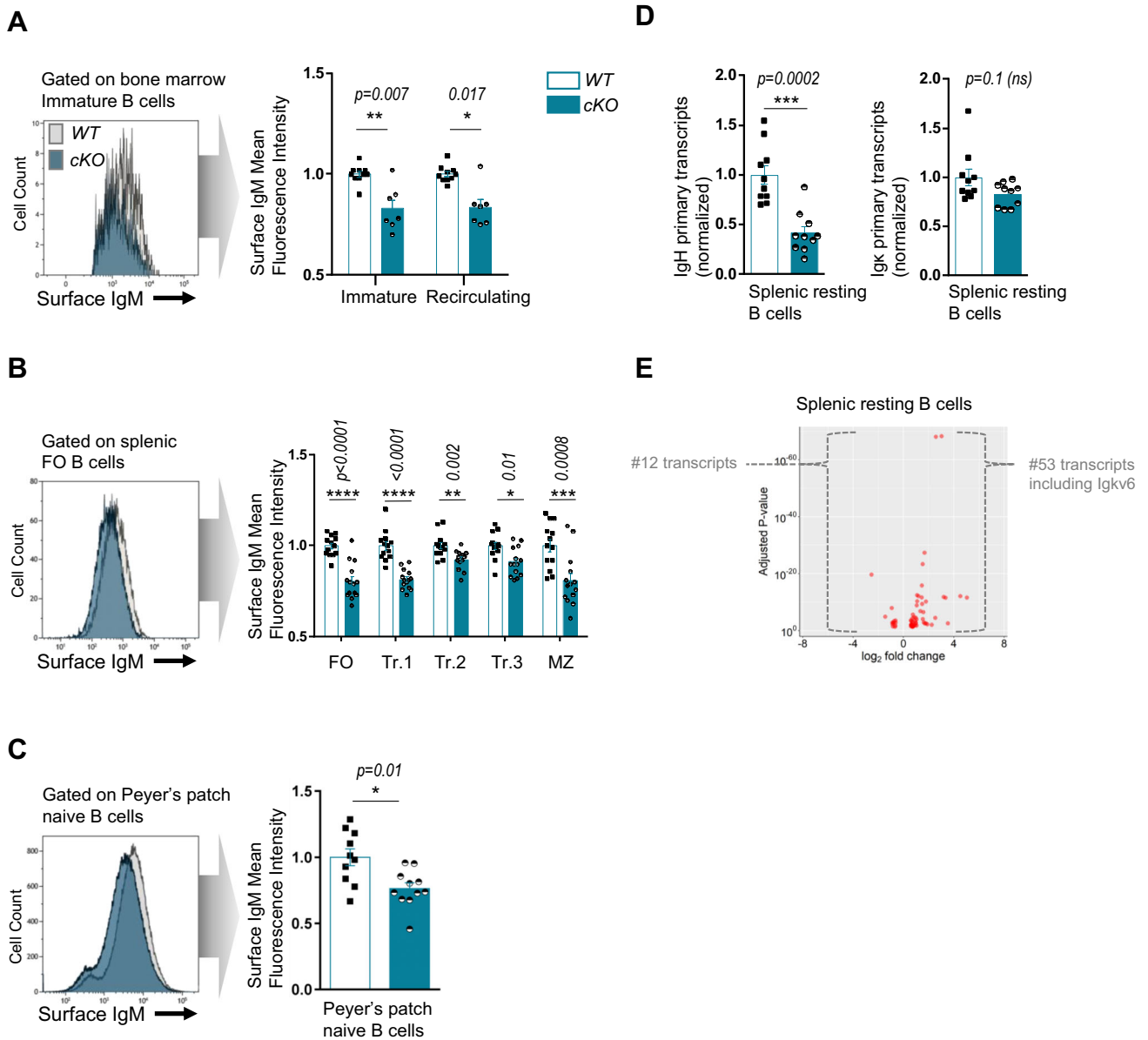
representation of  $V_H$  family usage indicated that SATB1 deletion did not hamper mechanisms leading to a diversified *IgH* VDJ repertoire in developing B cells. VDJ junction analysis in the preimmune repertoire of WT and *Satb1* cKO models revealed comparable lengths of CDR3 regions as well as a rather comparable distribution of P and N nucleotides (Supplementary Fig. S2E), suggesting that *IgHV* region assembly and end-joining occur normally in the absence of SATB1. Altogether, our results indicate that SATB1 is dispensable for establishing the *IgH* repertoire and does not affect *Rag* expression or influence *IgH* locus accessibility. Given that SATB1 deletion did not impair VDJ rearrangements and B-cell development, further studies could then be performed on peripheral B-cell subsets in this model without any bias. In line with this statement, we determined whether SATB1 deletion could impact Ig isotype production and secretion in mice. Sera from WT, cHet and cKO animals were collected at two months of age, and IgM, IgG1, IgG3, and IgA levels were quantified by ELISA. Homozygous and heterozygous KO mice displayed serum Ig levels comparable to WT for each isotype (Fig. 1F), suggesting that SATB1 does not influence global antibody production.

### SATB1 depletion decreases IgH transcription in resting B cells

Despite normal bone marrow B-cell development and Ig secretion in cKO mice, we suspected that SATB1 depletion might influence *IgH* locus expression, given its capacity to bind MAR sequences in vitro [55] and to modulate MAR-containing reporter genes [29]. We first examined surface IgM expression levels, as a component of B-cell receptors, on immature and mature B-cell subsets from bone marrow and spleen compartments in our model. Upon measuring IgM mean fluorescence intensity (MFI) by flow cytometry, *Satb1* cKO mice displayed consistent and significantly decreased IgM surface expression compared to WT mice on bone marrow immature and recirculating B cells (Fig. 2A) and on splenic transitional, marginal zone and follicular B-cell populations (Fig. 2B). The same significant decrease in IgM BCR expression, also observed on naive B cells sorted from Peyer's patches (Fig. 2C), confirmed a specific function for SATB1 as a positive regulator of BCR expression in resting B cells. This role for SATB1 in BCR surface expression could explain its function as a contributing factor to B-cell survival, as described recently by Ozawa et al. [35]. Indeed, the knock-in reporter mouse model used in their study leads to inactivation of one *Satb1* allele, and it is likely that such a deletion induces an intermediate surface IgM expression level similar to the one we observed in our heterozygous mice (Supplementary Fig. S3A). It is reasonable to suppose that reduced BCR expression on resting B cells reduces the ability of these cells to respond to anti-IgM-mediated in vitro stimulation. One fair interpretation of the survival defect described by Ozawa et al. [35] could be simply the consequence of reduced BCR expression in the knock-in model.

The decreased BCR expression observed in SATB1-deficient B lymphocytes raised the question of a potential Ig gene transcription defect in our model. Quantification of *IgH* primary transcription by RT-qPCR in resting splenic B cells showed a significant 2-fold decrease in transcripts running through the  $J_H4$  intron in cKO mice compared to WT mice (Fig. 2D). We also evaluated the respective proportion of membrane and secreted *Igμ* transcripts (Supplementary Fig. S3B). Our data showed that SATB1 deletion does not impair alternative splicing of the *Igμ* chain transcript. Although not statistically significant, a similar downward trend was also observed for *Igκ* transcription in mutants (Fig. 2D). MAR sequences flanking both sides of the *IgH*  $c\mu$  intronic enhancer are able to bind SATB1 [56]. Whereas binding of equivalent regions in the *Igκ* locus has never been shown, it has to be noted that only one upstream MAR is associated with the  $iEκ$  enhancer. The presence of either one or two MAR sequences surrounding these intronic enhancers could





**Fig. 2** **A** Comparison of surface IgM fluorescence intensity, evaluated by flow cytometry, on immature and recirculating bone marrow B-cell subsets from two-month-old WT and homozygous cKO animals. One representative histogram is shown (left), and the bar graph (right) displays normalized surface IgM MFI in WT and cKO animals ( $n=7-10$  from 4 independent experiments). **B** Comparison of surface IgM fluorescence intensity, evaluated by flow cytometry, on follicular (FO), marginal zone (MZ), and transitional (Tr.1, Tr.2, Tr.3) splenic B-cell subsets from two-month-old WT and homozygous cKO animals. One representative histogram is shown (left), and the bar graph (right) displays normalized surface IgM MFI in WT and cKO animals ( $n=13-14$  from 5 independent experiments). **C** Comparison of surface IgM fluorescence intensity, evaluated by flow cytometry, on naive B cells sorted from Peyer's Patches from two-month-old WT and homozygous cKO animals. One representative histogram is shown (left), and the bar graph (right) displays the normalized surface IgM MFI in WT and cKO animals ( $n=10-11$  from 3 independent experiments). **D** *IgH* and *Igk* primary transcripts quantified by RT-qPCR (normalized to *Hprt*) in splenic resting B cells sorted from WT and cKO mice ( $n=6-10$  from 5 independent experiments). **E** Volcano plot from RNA-seq indicating differentially expressed transcripts when comparing WT and cKO mouse ( $n=2$ ) resting B cells. Transcripts located within Ig gene loci are indicated. Error bars represent SEM;  $p$  value was determined by two tailed Mann-Whitney test; nonsignificant (ns) and significant differences are indicated (\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; \*\*\*\* $p < 0.0001$ )

potentially explain the differences observed for *IgH* and *Igk* transcription in our *Satb1* KO model. It is tempting to speculate that, through its DNA binding in proximity to intronic enhancers, SATB1 could physiologically potentiate their transcriptional effect in resting B cells. However, this hypothesis is unlikely since IgM BCR expression was never compromised in resting B-cell populations of mice devoid of either *E $\mu$* , *iE $\kappa$*  full regions or their associated-MARs [57–60]. In contrast, the literature reports that BCR expression on resting B cells was decreased upon

deletion of components of the 3' regulatory regions of both *IgH* [61] and *Igk* [62] loci, arguing for the importance of close contact between 3' enhancers and respective promoters of rearranged V exons. In resting B cells, such loops have been reported within the *IgH* locus in many studies [63]. In line with a proposed function for SATB1 as a promoter-enhancer loop regulator [54], a rational hypothesis could be that this MAR binding protein participates in physical interactions between rearranged *V<sub>H</sub>* exons and the 3'RR and consequently enhances Ig chain

transcription in resting B cells. To evaluate such chromatin loop interactions, we performed 3C-HTGTS experiments using a bait within the  $S\mu$  region [48] of resting B cells from wt and mutant mice. Our data showed that loop interactions between  $E\mu$  and  $V_H$  regions still occur at frequencies that do not significantly differ in resting B cells from both genotypes (Supplementary Fig. S4A); this suggests that SATB1 is not required to form chromatin loops in naive and LPS-activated B cells.

In line with the broad chromatin organization function proposed for SATB1 in T cells, we then performed total RNA-seq analysis of both WT and SATB1-deficient B-cell subsets. The resulting datasets were submitted to principal component analysis to validate reproducibility between samples (Supplementary Fig. S5A). When comparing datasets obtained from resting splenic B cells, our analysis only revealed 65 genes displaying significant changes in expression: 53 were overexpressed, and only 12 were downregulated ( $\log_2$  FC > +/−1.5; Fig. 2E and Supplementary Table S4). Indeed, our data support the hypothesis of an ambivalent function for SATB1, displaying both positive and negative regulatory actions for gene expression in resting B cells. Our study indicates that SATB1 depletion does not induce drastic transcriptional changes in resting B cells. In contrast to the T lineage [7, 27, 64], our data suggest that the intrinsic function of SATB1 in B cells may be relatively focused. In this cell type, SATB1 would have a distinctly different role from that of a major genome organizer.

### SATB1 depletion increases IgH transcription and Ig synthesis in activated B cells

We sought to evaluate the effect of SATB1 depletion on B-cell activation in response to mitogenic and antigenic stimuli. We first performed in vitro stimulation of splenic B cells from both genotypes with LPS and carried out RNA-seq in a time-course manner on resting B cells (sorted at Day 0), activated B cells (bulk culture at Day 2) and plasmablasts (sorted at Day 4). Once sample reproducibility was validated (Supplementary Fig. S5A), we compared transcriptional programs induced by the transition from “resting to activated” and “activated to plasmablast” cell stages. Venn diagrams mainly indicate that both B-cell activation and differentiation programs induced by LPS were not drastically impaired upon SATB1 depletion. Indeed, following in vitro culture, a vast majority of transcripts displayed a common regulatory profile in WT and SATB1-deficient models, *i.e.*, 5460 transcripts during the “resting to activated” transition and 2758 during the “activated to plasmablast” transition (Supplementary Fig. S5B). Nevertheless, we submitted differentially expressed genes from both transition programs to a gene ontology analysis for biological processes relevant to B-cell activation (see list in Supplementary Table S3). During the “resting to activated” transition, SATB1 depletion decreased the expression of a few genes involved in humoral immune responses mediated by circulating immunoglobulins (Fig. 3A, red dots), including *Lta*, a gene recently described as a SATB1 target in T cells [54], as well as *Tgfb1* encoding TGF $\beta$ , the cytokine involved in the regulation of isotype switching to IgA. Strikingly, the analysis specified that SATB1 depletion increased *Tcf3* expression in activated B cells (Fig. 3A, blue dot), while its gene expression, encoding the E2A transcription factor, was unchanged in hematopoietic stem cells devoid of SATB1 [23]. During the “activated to plasmablast” transition, SATB1 depletion downregulates *Cd40* and a few genes involved in the Ig class switching pathway (Fig. 3B, red dots); this effect on *Cd40* targeting was expected since it was already observed upon SATB1 deletion in hematopoietic stem cells and T cells [23, 54].

In parallel, in such in vitro stimulation assays, B cells were tested for intrinsic abilities to proliferate, to undergo LPS-driven class switch recombination and to differentiate into plasmablasts. When measured by EdU incorporation, the ability of SATB1-deficient B cells to proliferate was similar to that of WT cells (Supplementary

Fig. S6A). This suggested that while decreasing IgM BCR expression on naive cells, SATB1 deficiency is dispensable for cell cycle entry of splenic B cells upon activation of TLR pathways. The same normal proliferation capacity in response to TLR or CD40 triggering was also recently reported for B cells devoid of the entire *Satb1* gene region [35]. In our model, normal in vitro proliferation observed in B cells devoid of SATB1 was consistent with the normal proportion of spontaneous GC B cells in Peyer's patches of KO animals (Fig. 3C) or the equivalent proportions of splenic GC centroblasts and centrocytes obtained after NP-CGG immunization (Supplementary Fig. S6B). More strikingly, in clear contrast with the decreased *IgH* expression in resting B cells, such in vivo-activated cells express higher BCR levels, as evidenced by IgM surface labeling (Fig. 3D). This highlights that the ambivalent function for SATB1 clearly applies to Ig gene expression in developing mature B cells.

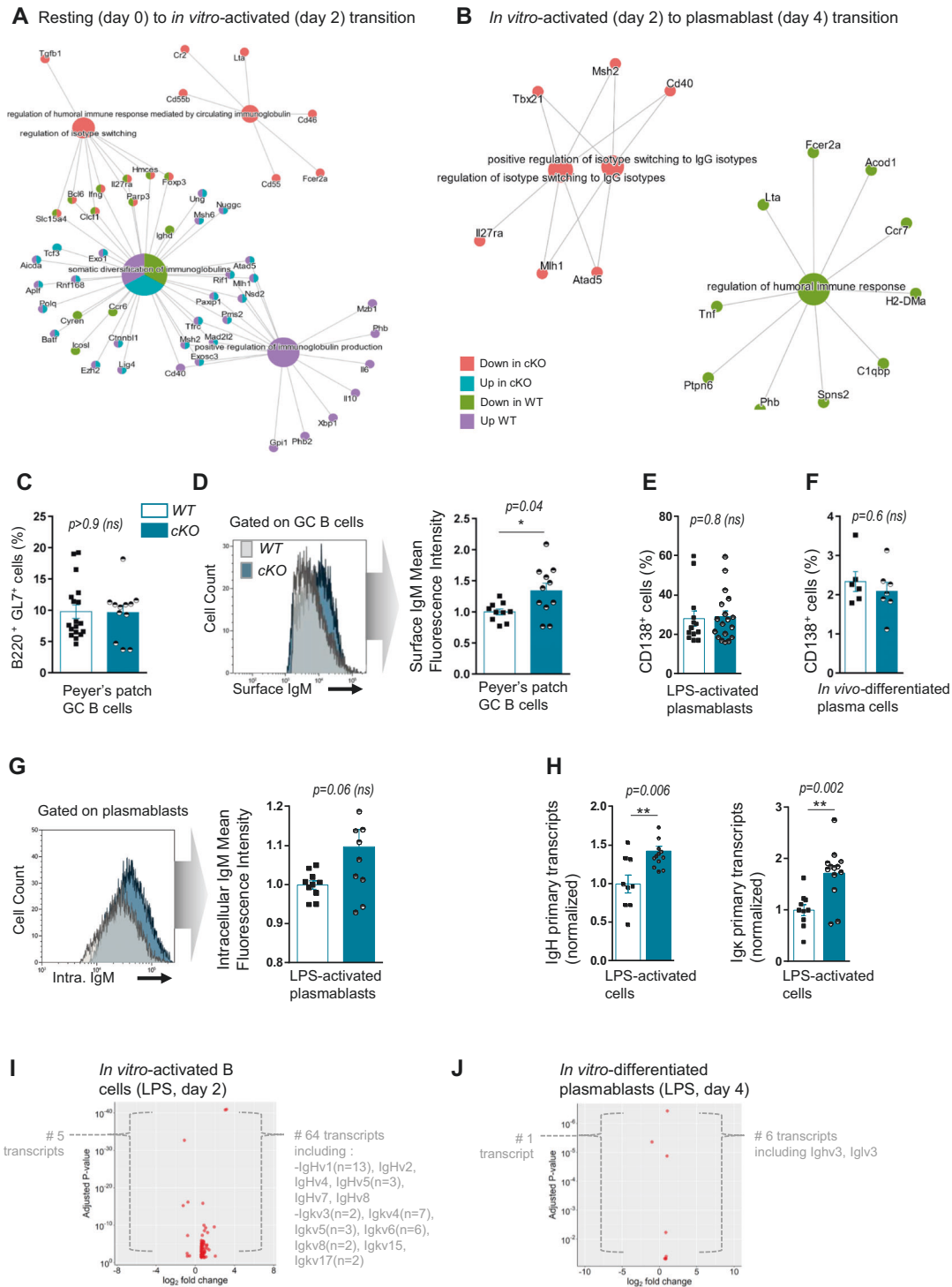
Class switching to IgG3 was examined by surface labeling of in vitro-activated splenic B cells by LPS at Day 4. The data showed equivalent numbers of IgG3-expressing cells in both WT and SATB1 KO models (Supplementary Fig. S6C), indicating that SATB1 is not required for IgG3 class switching. RNA-seq datasets from LPS-activated cells at Day 2 also provided relevant portrayals of the germline transcription landscape in the *IgH* locus of WT and SATB1-deficient models (Supplementary Fig. S6D). When examining germline transcription within the *I $\mu$* -donor and *Iy3*- or *Iy2b*-acceptor regions, both WT and SATB1-deficient models displayed identical profiles, suggesting that, in agreement with normal CSR to IgG3 observed at Day 4, SATB1 depletion does not impede LPS-induced germline transcription occurring prior to CSR.

When measured in the same in vitro assays, the intrinsic ability of splenic B cells to differentiate into plasmablasts (CD138<sup>+</sup>) upon LPS activation was also similar in WT and our SATB1 KO model (Fig. 3E). In agreement with the in vitro data, NP-CGG immunization of mice induced the generation of plasma cells in normal proportions in the absence of SATB1 (Fig. 3F). Altogether, these data suggest that SATB1 is dispensable for efficient differentiation into antibody-secreting cells.

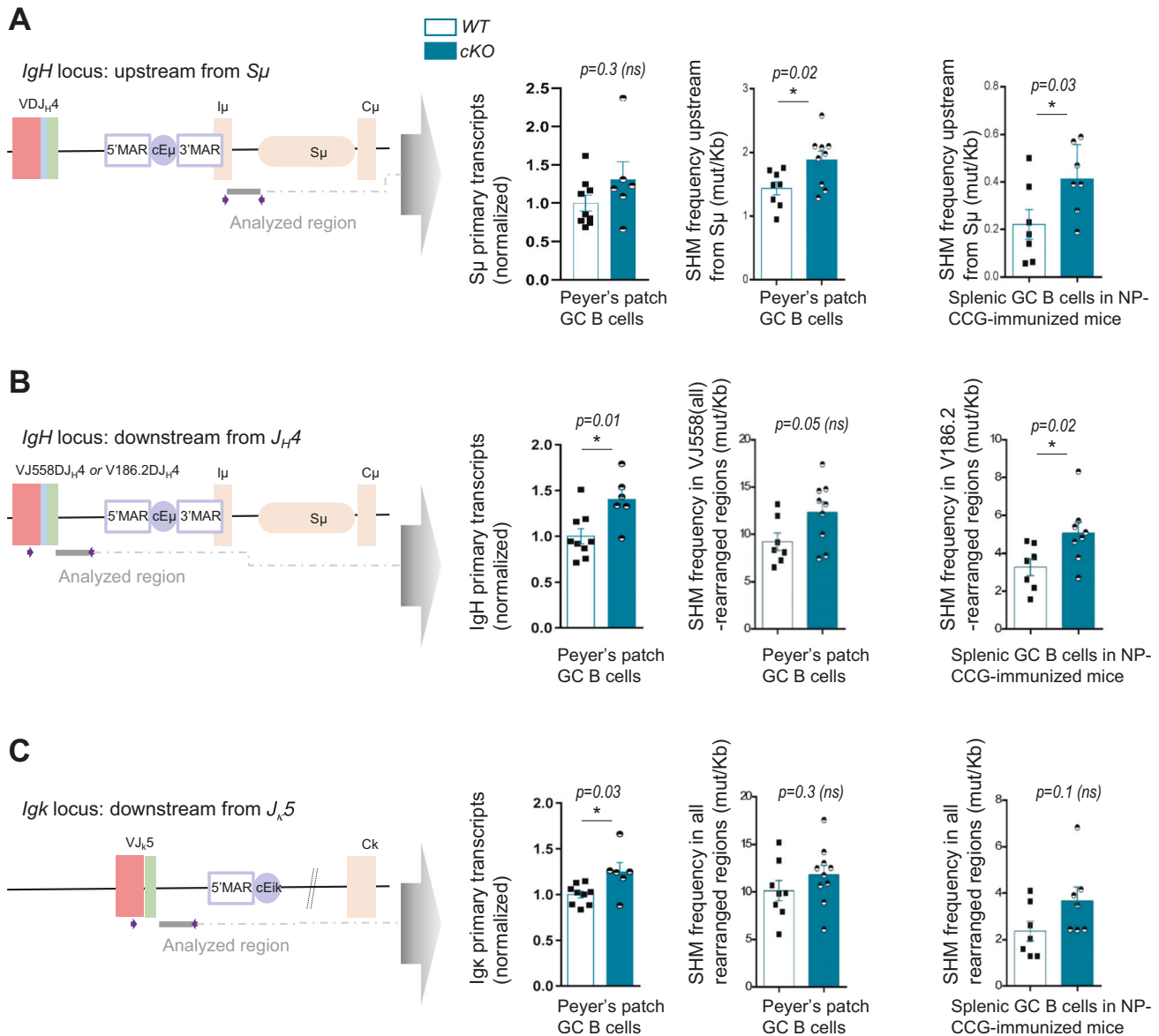
However, once differentiated in vitro into plasmablasts, SATB1-deficient cells displayed higher levels of intracellular IgM when analyzed by flow cytometry (Fig. 3G). This increase in Ig production by plasmablasts was correlated with a significant increase in both *IgH* and *Igk* primary transcripts (1.4 and 1.7, respectively) in LPS-activated cells from SATB1-deficient animals (Fig. 3H).

The effect of SATB1 deletion on global transcription of both IgH and IgL chains was also confirmed by comparing RNA-seq data from splenic B cells activated by LPS at Day 2 (Fig. 3J) or from sorted plasmablasts at Day 4 (Fig. 3K). In the first case, among the 64 genes found to be upregulated in SATB1-deficient models ( $\log_2$  FC > +/−1.5), 20 corresponded to *IgV* H chain products, and 23 others were identified as *IgV*  $\kappa$ L chain products (Fig. 3J and Supplementary Table S5). When comparing plasmablasts, *IgV* H chain and *IgV*  $\lambda$ L chain products were strongly increased upon SATB1 depletion (Fig. 3K and Supplementary Table S6). Our data indicate that, beyond its modest effect on B-cell activation programs induced by LPS in vitro, the absence of SATB1 induces a more pronounced transcription of Ig genes. Moreover, SATB1 is dispensable for proliferation, CSR and differentiation into plasma cells.

In parallel to LPS stimulation, we performed in vitro stimulation with anti-CD40 + IL4 in our model to evaluate, in another activation context, cell differentiation and Ig intracellular content in the absence of SATB1. Our data show, as in LPS stimulation, that SATB1-deficient B cells differentiate efficiently into plasmablasts in response to the anti-CD40 + IL4 cocktail (Supplementary Fig S6E left). Notably, plasmablasts devoid of SATB1 also contained high levels of intracellular IgM, although upregulation was not, in this case, increased significantly (Supplementary Fig S6E right).



**Fig. 3** **A** Visual representation of Gene Ontology enrichment analysis of genes significantly associated with biological pathways upon B-cell transition from resting (Day 0) to *in vitro*-activated (Day 2) subsets in WT and cKO mice. Datasets from RNA-seq experiments were used for the analysis. **B** Same representation as above for B-cell transition from *in vitro*-activated (Day 2) to plasmablast (Day 4) subsets. **C** Germinal center (GC) B-cell subsets evaluated by flow cytometry in Peyer's patches of WT and cKO animals ( $n = 12-18$  from 4 independent experiments). **D** Comparison of surface IgM fluorescence intensity evaluated by flow cytometry on Peyer's patch GC B cells from two-month-old WT and homozygous cKO animals. One representative histogram is shown (left), and the bar graph (right) displays normalized surface IgM MFI in WT and cKO animals ( $n = 10-11$  from 3 independent experiments). **E** Proportion of plasmablasts evaluated by flow cytometry, obtained following *in vitro* LPS activation (Day 4) of splenic B cells from WT and cKO animals ( $n = 13-18$  from 6 independent experiments) (left). **F** Proportion of plasma cells in the spleen evaluated by flow cytometry following NP-CGG immunization of WT and cKO animals ( $n = 6-7$  from 1 experiment). **G** Comparison of intracellular IgM fluorescence intensity evaluated by flow cytometry in *in vitro*-differentiated plasmablasts (Day 4) from WT and homozygous cKO animals. One representative histogram is shown (left), and the bar graph (right) displays normalized intracellular IgM MFI in WT and cKO animals ( $n = 9-10$  from 4 independent experiments). **H** *IgH* and *Igk* primary transcripts quantified by RT-qPCR (normalized to *Hprt*) in *in vitro* LPS-activated B-lineage cells (Day 4) of WT and cKO animals ( $n = 9-13$  from 5 independent experiments). **I** Volcano plots from RNA-seq indicating differentially expressed transcripts when comparing WT and cKO mice ( $n = 2$ ) in *in vitro* LPS-activated B cells (Day 2). Transcripts located within Ig gene V regions are indicated. **J** Same analysis as above in *in vitro* LPS-differentiated plasmablasts (Day 4). Error bars represent SEM;  $p$  value was determined by two tailed Mann-Whitney test; nonsignificant (ns) and significant differences are indicated (\* $p < 0.05$ ; \*\* $p < 0.01$ )



**Fig. 4** Comparison of primary transcription and somatic hypermutation in WT and cKO animals, taking place in various regions of Ig genes targeted by AID (left schemas), in spontaneous GC B cells sorted from Peyer's patches ( $n = 7-10$  from 4 independent experiments, left and middle bar graphs). Somatic hypermutation in splenic GC B cells sorted after NP-CCG immunization ( $n = 7-8$  from 4 independent experiments, right bar graphs). Data were obtained by NGS (Ion proton) combined with DeMinEr filtering, an accurate dedicated method to quantify AID-induced substitutions in B-lineage cells (50). **A** Transcription and SHM upstream from the *IgH* *S $\mu$*  region. **B** Transcription and SHM within the *IgH* intron downstream from VJ558 (consensus) to *J<sub>H</sub>4*-rearranged exons (in spontaneous GC B cells) or in the same region downstream from V186.2 to *J<sub>H</sub>4*-rearranged exons (in NP-CCG-induced GC B cells). **C** Transcription and SHM within the *Igk* intron downstream from all-rearranged *J<sub>k</sub>5* segments. Error bars represent SEM;  $p$  value was determined by two-tailed Mann-Whitney test; nonsignificant (ns) and significant differences are indicated (\* $p < 0.05$ )

However, our data unveil a switch in the regulatory function of SATB1 for Ig chain transcription, from positive in naive to negative in activated B cells. This striking effect persists until the terminal stage since SATB1-deficient plasma cells display a high content of Ig chains. However, SATB1 depletion does not allow plasma cells to produce more Ig since our deficient mice display broadly normal levels of serum antibody isotypes.

#### SATB1 depletion increases somatic hypermutation

Since our group recently reported a critical function of *cE $\mu$* -associated MARs for SHM of the *IgH* locus in mouse models [65], and according to historic studies in the *Igk* locus [59], it was questionable whether MAR binding proteins, such as SATB1, could

be involved in targeting somatic mutations of Ig genes. We first evaluated the ability of SATB1-deficient B cells to support somatic mutations by analyzing global SHM within intronic regions, not subject to antigen selection, located immediately downstream from V exons of both *IgH* and *Igk* loci and upstream from *IgH* *S $\mu$*  regions in spontaneous GC B cells from Peyer's patches. As a possible regulatory mechanism, transcription taking place in the corresponding regions targeted by AID was also quantified in this subset. Second, to more accurately evaluate the ability of SATB1-deficient B cells to undergo SHM in response to specific antigens, we immunized mice with NP-CCG [66]. By analogy to B cells devoid of *MAR<sub>S $\mu$</sub>*  regions, in which SHM machinery gains access more frequently to the region downstream from *cE $\mu$*  (upstream



from  $S\mu$  [67], we carefully quantified SHM in this same region (Fig. 4A, scheme). In GC B cells sorted from Peyer's patches, while  $S\mu$  transcription was not significantly increased (Fig. 4A, left bar graph,  $p = 0.3$ ), we found a significant increase in mutations in the absence of SATB1 (1.88 mut/Kb) compared to WT SATB1 (1.44 mut/Kb) (Fig. 4A, middle bar graph,  $p = 0.02$  and Supplementary Table S7A). A similar and significant increase in SHM in this region was also observed in splenic GC B cells sorted after NP-CGG immunization: while WT cells barely reached 0.22 mut/Kb, SATB1-deficient cells underwent twofold more mutations, reaching 0.41 mut/Kb (Fig. 4A, right bar graph,  $p = 0.03$  and Supplementary Table S7B). This common feature shared by  $MAR_{S_{E\mu}}$  KO- and *Satb1* KO-B cells suggests that SATB1 could be involved in limiting access of the SHM machinery to donor  $S\mu$  regions in cells undergoing SHM. When transcription and SHM were quantified downstream from the rearranged *IgH* variable regions (Fig. 4B, scheme) in GC B cells sorted from Peyer's patches, homozygous *Satb1* deletion increased transcription (Fig. 4B, left bar graph,  $p = 0.01$ ) downstream from  $J_H4$  and raised, although not significantly, SHM frequencies (Fig. 4B, middle bar graph,  $p = 0.05$  and Supplementary Table S7C). These findings were clearly confirmed upon NP-CGG immunization. Since the B-cell response to NP-CGG challenge is preferentially dominated by mutated clones expressing the  $V_H186.2$  segment [66], quantification of base substitutions in the  $J_H4$  intron downstream from the  $V_H186.2$ -rearranged exons is considered a reliable hallmark of antigen-induced SHM. Indeed, in SATB1-deficient GC B cells from immunized mice, NP-CGG-induced mutations were significantly increased within the  $J_H4$  intron: mutant cells displayed 5.06 mut/Kb, while WT cells only reached 3.28 mut/Kb (Fig. 4B, right bar graph,  $p = 0.02$  and Supplementary Table S7D). When, in the intronic region downstream from *Jk5* of the *Igk* locus (Fig. 4C, scheme), global transcription was significantly increased in mutant samples, (Fig. 4C, left bar graph,  $p = 0.03$ ), both spontaneous and NP-CGG-induced GC B cells displayed an increased trend of mutation frequencies upon SATB1 depletion (Fig. 4C, middle and right bar graphs and Supplementary Table S7E, F). Since B-cell responses to DNP hapten preferentially involve Ig composed of  $Ig\lambda 1$  light chains [68], it is likely that changes in mutation frequency within the *Igk* light chain loci, in this case not significant, probably underestimated any potential SHM increase induced by SATB1 depletion. While our Gene Ontology enrichment analysis detected the "Somatic diversification of Igs" as a differentially regulated pathway during the "resting to activated" transition, the expression of major actors of mismatch repair (MMR) and base excision repair (BER) pathways, such as *Ung*, *Msh2*, *Msh6*, *Pms2*, and *Polq*, was similarly upregulated in both wt and cKO samples (Fig. 3A). Beyond global pathway analysis, it was also necessary to verify in detail whether SATB1 depletion could affect, through its transcription factor function, the expression of any potential factors, including AID, involved in SHM. Indeed, when extracted from RNA-seq datasets, the expression of 16 genes involved in SHM was unchanged upon SATB1 deletion in both LPS-activated B cells and LPS-induced plasmablasts (Supplementary Fig. S7A). Within Ig gene regions submitted to SHM analysis, substitution frequencies calculated at each base displayed a globally increased pattern (although not significant) that did not offer any hypothesis regarding the origin of the changes (Supplementary Fig. S7B). It is well established that AID targeting SHM is coupled to transcription initiated at V promoters [69]. Given the significant increase in primary transcription occurring at *IgH* and *Igk* in GC B cells devoid of SATB1 (Fig. 4B, C), one straightforward hypothesis to explain the global SHM increase in our model could be an overall increase in AID targeting of Ig genes. Interestingly, in line with this hypothesis, we recently highlighted increased AID deamination coupled to increased transcription in our mouse model devoid of  $MAR_{S_{E\mu}}$  regions when bred into a DNA repair-deficient background [67]. Surprisingly,

while genomic deletion of  $MAR_{Eik}$  or  $MAR_{S_{E\mu}}$  regions in the mouse decreased SHM [59, 67], suppression of SATB1 led to the opposite effect in the regions downstream from the rearranged variable segments of the *IgH* and *Igk* loci. This finding also suggests a protective function for SATB1 against mutations. In this case, since the regions involved are located upstream from  $E\mu$  and  $Eik$ , one could propose that the potential protective effect of SATB1 could take place in resting B cells in which SHM is not expected to happen. Moreover, the discrepancy between the effect of  $MAR_{S_{E\mu}}$  deletion and SATB1 depletion on targeted mutations upstream from intronic enhancers suggests that other interacting factors, beyond SATB1, participate in the complex regulation of cis-acting MARs. While we recently proposed that  $MAR_{S_{E\mu}}$  favors error-prone repair in its upstream regions to optimize SHM [67], our current study suggests that SATB1 does not participate in the unfaithful repair processes associated with SHM.

Taken together, our data suggest that SATB1 deletion increases the SHM of Ig genes through a transcription-coupled mechanism that probably favors AID targeting.

## DISCUSSION

The fact that SATB1 plays a major role as a "genome organizer" in hematopoietic- and T-lineage cells [7, 28, 30] suggested that it is also important in B-lineage cells, a cell type that also undergoes fine developmental regulation of its expression [35]. By using a conditional deletion model in B cells, our study fills in gaps concerning the function of SATB1 in this lymphocyte lineage. In contrast to its function in early-developing T cells, SATB1 is dispensable for the establishment of the Ig gene repertoire and overall early B-cell development in the bone marrow; this result is consistent with the pioneer SATB1-null mouse model in which B-cell populations could be observed in the spleen [2].

When looking at regions involved in *Rag* gene regulation, accessibility data suggest that, unlike its critical function in T-cell precursors, the ASE region is not active in B-lineage cells. This difference suggests that the *Rag* gene regulation process evolved separately in the B and T lineages, at least with regard to the regulatory function of SATB1. However, our model reveals that SATB1 is involved in the control of Ig gene transcription in mature B cells. Although previously unknown in the B lineage, our findings once again point to a dual function for SATB1 depending on the activation stage. We show that SATB1 promotes Ig gene transcription in resting B cells, while in activated B cells, it acts as a repressor. In contrast to the T-cell lineage, where SATB1 is considered a major regulatory factor of the enhancer/super-enhancer network [53, 54], SATB1 deletion does not induce major disruptions in the B-cell transcriptome. Our invalidation model shows that only a reduced number of genes expressed during B differentiation are impacted by SATB1 deletion. In agreement with the repressive function of this factor in activated B cells, our study also shows that Ig genes are the predominant targets of SATB1 in activated B cells. Our 3C-HTGTS data indicate that SATB1 is not required to establish chromatin loops between  $E\mu$  and  $V_H$  regions in resting and activated B cells. A rational hypothesis could be that the effect of SATB1 on *IgH* transcription is mainly due to its function as a transcription factor rather than a genome organizer. Given the critical effect of SATB1 depletion on Ig gene transcription, it would certainly be interesting to evaluate SATB1 physiological binding to regulatory regions in Ig gene loci in resting and activated B cells.

By clarifying that SATB1 is an essential activator of Ig chain transcription and consequently BCR expression in resting B cells, our study extends recent work published by Yokota's team [35] and proposes a rational explanation for the survival function of SATB1 at this stage. These two studies also point out the importance of SATB1 expression levels, which could, as demonstrated in HSCs [22, 23], fine-tune some critical genes or regulatory

pathways in lymphocytes. Beyond the physiological decrease in *Satb1* expression between naive and activated B stages, it is possible that this gene undergoes splicing modifications in the B-cell lineage, as proposed in developing T cells by Spilianakis' group [70]. The stability of the SATB1 protein, according to its transcript variants, might also be critical for its function in mature B-cell subsets, as already proposed in T-lineage cells by Galande's group [19]. Interestingly, the dual function of SATB1 also observed in the B lineage, switching from activator to repressor of transcription in activated B cells, is reminiscent of the molecular switch observed upon phosphorylation and acetylation of this factor [16].

A surprising observation from our study is that SATB1 deletion increases somatic hypermutation of the *IgH* locus, extending into the  $S\mu$  region. It is very likely that the *Igk* locus is similarly influenced. A direct impact of SATB1 on key factors involved in SHM, such as AID and proteins involved in BER and MMR pathways, can be excluded since normal induction of their transcripts was observed in the absence of SATB1. In this respect, SATB1 could play a protective role against SHM in resting cells. This finding should be read in conjunction with our recent study of the MAR regions associated with the  $E\mu$  enhancer of the *IgH* locus [67], which is known to bind SATB1 [1]. The SATB1-induced increase in SHM could logically be correlated with the observed increase in primary transcription of the variable regions of Ig H and L chains. However, it is not excluded that SATB1 modifies the accessibility of these regions to AID or its cofactors. Since it is proposed that SATB1 stabilizes the unpaired DNA regions against unwinding [13], such an action of SATB1 in the *MARs-E $\mu$*  region could contribute to the protection of its surrounding regions against unwanted SHM. In line with the putative protective function of SATB1, it has been shown that this protein acts as an accessory factor of BER through its interaction with Oxo-guanine-glycosylase 1 (OGG1) [71], a DNA glycosylase that usually promotes error-free repair and is not involved in Ig gene SHM [72].

Further study of these modifications of SATB1 in B cells will be necessary to clarify the origin of its versatility. It has recently been proposed that SATB1 isoforms are subject to phase separation in T cells [70]. Consistent with its localization to PML nuclear bodies [73, 74], this mode of regulation deserves to be explored in the context of the B-cell nucleus. Moreover, since chromatin loop extrusion mechanisms are now proven to be critical for Ig gene regulation [75], it might be relevant to evaluate a potential implication of SATB1 in this process.

## DATA AVAILABILITY

Raw data from RNA-seq, SHM, Rep-Seq, and 3C-HTGTS have been deposited in the European Nucleotide Archive database under accession number PRJEB52320.

## REFERENCES

- Dickinson L, Tadashi J, Yoshinori K, Terumi K-S. A tissue-specific MARSAR DNA-binding protein with unusual binding site recognition. *Cell*. 1992;70:631–45.
- Alvarez JD, Yasui DH, Niida H, Joh T, Loh DY, Kohwi-Shigematsu T. The MAR-binding protein SATB1 orchestrates temporal and spatial expression of multiple genes during T-cell development. *Genes Dev*. 2000;14:521–35.
- Cai S, Han H-J, Kohwi-Shigematsu T. Tissue-specific nuclear architecture and gene expression regulated by SATB1. *Nat Genet*. 2003;34:42–51.
- Yasui D, Miyano M, Cai S, Varga-Weisz P, Kohwi-Shigematsu T. SATB1 targets chromatin remodelling to regulate genes over long distances. *Nature*. 2002;419:641–5.
- Cai S, Lee CC, Kohwi-Shigematsu T. SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nat Genet*. 2006;38:1278–88.
- Kohwi-Shigematsu T, Poterlowicz K, Ordinario E, Han H-J, Botchkarev VA, Kohwi Y. Genome organizing function of SATB1 in tumor progression. *Semin Cancer Biol*. 2013;23:72–9.
- Zelenka T, Spilianakis C. SATB1-mediated chromatin landscape in T cells. *Nucl Austin Tex*. 2020;11:117–31.

- Wang Z, Yang X, Chu X, Zhang J, Zhou H, Shen Y, et al. The structural basis for the oligomerization of the N-terminal domain of SATB1. *Nucleic Acids Res*. 2012;40:4193–202.
- Purbey PK, Singh S, Kumar PP, Mehta S, Ganesh KN, Mitra D, et al. PDZ domain-mediated dimerization and homeodomain-directed specificity are required for high-affinity DNA binding by SATB1. *Nucleic Acids Res*. 2008;36:2107–22.
- Yamasaki K, Akiba T, Yamasaki T, Harata K. Structural basis for recognition of the matrix attachment region of DNA by transcription factor SATB1. *Nucleic Acids Res*. 2007;35:5073–84.
- De Belle I, Cai S, Kohwi-Shigematsu T. The genomic sequences bound to special AT-rich sequence-binding protein 1 (SATB1) in vivo in Jurkat T cells are tightly associated with the nuclear matrix at the bases of the chromatin loops. *J Cell Biol*. 1998;141:335–48.
- Seo J, Lozano MM, Dudley JP. Nuclear matrix binding regulates SATB1-mediated transcriptional repression. *J Biol Chem*. 2005;280:24600–9.
- Ghosh RP, Shi Q, Yang L, Reddick MP, Nikitina T, Zhurkin VB, et al. *Satb1* integrates DNA binding site geometry and torsional stress to differentially target nucleosome-dense regions. *Nat Commun*. 2019;10. <https://doi.org/10.1038/s41467-019-11118-8>.
- Kumar PP, Purbey PK, Ravi DS, Mitra D, Galande S. Displacement of SATB1-Bound histone deacetylase 1 corepressor by the human immunodeficiency virus type 1 transactivator induces expression of interleukin-2 and its receptor in T cells. *Mol Cell Biol*. 2005;25:1620–33.
- Galande S, Purbey PK, Notani D, Kumar PP. The third dimension of gene regulation: organization of dynamic chromatin loopscape by SATB1. *Curr Opin Genet Dev*. 2007;17:408–14.
- Pavan Kumar P, Purbey PK, Sinha CK, Notani D, Limaye A, Jayani RS, et al. Phosphorylation of SATB1, a global gene regulator, acts as a molecular switch regulating its transcriptional activity in vivo. *Mol Cell*. 2006;22:231–43.
- Purbey PK, Singh S, Notani D, Kumar PP, Limaye AS, Galande S. Acetylation-dependent interaction of SATB1 and CtBP1 mediates transcriptional repression by SATB1. *Mol Cell Biol*. 2009;29:1321–37.
- Khare SP, Shetty A, Biradar R, Patta I, Chen ZJ, Sathe AV, et al. NF- $\kappa$ B signaling and IL-4 signaling regulate SATB1 expression via alternative promoter usage during Th2 differentiation. *Front Immunol*. 2019;10:667.
- Patta I, Madhok A, Khare S, Gottimukkala KP, Verma A, Giri S, et al. Dynamic regulation of chromatin organizer SATB1 via TCR-induced alternative promoter switch during T-cell development. *Nucleic Acids Res*. 2020;48:5873–90.
- Balamotis MA, Tamberg N, Woo YJ, Li J, Davy B, Kohwi-Shigematsu T, et al. *Satb1* ablation alters temporal expression of immediate early genes and reduces dendritic spine density during postnatal brain development. *Mol Cell Biol*. 2012;32:333–47.
- Will B, Vogler TO, Bartholdy B, Garrett-Bakelman F, Mayer J, Barreyro L, et al. *Satb1* regulates the self-renewal of hematopoietic stem cells by promoting quiescence and repressing differentiation commitment. *Nat Immunol*. 2013;14:437–45.
- Doi Y, Yokota T, Satoh Y, Okuzaki D, Tokunaga M, Ishibashi T, et al. Variable SATB1 levels regulate hematopoietic stem cell heterogeneity with distinct lineage fate. *Cell Rep*. 2018;23:3223–35.
- Satoh Y, Yokota T, Sudo T, Kondo M, Lai A, Kincade PW, et al. The *Satb1* protein directs hematopoietic stem cell differentiation toward lymphoid lineages. *Immunity*. 2013;38:1105–15.
- Wen J, Huang S, Rogers H, Dickinson LA, Kohwi-Shigematsu T, Noguchi CT. SATB1 family protein expressed during early erythroid differentiation modifies globin gene expression. *Blood*. 2005;105:3330–9.
- Hawkins SM, Kohwi-Shigematsu T, Skalnik DG. The matrix attachment region-binding protein SATB1 interacts with multiple elements within the *gp91<sup>Phox</sup>* promoter and is down-regulated during myeloid differentiation. *J Biol Chem*. 2001;276:44472–80.
- Fujii Y, Kumatori A, Nakamura M. SATB1 makes a complex with p300 and represses *gp91<sup>phox</sup>* promoter activity. *Microbiol Immunol*. 2003;47:803–11.
- Naik R, Galande S. SATB family chromatin organizers as master regulators of tumor progression. *Oncogene*. 2019;38:1989–2004.
- Papadogkonas G, Papamatheakis D-A, Spilianakis C. 3D genome organization as an epigenetic determinant of transcription regulation in T cells. *Front Immunol*. 2022;13:921375.
- Kohwi-Shigematsu T, Maass K, Bode J. A thymocyte factor SATB1 suppresses transcription of stably integrated matrix-attachment region-linked reporter genes. *Biochemistry*. 1997;36:12005–10.
- Burute M, Gottimukkala K, Galande S. Chromatin organizer SATB1 is an important determinant of T-cell differentiation. *Immunol Cell Biol*. 2012;90:852–9.
- Kakugawa K, Kojo S, Tanaka H, Seo W, Endo TA, Kitagawa Y, et al. Essential roles of SATB1 in specifying T lymphocyte subsets. *Cell Rep*. 2017;19:1176–88.
- Kitagawa Y, Ohkura N, Kidani Y, Vandenbon A, Hirota K, Kawakami R, et al. Guidance of regulatory T cell development by *Satb1*-dependent super-enhancer establishment. *Nat Immunol*. 2017;18:173–83.

33. Hao B, Naik AK, Watanabe A, Tanaka H, Chen L, Richards HW, et al. An antisilencer- and SATB1-dependent chromatin hub regulates Rag1 and Rag2 gene expression during thymocyte development. *J Exp Med*. 2015;212:809–24.
34. Feng D, Li Z, Qin L, Hao B. The role of chromatin organizer Satb1 in shaping TCR repertoire in adult thymus. *Genome*. 2021;64:821–32.
35. Ozawa T, Fujii K, Sudo T, Doi Y, Nakai R, Shingai Y, et al. Special AT-rich sequence-binding protein 1 supports survival and maturation of naive B cells stimulated by B cell receptors. *J Immunol*. 2022;208:1937–46.
36. Hobeika E, Thiemann S, Storch B, Jumaa H, Nielsen PJ, Pelanda R, et al. Testing gene function early in the B cell lineage in mb1-cre mice. *Proc Natl Acad Sci USA*. 2006;103:13789–94.
37. Martin OA, Garot A, Le Noir S, Aldigier J-C, Cogné M, Pinaud E, et al. Detecting rare AID-induced mutations in B-lineage oncogenes from high-throughput sequencing data using the detection of minor variants by error correction method. *J Immunol*. 2018;201:950–6.
38. Tinguely A, Chemin G, Péron S, Sirac C, Reynaud S, Cogné M, et al. Cross talk between immunoglobulin heavy-chain transcription and RNA surveillance during B cell development. *Mol Cell Biol*. 2012;32:107–17.
39. Pascal V, Dupont M, de Rouault P, Rizzo D, Rossille D, Jeannet R, et al. Demultiplexing Ig repertoires by parallel mRNA/DNA sequencing shows major differential alterations in severe COVID-19. *iScience*. 2023;26:106260.
40. Ouk C, Roland L, Gachard N, Poulain S, Oblet C, Rizzo D, et al. Continuous MYD88 activation is associated with expansion and then transformation of IgM differentiating plasma cells. *Front Immunol*. 2021;12:641692.
41. Turchaninova MA, Davydov A, Britanova OV, Shugay M, Bikos V, Egorov ES, et al. High-quality full-length immunoglobulin profiling with unique molecular barcoding. *Nat Protoc*. 2016;11:1599–616.
42. Javaugue V, Pascal V, Bender S, Nasraddine S, Dargelos M, Alizadeh M, et al. RNA-based immunoglobulin repertoire sequencing is a new tool for the management of monoclonal gammopathy of renal (kidney) significance. *Kidney Int*. 2022;101:331–7.
43. Boice M, Salloum D, Mourcin F, Sanghvi V, Amin R, Oricchio E, et al. Loss of the HVEM tumor suppressor in lymphoma and restoration by modified CAR-T cells. *Cell*. 2016;167:405–18.e13.
44. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29:15–21.
45. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30:923–30.
46. Varet H, Brillet-Guéguen L, Coppée J-Y, Dillies M-A. SARTools: a DESeq2- and EdgeR-Based R pipeline for comprehensive differential analysis of RNA-Seq data. *PLoS One*. 2016;11:e0157022.
47. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation*. 2021;2:100141.
48. Zhang X, Yoon HS, Chapdelaine-Williams AM, Kyritsis N, Alt FW. Physiological role of the 3'IgH CBEs super-anchor in antibody class switching. *Proc Natl Acad Sci USA*. 2021;118:e2024392118.
49. Hu J, Meyers RM, Dong J, Panchakshari RA, Alt FW, Frock RL. Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat Protoc*. 2016;11:853–71.
50. The Immunological Genome Project Consortium, Heng TSP, Painter MW, Elpek K, Lukacs-Kornek V, Mauermann N, et al. The immunological genome project: networks of gene expression in immune cells. *Nat Immunol*. 2008;9:1091–4.
51. Yoshida H, Lareau CA, Ramirez RN, Rose SA, Maier B, Wroblewska A, et al. The cis-regulatory atlas of the mouse immune system. *Cell*. 2019;176:897–912. e20.
52. Kohwi-Shigematsu T, Kohwi Y, Takahashi K, Richards HW, Ayers SD, Han H-J, et al. SATB1-mediated functional packaging of chromatin into loops. *Methods*. 2012;58:243–54.
53. Feng D, Chen Y, Dai R, Bian S, Xue W, Zhu Y, et al. Chromatin organizer SATB1 controls the cell identity of CD4<sup>+</sup> CD8<sup>+</sup> double-positive thymocytes by regulating the activity of super-enhancers. *Nat Commun*. 2022;13:5554.
54. Zelenka T, Klonizakis A, Tsoukatou D, Papamatheakis D-A, Franzenburg S, Tzerpos P, et al. The 3D enhancer network of the developing T cell genome is shaped by SATB1. *Nat Commun*. 2022;13:6954.
55. Dickinson LA, Dickinson CD, Kohwi-Shigematsu T. An atypical homeodomain in SATB1 promotes specific recognition of the key structural element in a matrix attachment region. *J Biol Chem*. 1997;272:11463–70.
56. Dickinson LA, Kohwi-Shigematsu T. Nucleolin is a matrix attachment region DNA-binding protein that specifically recognizes a region with high base-unpairing potential. *Mol Cell Biol*. 1995;15:456–65.
57. Marquet M, Garot A, Bender S, Carrión C, Rouaud P, Lecardeur S, et al. The E<sub>μ</sub> enhancer region influences H chain expression and B cell fate without impacting IgVH repertoire and immune response in vivo. *J Immunol Balt Md* 1950. 2014;193:1171–83.
58. Xu Y, Davidson L, Alt FW, Baltimore D. Deletion of the Ig kappa light chain intronic enhancer/matrix attachment region impairs but does not abolish V kappa J kappa rearrangement. *Immunity*. 1996;4:377–85.
59. Yi M, Wu P, Trevorrow KW, Claffin L, Garrard WT. Evidence that the Igkappa gene MAR regulates the probability of premature V-J joining and somatic hypermutation. *J Immunol Balt Md* 1950. 1999;162:6029–39.
60. Sakai E, Bottaro A, Davidson L, Sleckman BP, Alt FW. Recombination and transcription of the endogenous Ig heavy chain locus is effected by the Ig heavy chain intronic enhancer core region in the absence of the matrix attachment regions. *Proc Natl Acad Sci USA*. 1999;96:1526–31.
61. Garot A, Marquet M, Saintamand A, Bender S, Le Noir S, Rouaud P, et al. Sequential activation and distinct functions for distal and proximal modules within the IgH 3' regulatory region. *Proc Natl Acad Sci USA*. 2016;113:1618–23.
62. Inlay M, Alt FW, Baltimore D, Xu Y. Essential roles of the kappa light chain intronic enhancer and 3' enhancer in kappa rearrangement and demethylation. *Nat Immunol*. 2002;3:463–8.
63. Bruzeau C, Moreau J, Le Noir S, Pinaud E. Panorama of stepwise involvement of the IgH 3' regulatory region in murine B cells. *Adv Immunol*. 2021;149:95–114.
64. Yokota T, Kanakura Y. Role of tissue-specific AT-rich DNA sequence-binding proteins in lymphocyte differentiation. *Int J Hematol*. 2014;100:238–45.
65. Martin OA, Thomas M, Marquet M, Bruzeau C, Garot A, Brousse M, et al. The IgH E<sub>μ</sub>-MAR regions promote UNG-dependent error-prone repair to optimize somatic hypermutation. *Front Immunol*. 2023;14:1030813.
66. Cumano A, Rajewsky K. Clonal recruitment and somatic mutation in the generation of immunological memory to the hapten NP. *EMBO J*. 1986;5:2459–68.
67. Martin O, Thomas M, Marquet M, Garot A, Brousse M, Bender S et al. The IgH E<sub>μ</sub>-MAR regions promote UNG-dependent error-prone repair to optimize somatic hypermutation. *Immunology*. 2022. <https://doi.org/10.1101/2022.08.15.503996>.
68. Reth M, Hämmerling GJ, Rajewsky K. Analysis of the repertoire of anti-NP antibodies in C57BL/6 mice by cell fusion. I. Characterization of antibody families in the primary and hyperimmune response. *Eur J Immunol*. 1978;8:393–400.
69. Fukita Y, Jacobs H, Rajewsky K. Somatic hypermutation in the heavy chain locus correlates with transcription. *Immunity*. 1998;9:105–14.
70. Zelenka T, Tzerpos P, Panagopoulos G, Tsolis KC, Papamatheakis D-A, Papadakis VM et al. SATB1 undergoes isoform-specific phase transitions in T cells. *Front Cell Dev Biol*. 2022;11. <https://doi.org/10.3389/fcell.2023.1242481>.
71. Kaur S, Coulombe Y, Ramdzan ZM, Leduy L, Masson J-Y, Nepveu A. Special AT-rich sequence-binding Protein 1 (SATB1) functions as an accessory factor in base excision repair. *J Biol Chem*. 2016;291:22769–80.
72. Winter DB, Phung QH, Zeng X, Seeberg E, Barnes DE, Lindahl T, et al. Normal somatic hypermutation of Ig genes in the absence of 8-hydroxyguanine-DNA glycosylase. *J Immunol Balt Md* 1950. 2003;170:5558–62.
73. Kumar PP, Bischof O, Purbey PK, Notani D, Urlaub H, Dejean A, et al. Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. *Nat Cell Biol*. 2007;9:45–56.
74. Tan J-AT, Sun Y, Song J, Chen Y, Krontiris TG, Durrin LK. SUMO conjugation to the matrix attachment region-binding protein, special AT-rich sequence-binding protein-1 (SATB1), targets SATB1 to promyelocytic nuclear bodies where it undergoes caspase cleavage. *J Biol Chem*. 2008;283:18124–34.
75. Zhang Y, Zhang X, Dai H-Q, Hu H, Alt FW. The role of chromatin loop extrusion in antibody diversification. *Nat Rev Immunol*. 2022;22:550–66.

## ACKNOWLEDGEMENTS

The authors are grateful to the BISCEM unit (Univ. Limoges, UAR 2015 CNRS, US 42 Inserm, CHU Limoges) for technical support regarding DNA–RNA sequencing, cytometry experiments and the animal core facility. We are grateful to Emeline Lhuillier and the Genotoul Plateau GeT-Santé facility (<https://get.genotoul.fr>) for technical assistance with RNA sequencing, to Mehdi Alizadeh from Etablissement Français du Sang (Rennes; France) for assistance with Repertoire sequencing and to Christelle Oblet for technical help with Western blotting. This work benefitted from data assembled by the ImmGen consortium [50]. MT, CB and OM were supported by PhD fellowships from the French Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation and the Fonds Européen de Développement Régional (FEDER). This work was supported by La Ligue Contre le Cancer (comités 87, 23 to EP and SLN); the Fondation ARC pour la recherche sur le cancer (PJA 20181207918 to EP and PhD continuation fellowship to MT and CB), Institut CARNOT CALYM, INCA-Cancéropôle GSO Emergence (to EP). We are grateful to Drs. Charalampos Spilianakis, Jeanne Moreau and Amélie Bonaud for critical reading of the manuscript, helpful comments and edits.

## AUTHOR CONTRIBUTIONS

MT, CB, OM, SB, and CC performed the experiments. MT analyzed the data. EP and SLN conceived and supervised the study. MT and OM developed the experimental model. JP performed the bioinformatic analysis. MT, CB, EP, and SLN wrote the manuscript.

**COMPETING INTERESTS**

The authors declare no competing interests.

**ADDITIONAL INFORMATION**

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41423-023-01069-y>.

**Correspondence** and requests for materials should be addressed to Sandrine Le Noir or Eric Pinaud.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.





## Rôle des éléments *cis*-régulateurs du locus des chaînes lourdes d'immunoglobuline dans l'organisation du noyau et l'intégrité du génome des cellules B

---

Les stades tardifs de la différenciation lymphocytaire B sont caractérisés par des remaniements géniques, telles que l'hypermutation somatique et la recombinaison de classe, qui *in fine*, permettent la production d'immunoglobulines hautement spécifique. Ces processus, intervenant aux loci codant les immunoglobulines, introduisent des mutations ainsi que des cassures double brin de l'ADN. Tout dysfonctionnement de ces remaniements constitue un danger majeur pour l'intégrité du génome de la cellule B ; leur régulation est donc essentielle. La surveillance et le contrôle de ces événements sont notamment médiés par les régions *cis*-régulatrices du locus *IgH* : les régions *E $\mu$ MARs* et *3'RR*. Le rôle de ces régions à l'échelle transcriptionnelle a été largement caractérisé, grâce à différents modèles murins présentant des délétions totales ou partielles de celles-ci. La régulation spatio-temporelle des remaniements géniques a été également décrite au niveau nucléaire, avec des déplacements dynamiques des loci d'immunoglobuline, ainsi qu'au niveau supranucléosome par la formation de boucles de chromatine. Des données provenant du laboratoire et de la littérature suggèrent l'implication des régions *cis*-régulatrices du locus *IgH* dans ces mécanismes de régulation supplémentaires. Nous avons démontré que les core enhancers de la *3'RR* sont nécessaires et suffisants pour correctement adresser le locus *IgH* dans le noyau et pour permettre la formation des boucles de chromatine et ainsi induire une recombinaison de CSR optimale et fonctionnelle. Toutefois, nos données indiquent que ces core enhancers ne sont pas suffisants pour maintenir l'intégrité génomique de la cellule B comme en témoigne l'analyse des mutations sur des oncogènes tels que *Bcl6*. Ces travaux ouvrent de nouvelles perspectives pour la compréhension des mécanismes régissant le développement B tardif ainsi que le maintien de l'intégrité génomique *via* l'organisation du noyau.

---

Mots-clés : Cellules B, Organisation du Noyau, éléments *cis*-régulateurs, Locus *IgH*, Intégrité du génome

### Role of immunoglobulin heavy chain locus *cis*-regulatory elements in B cell nuclear organization and genome integrity

---

The late stages of B lymphocyte differentiation are characterized by gene rearrangements, such as somatic hypermutation and class recombination, which ultimately lead to the production of highly specific immunoglobulins. These processes, occurring at immunoglobulin-coding loci, introduce mutations and DNA double-strand breaks. Any malfunctioning of these rearrangements constitutes a major threat to the integrity of the B cell genome, and their regulation is therefore essential. Monitoring and control of these events are mediated by the *cis*-regulatory regions of the *IgH* locus: the *E $\mu$ MARs* and *3'RR* regions. The role of these regions at transcriptional level has been extensively characterized, using various mouse models with partial or total deletions of these regions. The spatio-temporal regulation of gene rearrangements has also been described at the nuclear level, with dynamic shifts of immunoglobulin loci, as well as at the supranucleosomal level through the formation of chromatin loops. Data from the laboratory and the literature suggest the involvement of *cis*-regulatory regions of the *IgH* locus in these additional regulatory mechanisms. We have demonstrated that core enhancers of the *3'RR* are necessary and sufficient to correctly address the *IgH* locus in the nucleus and to allow chromatin loop formation and thus induce optimal and functional CSR recombination. However, our data indicate that these core enhancers are not sufficient to maintain the genomic integrity of the B cell, as shown by the analysis of mutations in oncogenes such as *Bcl6*. This work opens up new perspectives for understanding the mechanisms governing late B development and the maintenance of genomic integrity via nuclear organization.

---

Keywords: B cells, Nuclear Organization, *Cis*-regulatory elements, *IgH* locus, Genome Integrity

