



HAL
open science

Régulation transcriptionnelle et épigénétique de la différenciation B normale et tumorale : rôle des enzymes Tet et du facteur de transcription SPI1

Marine Armand

► **To cite this version:**

Marine Armand. Régulation transcriptionnelle et épigénétique de la différenciation B normale et tumorale : rôle des enzymes Tet et du facteur de transcription SPI1. Biologie cellulaire. Université Paris-Saclay, 2020. Français. NNT : 2020UPASL035 . tel-03917487

HAL Id: tel-03917487

<https://theses.hal.science/tel-03917487>

Submitted on 2 Jan 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.

Régulation transcriptionnelle et épigénétique de la différenciation B normale et tumorale : rôle des enzymes TET et du facteur de transcription SPI1.

Thèse de doctorat de l'université Paris-Saclay

École doctorale n°582 : cancérologie : biologie-médecine-santé (CBMS)
Spécialité de doctorat : Science de la vie et de la santé
Unité de Recherche : INSERM UMR 1170, Dynamique moléculaire de la transformation hématopoïétique, Université Paris-Saclay, Inserm, Institut Gustave Roussy, Dynamique moléculaire de la transformation hématopoïétique, 94805, Villejuif, France
Référent : Faculté de médecine

Thèse présentée et soutenue à Villejuif le 29/10/2020, par

Marine ARMAND

Composition du Jury

Patricia KANNOUCHE DR, INSERM UMR 9019, Université Paris-Saclay	Présidente
Jean-Christophe BORIES DR, INSERM U1126, Université de Paris	Rapporteur
Jean FEUILLARD PU-PH, CHU de Limoges	Rapporteur
Fanny BARAN-MARSZAK PU-PH, CHU Avicenne, Bobigny	Examinatrice
Céline DELALOY CR1, INSERM U1236, Université de Rennes	Examinatrice
Olivier BERNARD DR, INSERM UMR 1170, Université Paris-Saclay	Directeur de thèse
Said AOUFOUCHI DR, INSERM UMR 9019, Université Paris-Saclay	Co-Directeur de thèse

Remerciements

Je tiens avant tout à remercier mes deux directeurs de thèse :

Olivier Bernard, pour m'avoir accueilli dans son équipe et intégrée à des projets passionnants. Merci d'avoir toujours pris le temps de répondre à mes doutes et mes questions, et d'avoir essayé d'affûter mon raisonnement scientifique (à l'impossible nul n'est tenu). Plus sérieusement, j'ai beaucoup appris en 3 ans, bien plus que je ne me serais permis d'espérer.

Said Aoufouchi, pour sa présence bienveillante, son oreille attentive et ses conseils. Merci de m'avoir accompagnée tout au long de cette aventure.

Je remercie le Dr Jean-Christophe Bories et le Pr Jean Feuillard d'avoir accepté d'être les rapporteurs de mon travail de thèse. Je remercie le Dr Patricia Kannouche, Dr Céline Delaloy et Pr Fanny Baran-Marszak d'avoir accepté d'examiner ce travail. Merci à l'ensemble de mon jury pour leur disponibilité et pour le temps qu'ils auront accordé à la lecture de mes travaux.

Merci à mon équipe, sans qui rien n'aurait pu être fait.

Véro : pourrais-je un jour honorer la dette que j'ai envers toi ? Bien sûr je parle du temps passé à aider, expliquer, consoler, motiver, tout à l'air facile et faisable avec toi. Je parle également de ces trajets Paris-Villejuif, peu importe le véhicule (petite pref pour la Smart), la compagnie est là et c'est toujours un bonheur. Je pense qu'à l'heure où tu liras ces lignes je n'aurai toujours pas réussi à te faire prendre le métro.

Merci à Hussein, pour son aide et sa patience, pour m'avoir expliqué et chaperonné pour une grande partie des manips qui ont permis de présenter ce travail.

Merci à Camille, Seb, Amaury, Philippe, Laure, Michaela, Mael, Paule. Si la motivation était là jusqu'au bout, c'est aussi grâce à vous. Je vous souhaite le meilleur pour vos projets perso et pro, ici ou ailleurs.

Je souhaite beaucoup de courage et de succès à Amina et Michael, qui auront déjà eu le courage de me supporter pendant une année, donc le reste devrait bien se passer. Amina, si je me crée un compte Insta je suivrai de près ta carrière d'influenceuse. Michael, essaye de ne pas combler mon absence dans le bureau par des quantités industrielles de tablettes de chocolat bio/équitable/vegan.

Merci à Christel Guillouf d'avoir pris le temps de m'aider, je suis contente que l'on puisse parler de sciences mais aussi de beaucoup d'autre chose ensemble.

Merci à Damien Roos-Weil pour son aide, c'est un plaisir de collaborer avec toi.

Merci à Julie Rivière pour son aide et à Thomas Mercher pour ses conseils.

Un grand merci également à Cécile Lopez (bouge pas, j'arrive outre-manche) pour, par ordre d'importance : la danse des légumes, sa bonne humeur contagieuse, son expertise scientifique.

Merci à Cyril, Yann et Philippe, de faire de la plateforme de cytométrie un endroit accueillant et convivial, merci d'avoir encaissé mes passages qui se suivent mais ne se ressemblent pas. Merci à Christophe pour son aide avec les souris, à quand la petite démo de Saxo ?

Merci à Yanis, Margaux et Chloé d'avoir été là. D'exister. Grâce à vous, la thèse sera associée à beaucoup d'excellents souvenirs. Yanis et Chloé, merci pour le petit coup de vieux, régulièrement, c'est quoi le bail avec toutes ces expressions que je ne connais plus ?

Margaux, merci d'être plus vieille que moi. Hâte que l'on se retrouve toutes les 4 pour reprendre nos discussions de haute portée scientifique.

John-Edward, merci d'être un compagnon de voyage exemplaire, que ce soit ici avec ton talent RATP qui m'a fait gagner beaucoup de temps de trajet, ou là-bas dans un ruin bar ou sur une grande roue. Merci pour ta veille scientifique permanente, tu feras un merveilleux doctorant.

Une pensée pour l'équipe de la Pitié-Salpêtrière : Frédéric Davi, qui suit avec bienveillance mon parcours, Magali Le Garff-Tavernier, Élise Chapiro, Florence Nguyen-Khac et pour les techniciens de Biologie Moléculaire, Marie-Thérèse, Patricia, Hervé Senior et Hervé Junior. J'espère avoir la chance de retravailler avec vous bientôt.

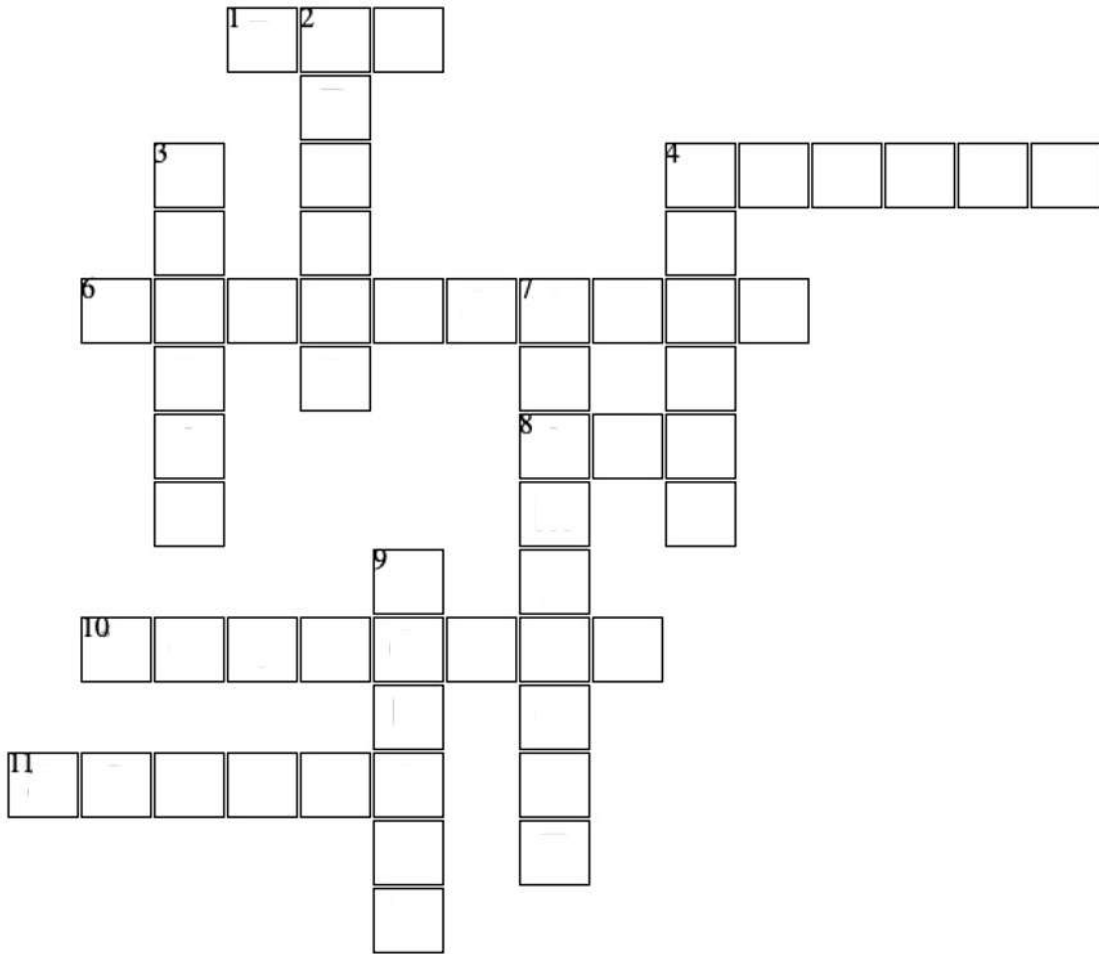
Merci à mes parents, de m'avoir toujours encouragée et soutenue, d'assister à une deuxième thèse encore moins compréhensible que la 1^{ère}, je ne sais pas si votre fierté est légitime mais elle a souvent été un moteur pour moi. Merci à mes trois grandes sœurs, pour qui je n'ai pas été très présente ces derniers temps mais qui comptent beaucoup pour moi.

Merci à Aurélie, qui a vécu les toutes premières lignes de ce manuscrit, pendant le confinement. Merci d'avoir rendu cette période plus douce. Merci également pour tes imitations de générique de séries, ton talent pour les arbres généalogiques, ta passion pour le voisinage, tout cela est très divertissant.

Merci à mes amies qui m'ont soutenue, entourée, reboostée : Marie (merci pour les bols d'air à St Jean), Chloe (très bon poisson-pilote), Lea, AF, Marine, Agnès, Anne-Laure, Emma, Tiffany, Anna, Marion, Julie...

Merci à tous ceux qui ont relu et corrigé ce manuscrit, ainsi qu'à ceux qui auront pu venir m'écouter pour cette soutenance, masqués, distancés. Mon plus grand regret sera de ne pouvoir vous faire une bise pour vous témoigner toute ma gratitude...

Si, d'aventure, vous aviez envisagé l'espace d'un instant de lire cette thèse, je vous propose une alternative ludique sur la page suivante (sur une idée originale d'Antonin et Nathan, merci à eux). Je ne peux néanmoins pas m'engager à fournir les solutions dans le prochain numéro... quoique...



Horizontal

- 1 réalisée un nombre incalculable de fois en 3 ans
- 4 on l'a dans l'os
- 6 cellule au coeur de ce travail
- 8 bases de la vie
- 10 expérience éprouvante mais enrichissante
- 11 pardon à celles qui ont croisé mon chemin

Vertical

- 2 technique d'édition du génome
- 3 couleur ocean pour la viabilité
- 4 n'est plus sauvage
- 7 coulera à flots si tout va bien
- 9 haut lieu où l'élite scientifique prospère

Table des matières

INTRODUCTION.....	9
1 Ontogénie des lymphocytes B.....	10
1.1 Phase indépendante de l'antigène.....	10
1.1.1 Etapes de différenciation.....	10
1.1.2 Rôle des cytokines et des récepteurs de cytokine.....	13
1.1.3 Rôle des facteurs de transcription.....	13
1.2 Phase dépendante de l'antigène.....	13
1.2.1 Départ de la MO et colonisation des organes lymphoïdes secondaires.....	13
1.2.2 Réaction du centre germinatif.....	14
1.2.3 Rôle des facteurs de transcription.....	18
1.2.4 Régulation épigénétique de la différenciation lymphoïde B.....	23
1.2.5 Rôle des signaux extrinsèques.....	26
2 Réaction du CG et lymphomagénèse.....	26
2.1 Remaniements génétiques à potentiel oncogénique.....	26
2.2 Stimulation antigénique chronique.....	27
2.3 Micro-environnement et signalisation.....	29
2.3.1 Signalisation du BCR.....	30
2.3.2 Signalisation CD40 et CD40L : activation des LB dépendante des LT.....	32
2.3.3 Signalisation BAFF et APRIL.....	32
2.3.4 Signalisation des TLR.....	33
2.3.5 Voie NF-kB.....	33
2.3.6 Coopération entre les différentes voies de signalisation des LB.....	34
2.4 Modélisation expérimentale de l'activation des LB <i>in vitro</i>.....	35
3 Protéines TET.....	36
3.1 Généralités.....	36
3.2 Expression et fonction.....	38
3.3 Régulation de l'expression et de l'activité.....	39
3.3.1 Contrôle transcriptionnel et post-transcriptionnel.....	39
3.3.2 Protéolyse par les caspases et les calpaïnes.....	39
3.3.3 Interactions avec d'autres facteurs.....	40
3.4 Implication dans les mécanismes de transformation tumorale.....	41
3.5 Perte de fonction de TET2 et TET3.....	42
3.5.1 Dans l'hématopoïèse globale.....	42
3.5.2 Dans la lignée lymphoïde B.....	44
4 Protéines ETS.....	47
4.1 Généralités.....	47
4.2 Liaison à l'ADN.....	48
4.3 Interactions protéiques.....	50
4.4 Sites génomiques de liaisons à l'ADN.....	51
4.5 Implication dans les mécanismes de transformation tumorale.....	51
4.6 SPI1/PU.1.....	52
4.6.1 Structure et expression.....	52

4.6.2	Fonctions.....	54
4.6.3	Partenaires protéiques de SPI1	57
4.6.4	Rôle dans la tumorigénèse.....	59
5	<i>Lymphomes non Hodgkiniens B</i>	60
5.1	Généralités	60
5.2	DLBCL	61
5.2.1	Épidémiologie.....	61
5.2.2	Evaluation morphologique et immunophénotypique.....	61
5.2.3	Evaluation clinique et thérapeutique.....	62
5.2.4	Classification des DLBCL.....	62
5.2.5	Principales anomalies cytogénétiques et moléculaires.....	64
5.3	Maladie de Waldenstrom	66
5.3.1	Épidémiologie.....	66
5.3.2	Evaluation morphologique et immunophénotypique.....	67
5.3.3	Evaluation clinique et thérapeutique.....	67
5.3.4	Principales anomalies cytogénétiques et moléculaires.....	68
5.3.5	Physiopathologie / Origine cellulaire	72
	OBJECTIFS DU TRAVAIL	74
	PARTIE 1 : Enzymes TET	76
	1. Résumé de l'article	76
	2. Résultats supplémentaires	106
	2.1 Contexte et objectifs.....	106
	2.2. Matériel et Méthodes supplémentaires	107
	2.3 Résultats	108
	2.4 Discussion	115
	PARTIE 2 : Facteur de transcription SPI1	122
	1. Résumé de l'article	122
	2. Résultats supplémentaires	140
	2.1 Contexte/question posée	150
	2.2 Matériel et méthodes supplémentaires	150
	2.3 Résultats	152
	2.4 Discussion	165
	CONCLUSION et PERSPECTIVES	172
	REFERENCES	174

Liste des abréviations

5mC : 5-méthylcytosine

5hmC : 5-hydroxyméthylcytosine

A

ABC : Activated B-Cell

Ac : Anticorps

Ag : Antigène

AID : Activation-Induced cytidine Deaminase

AITL : Angioimmunoblastic T-cell Lymphoma

ASC : Antibody-Secreting Cell

B

BCR : B-Cell Receptor

BGC : B-cell of the Germinal Center

BTK : Bruton Tyrosine Kinase

C

CDF : Cellule Dendritique Folliculaire

CG : Centre Germinatif

CHIP : Clonal Hematopoiesis of
Indeterminate Potential

CRISPR : Clustered Regularly Interspaced
Short Palindromic Repeats

CSE : Cellule Souche Embryonnaire

CSH : Cellule Souche Hématopoïétique

CSR : Class Switch Recombination

D

DKO : Double Knock-Out

DLBCL : Diffuse Large B-Cell Lymphoma

DZ : Dark Zone

E

EICE : Ets-Interferon Consensus Elements

F

FT : Facteur de Transcription

G

GFP : Green Fluorescent Protein

I

Ig : Immunoglobuline

IL : Interleukine

IP : Immunoprécipitation

L

LAL : Leucémie Aigue Lymphoblastique

LAM : Leucémie Aigue Myéloïde

LB : Lymphocyte B

LH : Lymphome de Hodgkin

LLC : Leucémie Lymphoïde Chronique

LMMC : Leucémie Myélomonocytaire
Chronique

LNH : Lymphome Non Hodgkinien

LPL : Lymphome Lymphoplasmocytaire

LPS : Lipopolysaccharide

LT : Lymphocyte T

LZ : Light Zone

M

MALT : Mucosa-Associated Lymphoid Tissu
MO : Moelle Osseuse
MSCV : Murine Stem Cell Virus
MW : Maladie de Waldenstrom

P

PAM : Protospacer Adjacent Motif
PB : Plasmablaste
PC : Plasmocyte
PM : Point Mutation

R

RAG : Recombination Activating Gene

S

SHM : Somatic Hypermutation
SMD : Syndrome Myélodysplasique
SMP : Syndrome Myéloprolifératif
SRBC : Sheep Red Blood Cells

T

TFH : T Follicular Helper
TNF : Tumor Necrosis Factor
TSS : Transcription Start Site

U

URE : Upstream Regulatory Element
UTR : Untranslated Region

W

WES : Whole Exome Sequencing
WGS : Whole Genome Sequencing
WT : Wild Type

INTRODUCTION

L'hématopoïèse est le processus biologique qui génère en continu les cellules sanguines et les plaquettes. Chez l'Homme, elle débute au stade embryonnaire dans le sac vitellin, puis dans le foie et la rate avant de gagner progressivement la moelle osseuse (MO). Chez l'adulte, la MO devient le site principal de l'hématopoïèse, donnant naissance aux cellules des lignées myéloïdes, lymphoïdes, mégacaryocytaires et érythroïdes. Le modèle communément accepté est hiérarchique, avec au sommet de la pyramide une population de cellules appelées cellules souches hématopoïétiques (CSH), caractérisées par des capacités d'auto-renouvellement et de multipotence (**Figure 1**). L'engagement vers une lignée plutôt qu'une autre est dicté par l'expression coordonnée de facteurs de transcription (FT) spécifiques (1-3). Ce processus est également contrôlé par des signaux extracellulaires provenant de l'environnement, comprenant des molécules de la matrice extracellulaire, des molécules d'adhérence, des facteurs de croissance et des cytokines capables de moduler des voies de signalisation intracellulaire (4).

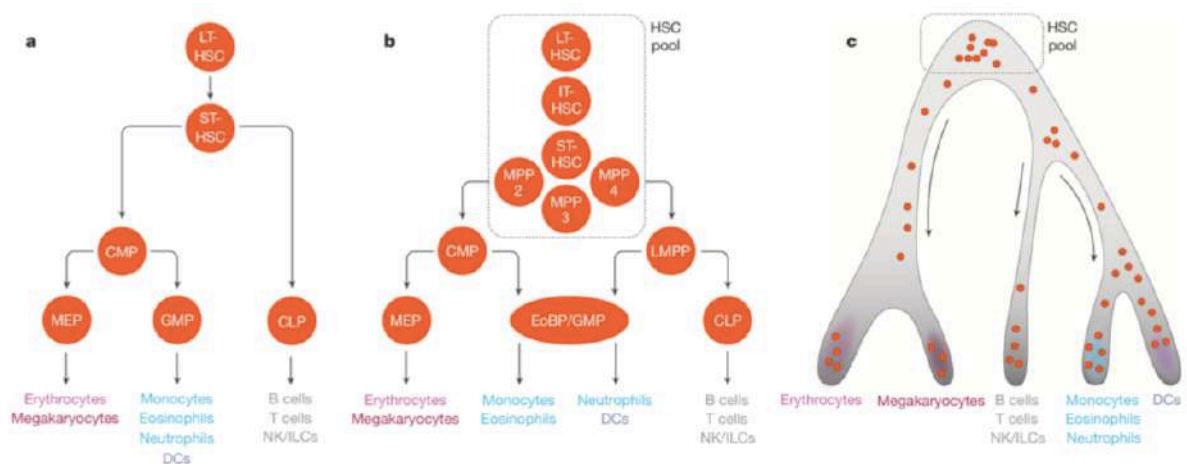


Figure 1 : Évolution des modèles hiérarchiques d'hématopoïèse (2).

a) Les CSH étaient à l'origine considérées comme une population homogène, se divisant ensuite en branches myéloïdes et lymphoïdes. b) Par la suite, le pool de CSH a été reconnu comme étant plus hétérogène à la fois en termes d'auto-renouvellement (axe vertical) et de propriétés de différenciation (axe horizontal), les branches myéloïdes et lymphoïdes restent associées plus bas dans la hiérarchie via le progéniteur multipotentiel engagé vers la voie lymphoïde (LMPP). c) À partir de 2016, les données transcriptomiques unicellulaires indiquent un continuum de différenciation. Chaque cellule (point rouge) se localise le long d'une trajectoire de différenciation.

Ce travail se concentrera sur une lignée en particulier : la lignée lymphoïde B.

1 Ontogénie des lymphocytes B

En tant que seule source d'anticorps (Ac) dans l'organisme, le développement de lymphocytes B (LB) fonctionnels est essentiel pour assurer un système immunitaire adaptatif. L'immunité adaptative (ou acquise), par opposition à l'immunité innée, est une immunité spécifique car la réaction immunitaire est basée sur la reconnaissance d'un antigène (Ag) par les récepteurs de surface de deux types de cellules immunitaires : les LB et les lymphocytes T (LT). Les LB participent à l'immunité à médiation humorale en sécrétant des Ac tandis que les LT participent à l'immunité à médiation cellulaire.

L'immunité à médiation par les Ac repose sur la capacité des LB à répondre à de multiples stimuli environnementaux. Après exposition à un Ag, les LB rentrent dans le cycle cellulaire et subissent de multiples divisions, ce qui leur permet de se différencier en plasmocytes à courte durée de vie, constituant une première ligne de réponse rapide, mais de faible affinité. En parallèle, en réponse à l'Ag et avec l'aide des LT, les LB activés peuvent entrer dans des structures transitoires appelées centres germinatifs (CG), où ils subissent une amplification clonale et une différenciation en plasmocytes qui sécrètent des Ac de haute affinité. Les CG produisent également des cellules B mémoires qui peuvent se différencier rapidement en plasmocytes lors d'une seconde exposition à l'Ag.

L'ontogénèse B chez l'Homme peut être divisée en deux grandes parties : une phase indépendante de l'antigène, qui se déroule dans la MO et la phase dépendante de l'Ag, qui se déroule dans les organes lymphoïdes secondaires. La première phase va permettre de générer des LB dits « naïfs », possédant une immunoglobuline (Ig) de surface fonctionnelle capable de reconnaître un futur Ag. La seconde phase va consister à activer et sélectionner les LB qui auront su le mieux s'adapter à l'Ag rencontré, pour aboutir à la génération de LB mémoires et de plasmocytes sécréteurs d'Ac spécifiques et de haute affinité.

La différenciation lymphoïde B est un processus finement régulé, qui aboutit en situation physiologique à une grande efficacité de réponse contre l'immense diversité des Ag existants. Le revers de la médaille est que les LB sont particulièrement sujets aux événements transformants lorsque la machinerie qu'ils utilisent pour la diversification des Ac est détournée de ses fonctions premières.

1.1 Phase indépendante de l'antigène

1.1.1 Etapes de différenciation

Les étapes de différenciation qui conduisent de la CSH au LB immature se déroulent dans la MO, en l'absence de stimulation antigénique. Les principaux stades du développement lymphocytaire B précoce sont caractérisés par l'acquisition de l'expression de marqueurs

de surface spécifiques et par une succession de réarrangements associant les différents segments des gènes codant pour les chaînes lourdes (IGH ; H, heavy) et légères (IGL ; L, light) d'Ig (**Figure 2**). Ces étapes s'articulent autour de la réorganisation fonctionnelle des gènes du récepteur à l'Ag des LB (B-cell receptor, BCR) et sont régulées par des étapes de sélection positive et négative (5).

Les précurseurs lymphoïdes communs (CLP), issus des CSH, possèdent la capacité de reconstituer les lignées lymphoïdes T, B et Natural Killer (NK). Les précurseurs B les plus immatures constituent une sous-population de cellules appelées pré-pro-B qui ne sont pas totalement engagées dans la voie B et n'ont pas encore réarrangé les gènes des Ig. Les cellules pré-pro-B expriment très faiblement les enzymes RAG (Recombination Activating Gene). En revanche, l'expression du gène codant pour Iga (*CD79A*) est détectée dès ce stade sous forme de protéines CD79a intracytoplasmiques.

C'est au stade pro-B que les réarrangements des gènes d'Ig commencent à se mettre en place. Les réarrangements débutent au locus de la chaîne lourde d'Ig (IGH) par jonction d'un segment DH avec un segment JH. Ces événements caractérisent le stade pro-B précoce au cours duquel apparaît le marqueur CD19. Ces premiers réarrangements sont suivis dans les cellules pro-B tardives par l'assemblage d'un segment VH avec les segments DJH réarrangés (**Figure 2**). Seuls les segments VHDJH en phase de lecture correcte et sans codon-stop codent pour une région variable fonctionnelle, et permettent la synthèse d'une chaîne lourde μ intracytoplasmique.

Le stade suivant, pré-B, est marqué par l'expression de la chaîne lourde μ à la surface des cellules, en association avec une pseudo-chaîne légère formée de la liaison non covalente des protéines $\lambda 5$ et Vpré-B. Ce complexe forme le pré-BCR qui permet à la cellule de poursuivre sa différenciation et d'entrer dans une phase d'expansion clonale. Le pré-BCR joue aussi un rôle critique dans l'exclusion allélique en induisant une diminution transitoire de l'expression des gènes *RAG* ce qui arrête la recombinaison des gènes de chaînes lourdes sur l'autre copie.

Suite à cette expansion clonale, les gènes *RAG* sont réexprimés pour réaliser les réarrangements VJ des gènes des chaînes légères, le lymphocyte pré-B perd alors l'expression du CD34 et exprime le CD20.

Le stade B immature est caractérisé par la production d'une chaîne légère classique qui remplace la pseudo-chaîne légère et donne naissance à une IgM de surface complète. Celle-ci est formée par l'association de deux chaînes lourdes identiques et de deux chaînes légères (kappa ou lambda) identiques, associées de manière non covalente aux CD79A/CD79B et confère à la cellule sa spécificité de reconnaissance de l'antigène.

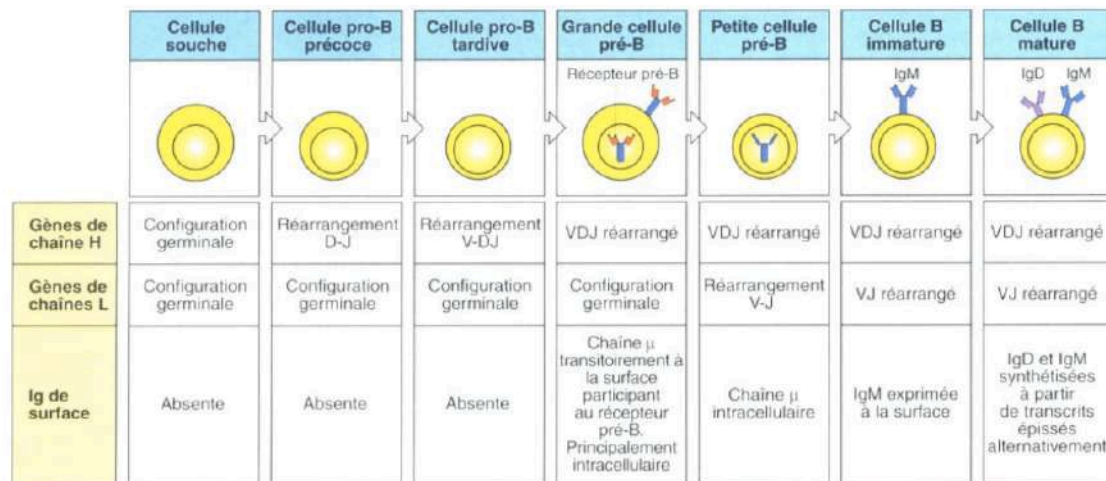


Figure 2 : Différenciation des lymphocytes B à partir de la CSH.

Les différents stades de développement sont marqués par le réarrangement et l'expression des gènes des immunoglobulines (Immunobiology, 6^{ème} édition, Garland Science 2005).

Toute cellule incapable d'assembler son pré-BCR ou d'associer et de transporter à la membrane une chaîne lourde μ avec une chaîne légère sera éliminée. A cette sélection positive, s'ajoute un processus de sélection négative selon lequel les cellules immatures exprimant des Ig membranaires spécifiques pour les antigènes du soi (autoréactifs – induisant un signal fort) seront également éliminées (de 55 à 75% des cellules) (6). Des réarrangements secondaires produisant de nouvelles chaînes légères (receptor editing) pourront aussi à ce stade modifier la spécificité du BCR afin qu'il ne soit plus autoréactif (7,8). Enfin, une cellule immature conservant son autoréactivité malgré l'editing peut encore entrer en anergie (diminution de la prolifération, désensibilisation et sous-expression du BCR (9,10)) ou être éliminée par apoptose (11,12).

Ces cellules B immatures vont alors produire un long transcrite d'ARN pré-messager couvrant les régions constantes des chaînes μ et δ . Un épissage de ce pré-ARNm associe la région variable VDJ recombinaison aux domaines constants de l'un ou l'autre isotype. Ces cellules B immatures naïves coexpriment ainsi deux types d'Ig avec la même spécificité : l'IgM et l'IgD. Ces cellules quittent ensuite la MO pour se rendre dans les organes lymphoïdes secondaires où elles pourront subir les dernières étapes de maturation.

1.1.2 Rôle des cytokines et des récepteurs de cytokine

Les premières étapes du développement sont strictement dépendantes du micro-environnement particulier apporté par les cellules stromales de la MO. Ces cellules stromales régulent la croissance, la maturation et la survie des précurseurs par l'intermédiaire de facteurs solubles (IL7, Stem Cell Factor ou SCF, SDF-1 et son récepteur CXCR4) et de contacts directs avec les cellules en développement (13).

1.1.3 Rôle des facteurs de transcription

Au cours de la différenciation lymphocytaire, la fixation de FT sur différentes séquences régulatrices de gènes cibles est impliquée dans la quiescence, la survie des progéniteurs B, ainsi que dans l'engagement des cellules dans une lignée spécifique. Le FT Ikaros (codé par le gène *IKZF1*) induit l'expression des gènes *RAG* et module l'accessibilité de la chromatine nécessaire à la recombinaison V(D)J (réarrangement des gènes des Ig) (14). *IKZF1* et *SPI1/PU.1* vont conditionner la différenciation des CSH en CLP, en régulant l'expression du récepteur à l'IL7 (*IL7-R α*), essentiel à l'initiation de la lymphopoïèse précoce (15,16). L'expression des facteurs EBF (Early B-cell factor) et E2A permet la différenciation du CLP en cellule pré-pro-B (17,18). Cette cellule perd alors toute capacité de différenciation en cellules T ou NK. Les FT *PAX5* (paired box protein 5) qui induit l'expression des gènes spécifiques des cellules B (*CD19* par exemple) et fige l'engagement vers la différenciation lymphoïde B et *STAT5* (signal transducer and activator of transcription 5) sont nécessaires à la maturation en cellule pro-B (19,20). La première recombinaison des gènes des chaînes lourdes d'Ig se produit à ce stade, générant ainsi le pré-BCR qui active deux axes : d'un côté, en stimulant la voie Ras/ERK, il potentialise l'expression de E2A, et de l'autre côté, il induit l'expression du FT *IRF4*. *IRF4* induit une acétylation locale des histones H3 et H4 sur les sites proches des gènes des Ig, favorisant ainsi les réarrangements ultérieurs qui aboutiront au BCR mature (21). *SPI1* régule positivement l'expression de *PAX5* et les deux facteurs restent exprimés pendant la plupart des étapes restantes de la différenciation des cellules B (22,23).

1.2 Phase dépendante de l'antigène

1.2.1 Départ de la MO et colonisation des organes lymphoïdes secondaires

Parmi les plusieurs millions de LB qui se développent chaque jour dans la MO, seuls 10 % environ sortent en périphérie (24) et passent par un stade intermédiaire, le stade B transitionnel, dernière étape avant la rencontre avec des antigènes exogènes dans les organes lymphoïdes secondaires (rate, ganglions).

Il existe trois sous-populations principales de cellules B chez la souris et l'homme : Les cellules B-1, les cellules B folliculaires (également appelées B-2), essentiellement impliquées dans les réponses humorales dépendantes des LT et les cellules de la zone marginale B (MZB), impliquées dans les réponses humorales indépendantes des LT. Les cellules B-1, divisées en B-1a (CD5+) et B-1b (CD5-), sont principalement présentes dans le péritoine où elles sont maintenues par auto-renouvellement. On pense que les cellules B-1a et une partie des cellules B-1b sont d'origine fœtale (25). En revanche, les cellules B folliculaires résident dans les follicules des organes lymphoïdes secondaires et sont générées à partir de précurseurs formés dans la MO (26). Enfin, les cellules MZB sont localisées dans les zones marginales des follicules spléniques et on pense qu'elles sont recrutées à partir de cellules B folliculaires matures (27). Alors que les cellules B folliculaires et MZB expriment l'Ag B220, les cellules B-1 sont B220neg/faible et expriment le CD43.

1.2.2 Réaction du centre germinatif

1.2.2.1 Structure et développement des centres germinatifs

Dans les organes lymphoïdes secondaires au repos, les cellules B et T polyclonales sont compartimentées et ne se divisent pas. Suite à une immunisation avec un antigène, les CG vont apparaître en quelques jours et persister jusqu'à plusieurs semaines (**Figure 3**).

Lorsqu'un Ag entre dans l'organisme, il est capturé par des cellules présentatrices d'Ag comprenant les LB, les macrophages et les cellules dendritiques folliculaires (CDF). Le LB reconnaît l'antigène pour lequel il a une affinité via son BCR membranaire. Cela entraîne l'activation du LB et induit l'expression de CCR7, favorisant sa migration à l'interface des zones B et T. En effet, la seule interaction entre l'antigène et le BCR n'est pas suffisante pour activer le LB et déclencher la synthèse d'Ac. Les LB ont besoin d'un second signal, qui peut être apporté par les lymphocytes T folliculaires auxiliaires (ou T follicular helper, Tfh) dans le cadre d'une coopération T-B. Simultanément à l'activation des LB, les lymphocytes T CD4+ qui ont détecté leur Ag apparenté présenté par la CDF, commenceront à exprimer le répresseur transcriptionnel BCL6 (28), régulateur principal de la lignée des Tfh (29). La sécrétion des cytokines TGF- β , d'IL12, d'IL23 et l'expression d'ICOS par les lymphocytes T CD4+ activés favorise leur différenciation en Tfh. Ces derniers continuent d'exprimer BCL6 et produisent de l'IL21, perdent l'expression de CCR7 au profit de CXCR5, ce qui leur permet de migrer vers la zone B pour y rencontrer le LB qui vient lui aussi d'être activé (30).

Par la suite, d'importantes interactions récepteur-ligand sont formées, notamment CD40:CD40L, ICOSL:ICOS (31). Ces interactions garantissent que les LB et LT restent en

contact pour déclencher la phase suivante de la réponse, un à deux jours plus tard. La Tfh mature, spécifique de l'Ag, se déplace vers le centre du follicule. Les cellules B activées ont plusieurs destins à leur disposition (28). Certaines cellules B se déplacent vers la zone inter-folliculaire et se différencient en cellules sécrétrices d'anticorps (ASC) à courte durée de vie, fournissant une réponse rapide à l'infection, mais de faible affinité. D'autres cellules B se déplacent vers la zone pré-CG, probablement sous l'effet de la régulation positive de BCL6 et des modifications de l'expression des récepteurs de chimiokines. Elles y subiront une maturation d'affinité dans le CG en développement. Enfin, certaines cellules B se différencieront en cellules B mémoires précoces après cette interaction initiale entre les cellules T et les cellules B, en amont de la formation du CG (32,33) (**Figure 3**). Il est néanmoins admis que la majorité des cellules B mémoire sont issues de la réaction du CG (34).

Après quelques jours de prolifération massive, la structure caractéristique du CG devient apparente : une zone sombre (DZ, Dark Zone) constituée presque exclusivement de cellules B proliférant en masse et connues sous le nom de centroblastes, et une zone claire (LZ, Light Zone) composée de centrocytes plus petits, non divisés, situés dans un maillage de CDF, de cellules T et de macrophages (**Figure 3**). Le CG atteint sa taille maximale dans un délai d'environ deux semaines. Les centroblastes diversifient leurs gènes IgV par le processus d'hypermutation somatique (SHM, voir paragraphe 1.2.2.2), et les cellules qui expriment des Ac modifiés nouvellement générés sont sélectionnées dans la LZ sur l'affinité de leur Ig pour l'Ag. Certains centrocytes se différencient en cellules B mémoires ou en plasmocytes, après avoir subi une commutation isotypique (Class Switch Recombination, CSR, voir paragraphe 1.2.2.3).

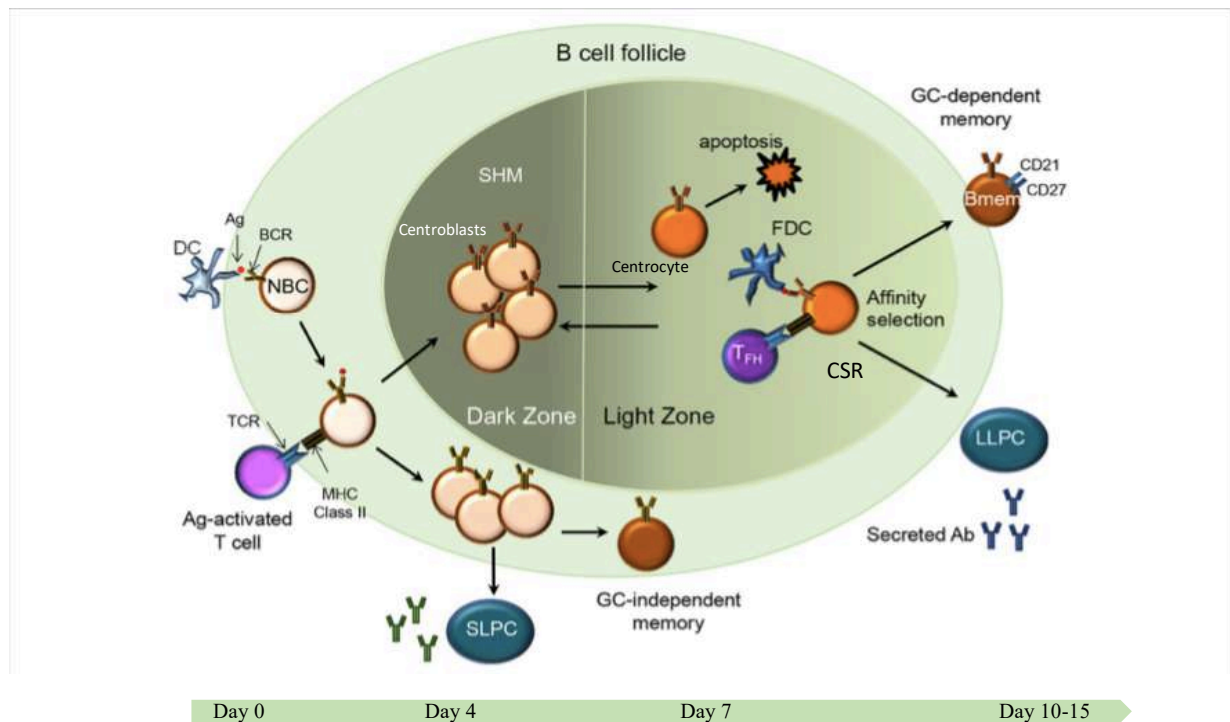


Figure 3 : Réaction du centre germinatif et génération des B mémoires.

Dans les organes lymphoïdes secondaires, les cellules T et les cellules B naïves (Naive B Cell, NBC) activées par l'Ag forment des interactions stables lors de leur migration vers le follicule primaire et la zone des cellules T, respectivement. Les lymphocytes B prolifèrent, une partie se différencie en B mémoires indépendants du centre germinatif (Germinal Center, GC) ou en plasmocytes à courte durée de vie (Short Lived Plasma Cell, SLPC), sécrétant des Ac dans un foyer extra-folliculaire. L'autre partie des cellules B spécifiques de l'Ag subissent une expansion clonale et des hypermutations somatiques (SHM) pour diversifier leurs gènes d'immunoglobulines dans la zone sombre du follicule primaire. Les cellules B avec une BCR diversifié peuvent se déplacer vers la zone claire où elles rencontrent des cellules dendritiques folliculaires (FDC) et des cellules T follicular helper (TFH). Grâce au processus de sélection d'affinité, les cellules B ayant une forte affinité pour l'Ag survivent tandis que celles ayant une faible affinité subissent l'apoptose. Après la maturation d'affinité, les lymphocytes B peuvent soit revenir dans le CG, soit en sortir sous forme de B mémoire ou de plasmocytes à longue durée de vie (Long Lived Plasma Cell, LLPC) sécrétant des ac. (Adapté de Shah et al (33))

Cette description classique de la réaction du CG englobe la plupart des événements cellulaires, mais s'est révélée trop schématique, car la compartimentation de la réaction du CG est souvent moins bien définie que le schéma classique DZ-LZ (35-37). En outre, le déplacement des cellules B dans le microenvironnement du CG ne suit pas toujours la direction de la DZ vers la LZ (38). En effet, les cellules B en prolifération peuvent être détectées dans la LZ des CG spléniques de souris (39,40), et des études en microscopie électronique ont montré que le trafic des BGC est bidirectionnel entre DZ et LZ (41,42). Certains progrès technologiques ont également été réalisés dans l'identification des mécanismes qui guident le trafic des cellules B entre les zones sombres et claires, qui semble être médié par des chimiokines qui sont peut-être produites par des cellules stromales (43,44). Les centroblastes expriment CXCR4 et migrent vers un gradient de CXCL12 (également appelé SDF1) qui provient de la LZ, tandis que les centrocytes

expriment le CXCR5, qui répond à un gradient de CXCL13 produit dans la DZ. L'alternance de la régulation à la hausse et à la baisse de l'expression de ces deux récepteurs de chimiokines sur les cellules BGC pourrait favoriser le cycle des cellules dans les deux sens entre la DZ et la LZ. Ceci est conforté par une étude récente où du RNA-seq unicellulaire a été réalisée dans les cellules BGC. Cela permet aux auteurs d'identifier plusieurs sous-populations, mettant en évidence une hétérogénéité importante au sein des BGC, notamment avec des populations dont le profil d'expression génique est intermédiaire entre les cellules caractéristiques de la LZ et DZ et qu'ils rattachent à des précurseurs des B mémoires ou des plasmablastes. En outre, ils confirment que bien que la prolifération soit une caractéristique principalement associée aux cellules BGC de la DZ, les cellules en prolifération et au repos se trouvent dans les compartiments DZ et LZ (37).

Une caractéristique reconnue depuis longtemps des centroblastes est leur taux de prolifération extrêmement rapide ; les estimations du temps nécessaire à l'achèvement d'un cycle cellulaire vont de 6 heures à 12 heures (42). Ce taux de prolifération rapide est nécessaire pour la génération à haute cadence d'un grand nombre d'Ig modifiées, parmi lesquelles sont sélectionnées les quelques cellules B présentant des Ac avec une meilleure liaison à l'Ag. En conséquence, l'analyse du profil d'expression des gènes a montré que la différenciation d'une cellule B activée par un Ag en un centroblaste s'accompagne d'une augmentation spectaculaire de l'expression des gènes associés à la prolifération cellulaire et d'une diminution des gènes codant pour des régulateurs négatifs de l'expansion clonale, tels que p21 (pour revue : (45,46)). Cependant, la physiologie des centroblastes diffère nettement de celle des autres cellules proliférantes de l'organisme. En particulier, la détection et la réponse aux dommages causés à l'ADN sont spécifiquement supprimées dans les centroblastes (47). En outre, les centroblastes expriment l'enzyme télomérase (48), ce qui garantit que la vigoureuse expansion clonale des cellules BGC n'entraîne pas une perte de leur potentiel de réplication, du fait du raccourcissement des télomères.

1.2.2.2 Hypermutations somatiques

Les centroblastes du CG subissent des mutations ponctuelles, des insertions ou des délétions dans la séquence nucléotidique codant pour la région variable de l'Ig de surface, modifiant ainsi son affinité pour l'Ag. À la suite de ce phénomène, seuls les centrocytes exprimant à leur surface des Ig de forte affinité pour l'antigène seront sélectionnés, grâce aux cellules folliculaires dendritiques et aux cellules T folliculaires auxiliaires présentes au sein de la zone claire des CG.

La CG est la structure principale, mais pas le site exclusif où la SHM se produit, car la SHM extrafolliculaire a également été observée chez la souris et chez l'homme (49,50). Le processus de SHM est initié par l'activité d'AID (Activation Induced cytidine Deaminase), enzyme convertissant les cytidines de l'ADN en uracile (51-53). Cela a pour conséquence

de déclencher des mécanismes de réparation de l'ADN infidèles qui conduisent à l'introduction de mutations somatiques (pour revue : (54)).

Dans les CG, la SHM se limite principalement aux centroblastes proliférants de la DZ. En effet, de la diversité intraclonale des IgV a été rapportée dans des clones de cellules B isolés par microdissection à partir de cellules individuelles de la DZ des CG (55), et est compatible avec l'expression d'AID dans les centroblastes (56).

1.2.2.3 Commutation de classe

La recombinaison par commutation de classe (CSR) est le processus par lequel la classe de chaîne lourde de l'Ac produit par un LB activé passe de l'IgM et de l'IgD à l'IgG, l'IgA ou l'IgE, qui ont des fonctions effectrices distinctes. Le processus se produit au sein des CG, mais est également observé en dehors (57,58). La CSR cible des zones de recombinaisons appelées régions Switch ou S, situées en amont de tous les domaines constants à l'exception de C δ . L'expression d'AID est nécessaire (52,53) et initie le processus de commutation en générant une déamination des cytosines en uraciles au niveau des régions S. Différentes enzymes de réparation de l'ADN agissent ensuite en cascade. La CSR peut se produire de manière dépendante ou indépendante des cellules T, et c'est la combinaison de cytokines spécifiques et de signaux co-stimulants qui détermine la nature de la classe d'Ig résultante.

1.2.3 Rôle des facteurs de transcription

1.2.3.1 Initiation et déroulement de la réaction du CG

Le répresseur transcriptionnel BCL6 orchestre la formation et le maintien du CG (59,60). Il permet aux cellules de surmonter le stress génotoxique physiologique associé aux phénomènes de SHM et CSR sans provoquer les réponses apoptotiques et d'arrêt de cycle cellulaire qui se produiraient dans d'autres conditions, principalement via la répression de TP53 (pour revue : (61)). BCL6 agit en coopération avec BACH2, un autre répresseur essentiel de la réaction du CG, par la co-régulation de plusieurs gènes cibles (62–64). L'une des fonctions majeures de BCL6 dans le maintien de la réaction du CG est l'inhibition du FT PRDM1/BLIMP1, reconnu depuis longtemps comme un régulateur crucial du développement des plasmocytes. Les modèles murins chez lesquels l'expression de BLIMP1 est conditionnellement éteinte ne peuvent pas générer de plasmocytes (65). Il a ensuite été démontré que BLIMP1 agissait en amont de la protéine de liaison à la X-box 1 (XBP1), un FT nécessaire au phénotype sécrétoire des plasmocytes (66,67). Enfin, IRF4 est un régulateur essentiel de la différenciation des plasmocytes, réprimant BCL6 et mettant ainsi fin au programme de transcription du CG (68,69). Des travaux récents ont suggéré que l'expression de BLIMP1, IRF4 et XBP1 est régulée indépendamment, mais que les trois FT

sont conjointement nécessaires à l'établissement des plasmocytes différenciés en phase terminale, comme le montrent les effets spectaculaires de leurs déficiences individuelles sur le développement des plasmocytes (65,66,68,69). La différenciation terminale d'une cellule B en plasmocyte nécessite l'inactivation fonctionnelle du répresseur transcriptionnel PAX5, qui est essentiel pour maintenir l'identité des cellules B matures, y compris les cellules B naïves, BGC et mémoires (23). La régulation à la baisse de PAX5 dans les cellules B matures, induite expérimentalement, conduit à la mise en place du programme de différenciation des plasmocytes (70,71). En conséquence, il est admis que l'inhibition de PAX5 représente la première étape vers la différenciation en plasmocytes (72) (**Figure 4**).

MYC, un autre facteur permettant la prolifération et contribuant à la transformation des cellules B, est nécessaire à la formation des CG (73). Son expression est induite précocement lors de l'initiation du CG puis s'éteint dans les LB de la DZ, ou il est reprimé par BCL6 et est enfin ré-exprimé de façon transitoire dans une sous-population de LB de la LZ destinés à retourner dans la zone sombre (74–76). Bien que l'importance biologique de l'extinction de l'expression de MYC dans les LB de la DZ reste inconnue, l'induction initiale de son expression active des programmes transcriptionnels qui sont associés à ses nombreuses fonctions, comme la réplication de l'ADN, l'entrée dans le cycle cellulaire et la maintenance des télomères (77).

IRF8, membre de la famille IRF, présente une forte homologie avec IRF4 mentionné plus haut. La redondance de leurs fonctions a été établie dans les stades précoces de développement B mais elles semblent adopter des rôles distincts voire antagonistes dans les stades plus tardifs de différenciation. L'expression d'IRF8 est maximale dans le CG pour s'éteindre dans les ASC, contrairement à IRF4 qui voit son expression augmenter des BGC vers les ASC (78) (**Figure 4**). Les souris déficientes pour IRF8 dans la lignée B sont capables d'engager une réaction du CG (79), contrairement aux souris déficientes pour IRF4 (80). IRF4 comme IRF8 interagissent avec d'autres FT, en particulier SPI1 (voir paragraphe 4.6.3.1). Les cellules B double knockout (DKO) *Spi1/Irf8* montrent une différenciation accélérée en ASC *in vivo* et *in vitro* (78), suggérant un rôle négatif des deux facteurs sur la différenciation en ASC.

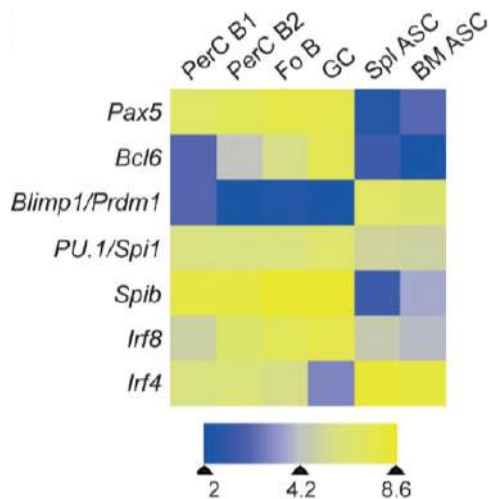


Figure 4 : Expression des facteurs de transcription clés de la différenciation B terminale (78).

Les données de la Heat-map sont extraites d'analyses de RNA-Seq.

PerC B1 (peritoneal B1 cells); *PerC B2* (conventional B2 cells in peritoneal cavity); *Fo B* (follicular B cells), *GC* (Germinal Center B cells); *Spl ASC* (spleen Antibody Secreting Cell), *BM ASC* (Bone Marrow Antibody Secreting Cell).

Le FT ETS1 a aussi été identifié comme jouant un rôle clé dans la différenciation terminale des LB. Les souris dépourvues d'ETS1 présentent un nombre nettement plus élevé de plasmocytes sécrétant des IgM dans leurs organes lymphoïdes (81,82). De plus, les cellules B dépourvues d'ETS1 se différencient davantage en plasmocytes lorsqu'elles sont cultivées en présence de CpG (83). En plus de sa fonction de régulation positive de l'expression de *Pax5* et d'autres gènes importants pour l'identité des cellules B, ETS1 bloque également la différenciation des plasmocytes par sa capacité à interagir directement avec BLIMP1 et à inhiber sa fonction. Il est donc établi que l'expression d'ETS1 doit être régulée à la baisse avant que la différenciation plasmocytaire puisse avoir lieu (84).

Un autre membre de la famille ETS, ELK1, intervient dans cette différenciation terminale. Une étude de Hipp *et al.* a montré que la voie de signalisation IL2-ERK-ELK1 régulait l'expression de BACH2 et donc la différenciation des plasmocytes (64). En effet, BACH2 est un répresseur transcriptionnel dont une des cibles majeures dans les LB est BLIMP1 (85). En l'absence de BACH2, la formation de LB matures n'est relativement pas affectée, mais lorsque ces LB sont activés, BLIMP1 s'exprime prématurément, ce qui entraîne une répression d'AID et une exacerbation de la différenciation plasmocytaire (86).

1.2.3.2 Transition du plasmablaste au plasmocyte

Un nombre important d'études se sont attachées à détailler les FT qui contrôlent l'engagement vers la différenciation terminale ; cependant, moins de données sont disponibles concernant la transition du plasmablaste (cellule en division et à courte durée de vie) vers le plasmocyte (cellule ne cyclant plus et à longue durée de vie). Phénotypiquement, la forte expression du CD138 et la diminution de l'expression du B220 et/ou CD19 est largement corrélée avec les étapes de finale maturation en plasmocytes (87). Plasmocytes et plasmablastes peuvent également être distingués sur la base de l'expression de BLIMP1, les plasmablastes présentant un niveau d'expression de BLIMP1 nettement inférieur à celui des plasmocytes (88,89). Les plasmocytes suppriment ainsi

l'expression de MYC, cible de la répression transcriptionnelle engagée par BLIMP1, ainsi qu'une partie du programme transcriptionnel induisant le cycle cellulaire (76,90).

1.2.3.3 Choix B mémoire / plasmocytes

Les cellules B mémoire forment un groupe minoritaire de cellules à longue durée de vie, persistant à l'état quiescent sans proliférer. Elles sont capables de générer une réponse accélérée et de forte amplitude aux pathogènes pour lesquelles elles sont spécifiques. En effet, elles peuvent présenter rapidement et efficacement l'antigène aux LT lors d'une réponse secondaire, proliférer puis se différencier en plasmocytes. Les cellules B mémoires peuvent ré-intégrer les CG pour de nouveaux cycles de sélection et d'hypermutation, permettant ainsi de perfectionner la mémoire immunitaire (74,91).

Chez l'Homme, la moitié des cellules B mémoire du sang périphérique a subi une recombinaison de la classe des Ig (~23% IgG et ~21% IgA), et l'autre moitié co-exprime encore les IgM et IgD de surface (~52% des cellules B mémoire), ou les IgM seules (<5% des cellules B mémoire). Les cellules B mémoires humaines représentent environ 30% des cellules B circulantes, le reste étant constitué de cellules B naïves. En revanche, il existe une différence notable chez la souris, où moins de 5% des B circulants sont des B mémoires (92,93). Il a longtemps été admis que le compartiment B mémoire humain était exclusivement ou principalement composé de cellules B switchées, mais des études ont démontré la présence de cellules B IgM+ ayant subi les hypermutations somatiques (92). Les LB mémoire issus du CG sont caractérisés par la perte de l'expression de l'IgD de surface et l'acquisition chez l'Homme du marqueur CD27. L'origine CG de cette population B mémoire IgM seule est maintenant acceptée, tandis que les B mémoire IgM+/IgD+/CD27+ correspondraient aux cellules B de la zone marginale splénique (contribuant à la protection contre les pathogènes T-indépendants transmis par le sang) (33). Il a en effet été montré que des cellules B mémoires pouvaient être générées en dehors des CG (94,95), pourtant longtemps considérés comme leurs seuls sites de génération. Des études récentes ont rapporté que des cellules B mémoire non mutées peuvent être générées chez des souris irradiées reconstituées avec de la MO déficiente en Bcl6, donc incapable de former des CG. Les cellules B mémoire en contexte déficient en Bcl6 se sont donc différenciées indépendamment des CG, ne portaient pas de mutations somatiques et n'ont pas subi de maturation d'affinité. Cependant, elles correspondaient bien à la définition des B mémoires : il s'agissait de cellules quiescentes, à longue durée de vie, capables de produire de plus grandes quantités d'Ac lors d'une réponse secondaire par rapport aux cellules B naïves (95,96).

Le destin cellulaire du centrocyte vers la voie mémoire ou plasmocyte serait dicté, entre autres, par l'affinité du BCR à sa surface. En effet, Shinnakasu *et al.* (34) ont montré chez

la souris que les cellules BGC de la LZ avec des BCR de moindre affinité ont une expression de BACH2 plus élevée que celle de leurs homologues avec des BCR de plus grande affinité et sont prédisposées à entrer dans le compartiment des cellules B de la mémoire. L'haploinsuffisance en BACH2 entraîne une diminution des B mémoire, et l'expression de BACH2 dans les cellules du CG est inversement corrélée à la force de l'aide fournie par la cellule T (34). Leurs données suggèrent un modèle dans lequel une aide faible de la part des cellules T maintient une expression relativement élevée du FT BACH2 dans les cellules BGC LZ, les prédisposant à entrer dans le compartiment des cellules B mémoire. Le mécanisme de cette coopération B-T reposerait sur la sécrétion d'IL2 par les TFH, entraînant une activation de la voie de signalisation ERK/ELK1 et *in fine* la répression de l'expression de BACH2, permettant l'engagement vers la différenciation plasmocytaire (64).

D'autre part, les TFH sécrètent également de l'IL9, qui oriente la différenciation des cellules B du CG vers les B mémoire, leurs précurseurs ayant une expression accrue du récepteur à l'IL9 (IL9-R) (97,98).

Il a été suggéré que STAT5 phosphorylé serve de médiateur pour la différenciation des cellules B mémoire par rapport aux plasmocytes en régulant l'expression de BCL6 (99), et auparavant, on a invoqué le rôle de BCL6 dans le développement des cellules B mémoire (100). Comme les cellules B mémoire continuent d'exprimer PAX5, il est probable que l'inactivation de PAX5 soit l'événement clé pour la différenciation d'un centrocyte en plasmocyte plutôt qu'en B mémoire.

1.2.3.4 Sortie des organes lymphoïdes secondaires

Les ASC vont quitter les organes lymphoïdes secondaires pour rejoindre la circulation sanguine et finalement rejoindre la MO, processus appelé 'homing', grâce à l'activation de S1PR1 (sphingosine-1-phosphate receptor 1). CXCR4 (CXC-chemokine receptor 4), récepteur de CXCL12 (chemokine CXC-chemokine ligand 12) est impliqué dans le recrutement des ASC dans la MO et dans leur maintien sur ce site (101,102).

MYB, facteur de transcription et protooncogène exprimé pendant le développement des cellules B, est essentiel pour que les plasmablastes dérivés du CG puissent migrer vers la MO. En l'absence de MYB, les plasmablastes ayant switché leur Ig de surface, générés après une immunisation ou une infection, n'ont pas pu rejoindre la circulation sanguine, et se sont accumulés dans la rate. L'effet de la perte de MYB n'a donc pas eu d'impact sur la formation de CG. Le principal défaut détecté était l'incapacité des plasmablastes issus du CG à migrer en réponse à CXCL12 et donc à contribuer au pool de plasmocytes médullaires à longue durée de vie (103).

1.2.4 Régulation épigénétique de la différenciation lymphoïde B

Un certain nombre de FT indispensables à la différenciation des cellules B ont été découverts, dont BCL6 pour la différenciation des cellules B du CG, ou BLIMP1, IRF4 et XBP1 pour la différenciation des plasmocytes (voir paragraphe 1.2.3 et **Figure 4**). Bien que d'une importance capitale pour assurer les programmes transcriptionnels du CG et des plasmocytes, ces FT doivent avant toute chose pouvoir se lier à leurs cibles. Ceci est dépendant de l'accessibilité de la chromatine, elle-même régulée par des modificateurs épigénétiques. Ainsi, la différenciation en B mémoires, plasmocytes à courte ou longue durée de vie repose sur une combinaison de régulateurs moléculaires, modulés de façon dynamique au cours de la réponse immunitaire. Ces dernières années, plusieurs études ont mis en évidence le rôle crucial que jouent les régulateurs épigénétiques dans la différenciation des cellules B lors de la réponse humorale, via la méthylation de l'ADN et les modifications des histones.

Dans la suite du manuscrit, le terme « enhancer » est utilisé pour désigner les éléments cis-activateurs, classiquement définis comme des séquences d'ADN situées à des distances variables du site de démarrage de la transcription, agissant en cis et capables d'augmenter la transcription des gènes. Le terme « promoteur », quant à lui, fait référence à des régions de l'ADN qui permettent l'initiation de la transcription d'un gène, généralement situées à proximité et en amont du site de démarrage de la transcription.

La méthylation de l'ADN consiste à ajouter des groupements méthyles aux résidus CpG, par l'action des méthyltransférases. Il existe trois ADN méthyltransférases (DNMT) : DNMT1, DNMT3A et DNMT3B. DNMT3A et DNMT3B assurent la méthylation *de novo* (104), tandis que DNMT1 permet la maintenance des profils de méthylation au cours des divisions cellulaires en reproduisant sur le brin néosynthétisé, lors de la réplication de l'ADN, la méthylation présente sur le brin matrice (105). La déméthylation de l'ADN est effectuée par la famille d'enzymes Ten Eleven Translocation (TET1, TET2 et TET3), en catalysant l'oxydation de la 5-méthylcytosine (5mc) en 5-hydroxyméthylcytosine (5hmc), conduisant *in fine* à la déméthylation (voir chapitre 3).

La méthylation des CpG est généralement associée à la répression des gènes, agissant comme une plateforme de recrutement pour les protéines qui possèdent des domaines de liaison au méthyle et initient la condensation de la chromatine. Il a été montré que la méthylation de l'ADN en elle-même favorise la condensation de la chromatine, en rigidifiant et compactant les nucléosomes (106). La méthylation de l'ADN peut également faciliter la répression des gènes en bloquant physiquement la liaison des régulateurs transcriptionnels.

La méthylation de l'ADN joue un rôle dans l'engagement des progéniteurs hématopoïétiques vers la lignée myéloïde ou lymphoïde (107,108). L'hypométhylation au

niveau des gènes clés de l'identité B dirige la cellule vers la lignée lymphoïde B (107). La comparaison des profils de méthylation des B naïfs, des BCG, des B mémoires et des plasmocytes humains indiquent que ces cellules subissent une déméthylation significative au cours de la réponse immunitaire. En d'autres termes, les sites qui étaient méthylés dans les cellules B naïves sont déméthylés lors des stades ultérieurs de différenciation, après stimulation antigénique (109–111). Orlanski *et al.* ont confirmé que l'expression de TET2 et TET3 pendant la différenciation des LB se traduisait par l'apparition de la marque 5hmc, première étape vers la déméthylation. Presque toutes les séquences déméthylées pendant l'engagement des cellules B restaient méthylées chez les souris déficientes en TET2/3, et un blocage partiel du développement des cellules B était observé (112). Les séquences cibles déméthylées par TET2/3 étaient enrichies en H3K27Ac et H3K4me1 et avaient une conformation de chromatine plus ouverte, suggérant une déméthylation préférentielle au niveau des enhanceurs (113). L'expression des protéines TET diminue au cours de l'engagement dans la différenciation des plasmocytes, ce qui est corrélé à des niveaux très faibles, voire indétectable de 5hmc dans les plasmocytes de la MO (114). Globalement, ces études montrent que l'engagement vers la différenciation plasmocytaire est associé à une perte de méthylation (109,110,113). Les cellules qui se différencient en plasmablastes/plasmocytes ont au préalable proliféré, cependant il a été montré que la perte de méthylation n'était pas liée à un phénomène de dilution suite aux divisions cellulaires mais qu'elle reposait sur l'enrichissement spécifique en 5hmC au niveau des gènes de l'identité plasmocytaire, ce qui conforte le rôle des protéines TET dans ce processus (114).

DNMT1 et dans une moindre mesure DNMT3B sont régulées à la hausse dans les cellules BGC. En revanche, DNMT3A est régulée à la baisse (115). Les souris hypomorphes pour DNMT1 ont des défauts de réaction du CG. La régulation de la réaction du CG par les enzymes DNMT a été confirmée par le traitement de souris sauvages immunisées avec la décitabine, un inhibiteur de l'ADN méthyltransférase (115). DNMT1 est recruté au niveau de l'ADN hémiméthylé par UHRF1, (membre des RING-finger type E3 ubiquitin ligases) pour réprimer l'expression des gènes (116). La formation des CG, la maturation d'affinité et la réponse des Ac étaient toutes défectueuses en l'absence d'UHRF1, ce qui reflète son rôle dans la régulation de la prolifération des cellules BGC via la méthylation des gènes régulateurs du cycle cellulaire tels que *Cdkn1a* (codant pour p21) (117). Il semble donc que le maintien de la méthylation de l'ADN par le recrutement d'UHRF1-DNMT1 soit important pour une régulation appropriée des cellules BGC. Barwick *et al.* ont étudié les effets du KO conditionnel de *Dnmt3a* et *Dnmt3b* dans les cellules B. Alors que le développement des cellules B semblait normal chez les souris DNMT3-déficientes, ils ont

constaté après stimulation antigénique une augmentation de la taille des CG, des titres d'Ac et des plasmocytes de la MO (118).

La liaison des FT à leur séquence de reconnaissance spécifique sur l'ADN contrôle souvent l'état épigénétique et l'activité des enhanceurs. Des signatures épigénétiques (profils de méthylation de l'ADN) sont retrouvées dans les différentes hémopathies B, des tumeurs pré-CG aux hémopathies à plasmocytes matures, ce qui aide à identifier les sous-types de cellules B physiologiques les plus « épigénétiquement » liées à un type de tumeur donné (119). La classification des patients selon leur degré de maturation épigénétique a montré un impact clinique dans plusieurs hémopathies B matures (120,121). La comparaison des patterns de méthylation de l'ADN entre cellules B tumorales et non tumorales a également permis de mieux comprendre les événements moléculaires associés à la transformation. Dans la leucémie lymphocytaire chronique (LLC), le rôle des FT NFAT, EGR, TCF3/E2A et SPI1 a ainsi été proposé (109).

Contrairement au petit nombre d'enzymes responsables de la modulation de la méthylation de l'ADN, il existe un plus large réseau d'enzymes en charge des modifications post-traductionnelles des histones, modulant la conformation et l'accessibilité de la chromatine. De façon générale, l'acétylation des résidus lysine des histones active l'expression des gènes, tandis que leur méthylation peut entraîner une activation ou une répression de la transcription en fonction du résidu modifié et du nombre de groupes méthyle présents. Par exemple, H3K4me3 active la transcription des gènes, tandis que H3K27me3 et H3K9me3 entraînent leur répression.

Les mutations somatiques affectant les gènes modificateurs d'histones sont fréquentes dans les lymphomes B diffus à grandes cellules (DLBCL) (122). Au cours des dix dernières années, des études ont mis en lumière le mécanisme par lequel les mutations de *KMT2D*, *EZH2*, *CREBBP* et *EP300* contribuent à la pathogénèse. Les mutations gain de fonction d'*EZH2* entraînent une répression des gènes de différenciation des PC via une augmentation de la marque H3K27me3 au niveau de leurs promoteurs (123,124). Les mutations perte de fonction de *KMT2D* entraînent une perte focale de H3K4me1 et donc une activation altérée des enhanceurs de gènes qui répondent normalement à la signalisation CD40 et contrôlent l'expression des gènes suppresseurs de tumeurs tels que *TNFAIP3* (125,126). Les mutations perte de fonction de *CREBBP* entraînent une perte focale de l'acétylation H3K27 au niveau des enhanceurs et donc une répression transcriptionnelle des gènes impliqués dans les signaux de sortie du CG (127,128).

1.2.5 Rôle des signaux extrinsèques

Parmi les signaux extrinsèques qui conduisent à la différenciation en ASC, on trouve le CD40L, les cytokines sécrétées par les Tfh, la signalisation par ERK1 et 2 et par la voie NF- κ B, largement impliquée dans l'activation, la croissance et la survie des LB via l'activation d'IRF4 et de BCL2 (129,130). En outre, un rôle important de la stimulation du CD40 dans la différenciation des cellules BGC en cellules B mémoire avait été suggéré sur la base d'expériences *in vitro* de cellules BGC humaines purifiées (131). Dans ces expériences, des cellules BGC cultivées sur des fibroblastes exprimant le CD40L avec une combinaison spécifique de cytokines (IL2 et IL10) ont finalement acquis un phénotype de cellule B mémoire, tandis que l'élimination du CD40L du système a entraîné la différenciation des cellules en plasmablastes (131).

2 Réaction du CG et lymphomagénèse

2.1 Remaniements génétiques à potentiel oncogénique

Au cours de son ontogénie, le LB est exposé à 3 situations de modifications génétiques physiologiques : i) le réarrangement VDJ initial dans la moelle osseuse ; ii) les mutations somatiques des gènes codant les parties variables des chaînes lourdes et légères d'Ig; iii) la commutation isotypique. Le premier risque est encouru par tous les LB, les 2 autres uniquement par des LB ayant été stimulés par leur BCR.

L'importante prolifération et le taux élevé de remaniements génétiques intervenant dans le CG rendent cet environnement particulièrement propice à la transformation tumorale, par l'acquisition de mutations et d'anomalies génomiques. Une contrepartie physiologique est identifiée dans la plupart des lymphomes B, en fonction des caractéristiques de la cellule on distingue des origines pré-CG, CG ou post-CG (pour revue : (132)) (**Figure 5**).

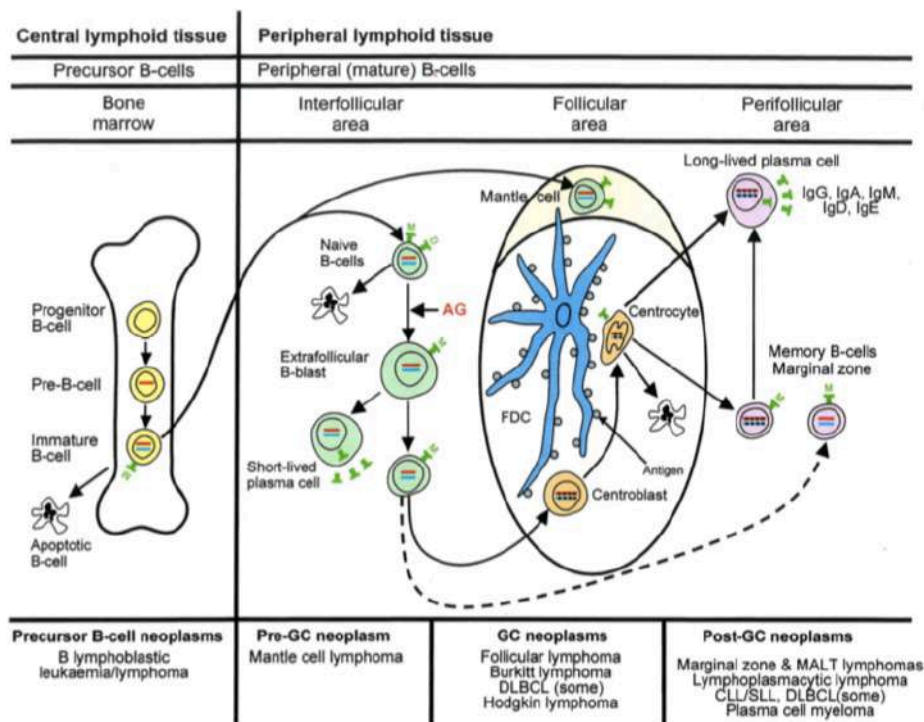


Figure 5 : Représentation de la différenciation lymphoïde B normale et de la correspondance avec les principaux lymphomes B (133).

Les hémopathies B correspondent à différents stades de différenciation des cellules B normales, bien que les contreparties physiologiques soient inconnues dans certains cas. Les cellules B précurseurs, qui se développent dans la moelle osseuse, peuvent subir une apoptose ou se transformer en cellules B naïves matures, qui, après exposition à l'antigène (AG), peuvent se transformer en plasmocytes à courte durée de vie ou entrer dans le centre germinatif (GC), où se produisent les hypermutations somatiques et la commutation de classe de la chaîne lourde d'Ig. Les centroblastes, les cellules transformées du GC, soit subissent une apoptose, soit évoluent en centrocytes. Les cellules post-GC comprennent à la fois des plasmocytes à longue durée de vie et des cellules mémoire/B de la zone marginale. La plupart des cellules B sont activées dans le GC, mais l'activation indépendante des cellules T peut avoir lieu en dehors du GC et conduit probablement aussi à des cellules B de type mémoire. Les barres rouges indiquent un réarrangement du gène de l'IGH et les barres bleues un réarrangement du gène de la chaîne légère de l'IG ; les points noirs dans les barres rouges et bleues indiquent les hypermutations somatiques.

En effet, à l'exception des lymphomes lymphoblastiques et des lymphomes du manteau, quasiment tous les lymphomes non hodgkiniens à cellules B (LNH-B) présentent des gènes d'IgV ayant subi des hypermutations somatiques ce qui indique qu'ils sont dérivés de cellules B bloquées au sein du CG ou l'ayant traversé (134,135). En outre, ces sous-types de LNH-B présentent deux types principaux de lésions génétiques - les translocations chromosomiques et les SHM aberrantes - qui représentent des erreurs ou des dysfonctionnements dans les mécanismes physiologiques de remaniement des gènes des Ig.

2.2 Stimulation antigénique chronique

Outre certains remaniements génétiques à haut risque oncogénique, un autre mécanisme favorise également la survenue des lymphoproliférations malignes en activant les voies de signalisation de la cellule B : la stimulation antigénique chronique. Elle a été d'abord décrite

dans les lymphomes gastriques du MALT (Mucosa-Associated Lymphoid Tissue) associés à l'infection chronique par la bactérie *Helicobacter pylori*. L'évolution de la pathologie est marquée par une période où la prolifération est sensible à un traitement antibiotique adapté puis, du fait d'événements oncogéniques secondaires, par l'acquisition d'une indépendance vis-à-vis de l'agent pathogène initiateur (pour revue : (136)).

L'infection chronique par le virus de l'hépatite C (VHC) représente un autre modèle favorisant la transformation des LB. De nombreuses études ont montré l'augmentation de la prévalence de l'infection par le VHC parmi les patients atteints de cryoglobulinémie mixte et de LNH-B (137). Un argument majeur pour affirmer le rôle du virus dans la lymphomagenèse réside dans l'observation de l'effet curatif du traitement antiviral sur les lymphoproliférations B (138). Les mécanismes par lesquels le VHC participe au développement de lymphome sont encore imparfaitement compris. Le rôle de la stimulation antigénique chronique, sur le modèle d'*Helicobacter pylori* dans les lymphomes des MALT, est suggéré par de nombreuses études tandis que d'autres proposent un rôle oncogénique direct du virus dans la transformation (139,140).

Cependant, pour la très grande majorité des lymphomes, Il n'existe pas d'antigène stimulant spécifique identifié. Les arguments en faveur du rôle d'un antigène viennent de l'état de différenciation de la cellule tumorale (presque jamais naïve), et du répertoire des segments VDJ utilisés par les cellules tumorales, qui dans certains cas sont communs à différents patients (répertoire biaisé).

2.3 Micro-environnement et signalisation

Les cellules B possèdent de multiples voies d'activation qui les aident à répondre aux agents pathogènes. En général, la liaison du BCR par l'Ag n'est pas suffisante pour amener la cellule B à proliférer et à se différencier, des signaux supplémentaires sont nécessaires pour une activation complète. Ces signaux de co-stimulation se répartissent en deux catégories : ceux qui sont associés à la reconnaissance innée des pathogènes et ceux qui proviennent des cellules T. Les signaux innés sont généralement transmis par la famille des récepteurs TLR (Toll-like receptor). La signalisation dépendante des cellules T nécessite l'interaction entre une cellule B spécifique de l'Ag avec une cellule T activée qui reconnaît les peptides internalisés et présentés par la cellule B. La cellule T fournit deux formes d'aide : des signaux co-stimulants, notamment grâce à l'expression du CD40L et la sécrétion de cytokines régulatrices solubles (141,142). De nombreuses études *in vitro* ont montré que le CD40L et les cytokines telles que l'IL4 délivrées à la cellule B en l'absence de tout signal Ag sont suffisantes pour stimuler la division des cellules B, la commutation de classe et la sécrétion d'Ac (143). Quel que soit le mode de stimulation, les signaux émis par l'engagement du BCR ne sont pas obligatoires pour que les cellules B répondent, mais ceux-ci modulent le comportement de la cellule, le BCR servant surtout à attirer et à concentrer les signaux de co-stimulation sur les cellules B spécifiques de l'Ag (144). Les Ac anti-Ig et l'Ag spécifique du LB délivrent de puissants signaux d'activation, affectant ainsi la viabilité des cellules et l'expression de molécules d'activation, dont le CD40 (145). Cela suggère que le signal Ag pourrait jouer un rôle important dans la régulation du taux d'interaction des cellules T-B et de leur division ultérieure.

Lorsqu'elles sont stimulées *in vivo* de manière T-dépendante (TD) ou T-indépendante (TI), les cellules B suivent le schéma typique de l'immunité adaptative : la plupart des cellules sélectionnées prolifèrent, se différencient, cessent de se diviser et meurent, tandis que d'autres deviennent des cellules effectrices et mémoires à longue durée de vie et persistent. Bien que toutes les cellules B aient initialement des caractéristiques similaires, une hétérogénéité apparaît après l'activation, certaines cellules B devenant des plasmablastes, d'autres changeant d'isotype et d'autres encore faisant les deux. Lors de la stimulation TD, une partie des cellules initie un CG et finit par donner naissance à des cellules mémoire ou des plasmocytes à longue durée de vie (146).

Les cellules B possèdent un récepteur à l'Ag et un ensemble de récepteurs additionnels qui modulent leur prolifération, leur survie et leur différenciation en réponse à la fois à la concentration et à la source de stimulation (**Figure 6**). Ces propriétés physiologiques sont détournées par les lymphocytes tumoraux, qui interagissent avec leur environnement pour assurer leur croissance, leur survie et pour se protéger des défenses immunitaires

antitumorales de l'hôte. Ces interactions sont médiées par des cytokines induisant des signaux prolifératifs, angiogéniques et/ou immunosuppresseurs et par la sécrétion de chimiokines recrutant des cellules variées.

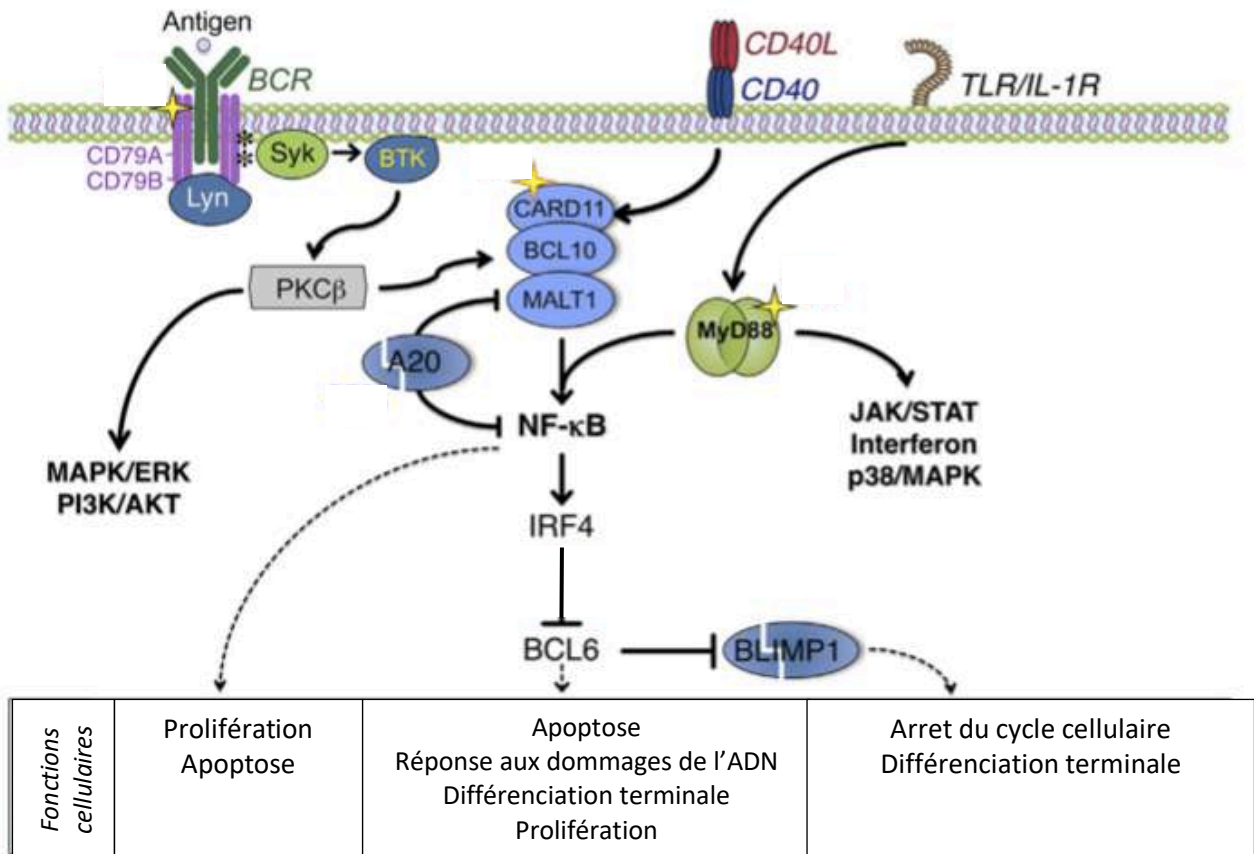


Figure 6 : Principaux acteurs intervenant dans la signalisation des lymphocytes B. Représentation schématisée des voies de signalisation induites après engagement du BCR par l'ag, interaction CD40-CD40L et activation des TLR. Tous ces signaux convergent vers l'activation de la voie NF-κB et conduisent à la régulation à la hausse de plusieurs gènes, dont IRF4. IRF4, à son tour, régule à la baisse l'expression de BCL6, ce qui permet de libérer l'expression de BLIMP1. (adapté de Dalla-Favera 2015 ((147))

2.3.1 Signalisation du BCR

Le BCR est un complexe composé d'une Ig de surface et de protéines CD79a/b chargées d'assurer la transduction du signal. L'Ig est un hétérodimère composé de deux chaînes lourdes (IGH) et de deux chaînes légères (IGK ou IGL), chaque chaîne étant composée d'une région variable, responsable de la reconnaissance antigénique et d'une région constante, dotée de fonctions effectrices (148). Après liaison de l'Ag au BCR, les sous-unités CD79a et CD79b transmettent des signaux à une variété de voies de signalisation en aval. Elles possèdent un motif ITAM (immunoreceptor tyrosine-based activation motif), composé de deux tyrosines au sein d'une séquence conservée d'acides aminés. L'agrégation du BCR favorise la phosphorylation des tyrosines de ces motifs ITAM par les

kinases de la famille Src, principalement LYN, FYN et BLK. Les motifs ITAM phosphorylés recrutent SYK par le biais d'interactions avec ses domaines SH2 (Src Homology 2) en tandem. L'activation de SYK déclenche une cascade de signalisation qui active les voies de signalisation NF- κ B, PI3K, NFAT, MAPK et RAS, conduisant à la survie et à la prolifération des LB (149) (**Figure 6**).

L'activation de la voie NF- κ B en aval du BCR implique la tyrosine kinase de Bruton (BTK, Bruton tyrosine kinase). BTK forme un complexe avec l'adaptateur BLNK (B-cell linker) et la phospholipase C γ (PLC γ). PLC γ produit ensuite le second messenger DAG (diacylglycérol), qui active la protéine kinase C β (PKC β), conduisant à la phosphorylation de CARD11, à l'activation du complexe CBM (CARD11-BCL10- MALT1) et à la signalisation NF- κ B (150).

La voie des MAP-kinases, également initiée après l'engagement du BCR, regroupe trois membres : ERK (Extracellulaire signal-regulated kinase), JNK/SAPK (c-JUN NH2-terminal kinase) et p38 MAPK (151). Après leur activation en cascade, ces kinases vont jouer leur rôle dans la signalisation en phosphorylant différents groupes de FT : ELK-1 et c-MYC (pour ERK), c-JUN et ATF-2 (pour JNK) et ATF-2 et MAX (pour p38 MAPK).

Des régulateurs négatifs contrôlent la signalisation du BCR. En plus de son rôle dans l'activation initiale de la signalisation BCR, la kinase LYN initie une boucle de rétrocontrôle négatif (152). En effet, LYN phosphoryle CD22 et le récepteur Fc γ , permettant le recrutement de la phosphatase SHP1 (SH2 domain-containing phosphatase 1), qui agit par déphosphorylation de diverses molécules contenant des motifs ITAM, dont CD79A et CD79B.

On distingue deux types de signalisation par le BCR : la première appelée « tonic BCR signaling », non liée à l'Ag, et la seconde appelée « chronic active BCR signaling », dépendante de l'Ag. La signalisation dite « tonique » est permanente et permet à la cellule B de survivre en l'absence de signal antigénique. Le signal induit par la fixation de l'Ag ou « signal actif chronique » quant à lui initie une cascade de transduction impliquant différentes tyrosines kinases et adaptateurs moléculaire au sein des voies citées précédemment pour promouvoir non seulement la survie, mais également la poursuite de la différenciation et la prolifération cellulaire (153).

En pathologie humaine, des mutations de la voie de signalisation du BCR ont été identifiées. Par exemple, des analyses de séquençage à haut débit ont rapporté dans près d'un quart des cas de DLBCL-ABC des mutations gain de fonction dans les motifs ITAM de CD79B (ou, plus rarement CD79A) (154), qui maintiennent la signalisation BCR en atténuant le

rétrocontrôle négatif impliquant l'activation par phosphorylation de la kinase LYN. La dépendance des cellules tumorales de DLBCL à la voie de signalisation du BCR est soulignée par des études cliniques récentes rapportant des taux de réponse encourageants avec l'utilisation d'agents inhibiteurs de BTK (155) (voir paragraphe 5.2.5).

2.3.2 Signalisation CD40 et CD40L : activation des LB dépendante des LT

Le CD40 appartient à la famille du TNF (Tumor Necrosis Factor) et est constitutivement exprimé sur les LB. Il interagit avec le CD40L, exprimé sur les LT auxiliaires activés. L'engagement de CD40 par CD40L favorise le recrutement des molécules TRAF (TNF receptor-associated factor), conduisant à l'activation de plusieurs voies de signalisation, telles que les voies NF- κ B (classique et alternative), MAPK, PI3K, et PLC γ (32) (**Figure 6**). La voie CD40 induit notamment l'expression d'AID, enzyme clé pour les processus de SHM et CSR. La délétion des gènes codant pour CD40 ou CD40L chez la souris provoque des défauts profonds de CSR et de production d'Ac. Chez l'homme, la mutation du gène codant pour CD40L est responsable du syndrome d'hyper-IgM. Dans les LB, la signalisation via CD40 favorise également la formation du CG et la différenciation des LB mémoire et des plasmocytes à longue durée de vie (32).

2.3.3 Signalisation BAFF et APRIL

Deux autres membres de la famille du TNF, BAFF (B-cell activating factor, codé par le gène *TNFSF13B*) et APRIL (ou *TNFSF13*) jouent un rôle central dans la fonction des LB. Des études chez des souris déficientes indiquent que BAFF est essentiel pour la maturation et la survie des LB et qu'APRIL est un facteur important pour diriger le CSR en IgA (156). BAFF et APRIL partagent deux récepteurs : TACI (transmembrane activator and CAML interactor) et BCMA (B-cell maturation antigen). BAFF et APRIL se lient à TACI avec une affinité modérée, alors que BCMA a une forte affinité pour APRIL mais une faible affinité pour BAFF. BAFF possède une haute affinité pour son récepteur BAFF-R. BAFF-R est un régulateur important de la survie des LB matures et il contrôle la production d'Ig au cours des réponses T-indépendantes. Les mutations de BAFF-R empêchent le développement correct des LB, la réponse T-indépendante ainsi que la production d'Ig (157). La liaison de BAFF à son récepteur active deux voies de signalisation favorisant la survie des LB, les voies NF- κ B et PI3K/AKT (158). TACI peut également interagir avec MYD88 (myeloid differentiation primary response gene-88), conduisant à l'activation de la voie NF- κ B classique, au CSR et à la différenciation plasmocytaire (159). En pathologie humaine, des mutations de TACI ont été identifiées dans les déficits immunitaires communs variables (160).

2.3.4 Signalisation des TLR

Les TLR sont une famille de récepteurs transmembranaires de type I qui jouent un rôle crucial dans la réponse immunitaire innée. Ils peuvent reconnaître une grande variété de composants bactériens, fongiques, protozoaires et viraux et enclencher la réponse immunitaire suite à leur contact. Ils sont abondamment exprimés sur plusieurs types de cellules de l'immunité, telles que les macrophages, les cellules dendritiques, les cellules T, les cellules B et les cellules épithéliales (pour revue : (161)). Actuellement, 10 TLR ont été trouvés chez l'homme et 12 chez la souris (162). Les TLR sont divisés en deux groupes selon leur localisation. Les TLR 1, 2, 4, 5, 6 et 10 sont localisés à la surface des cellules, et les TLR 3, 7, 8 et 9 sont localisés dans les compartiments intracellulaires.

Les TLR tiennent un rôle majeur dans la réponse des LB à l'Ag, indépendamment du BCR. Au moment de la fixation du ligand, les TLR de surface s'agrègent et initient la signalisation intracellulaire en activant divers adaptateurs intracytoplasmiques, en particulier la protéine adaptatrice MYD88 (163). MYD88 peut être recrutée par tous les TLR, sauf le TLR3 (164). Des mutations de *MYD88* ont été rapportées dans un grand nombre d'hémopathies lymphoïdes B où elles sont responsables de l'activation de la voie NF- κ B et de la survie cellulaire (165–167) (voir paragraphe 5.3.4).

2.3.5 Voie NF- κ B

NF- κ B est un FT présent dans les cellules sous forme homo ou hétérodimérique. Cinq protéines constituent les sous-unités du dimère NF- κ B : p65 (codé par *RELA*), RelB, c-Rel, p50 (codé par *NFKB1*), et p52 (codé par *NFKB2*). L'activation constitutive de la voie NF- κ B est un mécanisme fréquemment utilisé par les cellules tumorales pour éviter la mort cellulaire (168).

L'activation de la voie NF- κ B dans les LB peut se produire suite à l'activation de plusieurs récepteurs, tels que le BCR, le CD40, le BAFF-R et les TLR. La voie classique de l'activation de NF- κ B fait intervenir un complexe I κ B kinase (IKK) composé de sous-unités α , β et γ (NEMO). L'activation de ce complexe IKK par des signaux en amont est complexe, nécessitant la phosphorylation d'IKK β et l'ubiquitination d'IKK γ . Une fois actif, IKK β phosphoryle I κ B α , un inhibiteur qui à la fois séquestre les hétéro-dimères p50/p65 et p50/c-Rel dans le cytoplasme mais retire aussi ces activateurs transcriptionnels de la chromatine. I κ B α phosphorylé est ubiquitiné par l'ubiquitine ligase FBXW1A (F-box/WD repeat-containing protein 1A ; ou β TrCP) puis dégradé par le protéasome, permettant aux dimères NF- κ B de s'accumuler dans le noyau où ils activent la transcription de gènes cibles. Par ailleurs, la voie « alternative » de NF- κ B est initiée par les récepteurs CD40 et BAFF-R. Après liaison au récepteur, le complexe régulateur négatif TRAF3/MAP3K14-

TRAF2/BIRC3 est désolidarisé, permettant ainsi la libération cytoplasmique et la stabilisation de MAP3K14, kinase centrale de l'activation de cette signalisation alternative (169). Cette voie dépend d'un complexe IKK composé de deux sous-unités IKK α . Ce complexe IKK α phosphoryle p100, produit de traduction de l'ARNm de *NFKB2*, conduisant à sa protéolyse et transformation en sous-unité p52, qui se couple avec RelB et se déplace vers le noyau. Les voies classique et alternative de NF- κ B activent la transcription d'un module anti-apoptotique, comprenant les membres de la famille BCL2, comme BCL2L1 (BCL2 like 1 ; ou BCL-XL) et BCL2A1 (BCL2 Related Protein A1) ainsi que CFLAR (CASP8 and FADD like apoptosis regulator ; ou c-FLIP), BIRC2 (baculoviral IAP repeat containing 2 ; ou c-IAP1), BIRC3 (ou c- IAP2) et d'autres protéines (170,171).

La voie NF- κ B joue un rôle majeur au cours des différentes étapes de différenciation des LB (172). Pendant la réaction du CG, la signature transcriptionnelle de NF- κ B est absente des LB de la DZ mais présente dans une sous-population de LB de la LZ (46). Différentes sous-unités de NF- κ B, participant à des complexes protéiques distincts, interviennent à différents stades de la maturation lymphoïde B : REL est ainsi requis pour le maintien du CG alors que l'activation de RELA est indispensable à la différenciation en plasmocytes (129). Dans les LB de la LZ, NF- κ B active IRF4, qui régule à la baisse l'expression de BCL6 et favorise ainsi la différenciation post-CG (173). L'importance de l'axe NF- κ B-IRF4 dans la physiologie du CG a été bien documentée par les modèles murins où l'activation constitutive de NF- κ B conduisait à l'accumulation de plasmocytes alors que la délétion d'IRF4 diminuait leur formation (68,69,174).

En pathologie humaine, on retrouve une activation de la voie NF- κ B dans de nombreuses hémopathies lymphoïdes B comme les DLBCL-ABC, la Maladie de Waldenstrom (MW), le lymphome de Hodgkin (LH), le lymphome gastrique MALT et le myélome multiple. Un mécanisme récurrent de lymphomagénèse est l'inactivation de gènes codant pour des régulateurs négatifs de la voie NF- κ B. Parmi eux, on peut citer les gènes *IKBKA* (LH, (175)), *TNFAIP3* (LH, DLBCL-ABC et MALT (176,177), myélome (178)), *NFKBIE* (LLC (179,180)) et *BIRC3* (Splenic Marginal Zone Lymphoma (181,182)).

2.3.6 Coopération entre les différentes voies de signalisation des LB

La coopération entre la signalisation TLR et la signalisation BCR est un aspect essentiel de l'activation des cellules B mais dont le mécanisme n'est pas totalement élucidé (pour revue, (163)). *Pone et al.* ont montré une synergie entre la signalisation BCR activant la voie NF- κ B non canonique et la signalisation des TLR renforçant la voie NF- κ B canonique pour induire l'activation d'AID et la commutation de classe des cellules B (183).

Une étude de *Freeman et al.* suggère que la signalisation TLR contrôlait le seuil d'activation des BCR. Ils ont montré que la signalisation TLR modifie la rigidité du cytosquelette d'actine

de sorte que la mobilité des BCR est accrue, réduisant la quantité d'Ag nécessaire pour déclencher une signalisation BCR maximale (184).

D'autres études ont souligné l'importance de la voie de signalisation RAS/MEK/ERK (en aval du BCR) dans le contrôle de la différenciation des plasmocytes induite par TLR4, en fonction des signaux reçus par le BCR et des cytokines sécrétées par les cellules T (185,186).

Enfin, la kinase SYK, impliquée dans la transduction du signal du BCR, a été retrouvée activée en réponse à la signalisation par le TLR4 (185,187).

2.4 Modélisation expérimentale de l'activation des LB *in vitro*

Élucider les mécanismes de régulation de la maturation et de la différenciation dynamique des cellules BGC a longtemps représenté un challenge, en partie parce que la manipulation expérimentale des cellules BGC était limitée. Il est particulièrement difficile d'étudier la différenciation terminale des cellules B car l'affinité pour l'Ag et les réponses cellulaires aux divers stimuli induisant la différenciation sont hétérogènes et intégrés de manière dynamique à la fois dans l'espace et dans le temps (188). Malgré tout, de nombreux aspects de la différenciation plasmocytaire peuvent être reproduits dans des systèmes de culture mimant le signal du BCR, l'activation des TLR et/ou l'apport des cellules T (CD40L et sécrétion de cytokines) (**Table 1**).

Mode de stimulation / cytokines	Effet sur la différenciation des lymphocytes B	Reference
Anti-CD40	Mime l'activation T-dépendante des lymphocytes B. Nécessaire pour la réaction du GC et pour la génération des B mémoires	(189,190)
IL4	Secrétée par les T helper est également requise pour la formation des CG et la transition centroblaste-centrocyte	(191)
IL21	Nécessaire pour la production d'IgG1, la formation efficace de GC, le CSR, la maturation d'affinité et la formation de plasmocytes à longue durée de vie, mais pas pour le développement des cellules Bmem chez la souris.	(192-195)
Anti-IgM + anti-CD40	Mime l'activation T-dépendante des cellules B via l'engagement du BCR et la coopération B-T. Induisent une forte prolifération initiale. Celle-ci peut être suivie par la production d'anticorps en fonction de l'intensité du signal BCR et de la présence des cytokines appropriées.	(196)
LPS	Antigène T-indépendant issu des bactéries Gram-négatives. Induit l'activation des cellules B polyclonales, leur prolifération et la production de cytokines, telles que l'IL1, l'IL6, l'IL8 et le TNF- α . Effets médiés par la liaison à son récepteur, le TLR4. N'induit pas de CSR.	(197)
IL2	Cytokine produite par les LT. Augmente la prolifération des cellules B stimulées par le CD40L et coopère avec d'autres cytokines/facteurs stimulants pour augmenter la différenciation des cellules B activées	(198-200)
IL5	Cytokine produite par les LT. Facteur de croissance et de différenciation des lymphocytes B, favorise la commutation isotopique vers les IgA. Induit l'expression du CD25 (chaîne alpha du récepteur à l'IL2), agit par conséquent en synergie avec l'IL2 pour induire la différenciation des cellules B	(201)
CpG	Dinucléotide stimulant le TLR9, entraîne la prolifération des cellules B mais pas leur différenciation plasmocytaire	(188,202)

Table 1 : Molécules entraînant l'activation des lymphocytes B

Une équipe a publié il y a quelques années un nouveau système de culture *in vitro* sur lequel plusieurs des expériences de ce travail repose (**Figure 7**). Dans ce modèle, les cellules B murines naïves prolifèrent de façon importante, présentent un phénotype de cellule BGC et se développent *in vivo* après transfert adoptif en cellules B mémoires ou plasmocytes à longue durée de vie (203,204).

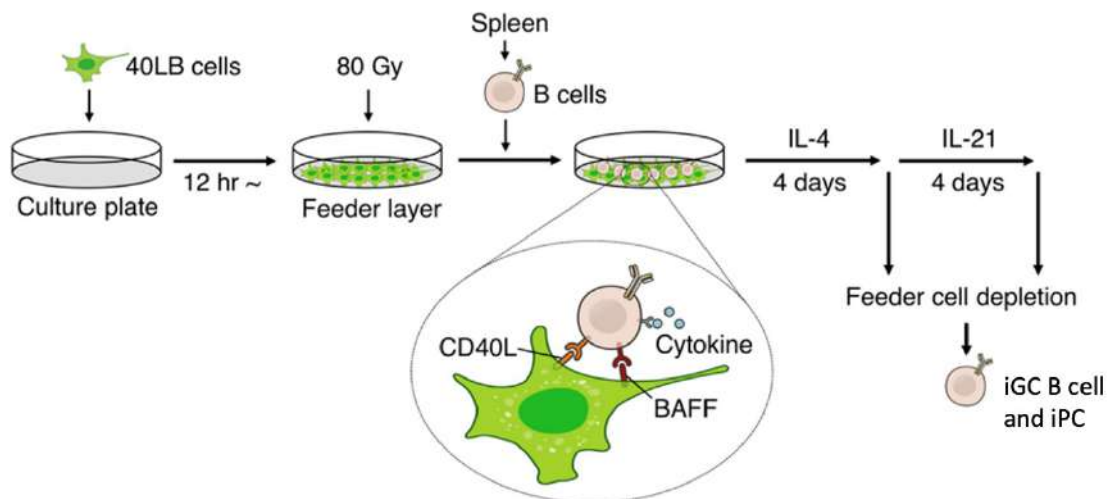


Figure 7 : Modèle de culture *in vitro* de différenciation des B naifs sur feeders

La lignée de fibroblastes 3T3 utilisée a été modifiée pour exprimer le CD40L et sécréter du BAFF (feeders 40LB). Après irradiation des feeders (80 Gray), les B naifs spléniques sont mis en culture selon deux phases successives de quatre jours chacune : une avec de l'IL4, la seconde avec de l'IL21. Dans ce modèle, l'IL21 augmente de façon importante la prolifération des cellules après la culture primaire avec l'IL4 et permet la différenciation en plasmocytes (PC). A jour 4 et jour 8, les cellules B du centre germinatif induites *in vitro* (iGC B cells) et les plasmocytes induits *in vitro* (iPC) sont analysées. (adapté de Nojima *et al.* (203))

3 Protéines TET

3.1 Généralités

La méthylation de l'ADN en 5' des cytosines (5-méthylcytosine ; 5mC) est l'une des principales marques épigénétiques, détenant un rôle crucial dans le développement et la régulation du génome (205–207). Il est connu que des régions génomiques distinctes sont différemment méthylées selon le type de cellule ou de tissu et le stade de développement (207). Outre l'établissement de "méthylomes" spécifiques à une lignée au cours du développement des mammifères, deux vagues de déméthylation globale de l'ADN ont également été signalées : l'une dans le zygote fécondé et l'autre au cours du développement primaire des cellules germinales (207,208).

TET1 (Ten-eleven translocation 1) a été le premier membre de la famille des protéines TET identifié comme une dioxygénase responsable de la catalyse de la conversion de 5mC en 5hmC (209). La famille TET comprend trois membres, TET1, TET2 et TET3, qui partagent tous un degré élevé d'homologie dans leur domaine catalytique C-terminal (210,211). La découverte de cette famille d'enzymes a suggéré un mécanisme potentiellement nouveau pour la régulation de la méthylation de l'ADN, le 5hmC agissant comme intermédiaire pendant la déméthylation de l'ADN. En outre, il a été démontré que les protéines TET peuvent oxyder davantage la 5mC ou 5hmC et les convertir en 5-formylcytosine (5fC) et/ou 5-carboxylcytosine (5caC) (212,213) (**Figure 8**). Ces formes intermédiaires ont

longtemps été considérées comme de seuls précurseurs de l'état déméthylé, mais auraient peut-être leur propre rôle de marque épigénétique (214,215).

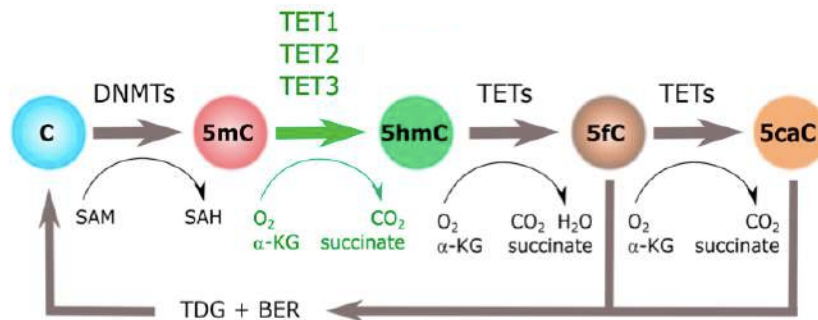


Figure 8 : Oxydation des cytosines méthylées par les dioxygénases TET. Les cytosines sont méthylées (5mC) par les ADN méthyltransférases DNMT en présence du co-facteur S-Adenosyl Methionine (SAM). Elles sont ensuite oxydées en hydroxyméthyl- (5hmC), formyl- (5fC) et carboxyl- (5caC) cytosines par les dioxygénases de la famille TET, qui ont besoin d' α -kétoglutarate (α -KG) comme co-substrat de l'activité enzymatique. Les 5fC et 5caC peuvent ensuite être excisés par des DNA glycosylases comme la thymine DNA glycosylase (TDG), conduisant à la formation d'un site abasique. Ce site est par la suite reconnu par le complexe Base Excision Repair (BER) qui réintroduit une cytosine. (adapté de Ecsedi *et al.* (216))

Toutes les protéines TET contiennent un domaine catalytique C-terminal « CD » (contenant des régions riches en Cys et des régions double-stranded β -helix, DSBH) qui présente une activité de dioxygénase dépendante du 2-oxoglutarate (2-OG) et du fer (II) (**Figure 9**). Les protéines TET oxydent 5mC en 5hmC via ces domaines CD et ont besoin d' α -kétoglutarate comme co-substrat de l'activité enzymatique (209). Une autre caractéristique des protéines de la famille TET est le domaine CXXC en doigt de zinc, présent à l'extrémité N-terminale de TET1 et TET3, mais pas de TET2. Il sert de médiateur pour leur capacité de liaison directe à l'ADN (217,218). En plus des domaines fonctionnels décrits ci-dessus, responsables de l'activité catalytique de la protéine, le domaine CD comprend également une région « spacer », dont la longueur varie entre les membres de la famille TET. Son rôle fonctionnel n'est pas totalement élucidé mais des mutations non-sens ou entraînant un décalage du cadre de lecture ont été trouvées dans cette région dans les syndromes myélodysplasiques (SMD) et les lymphomes T angioimmunoblastiques (AITL) (219,220).

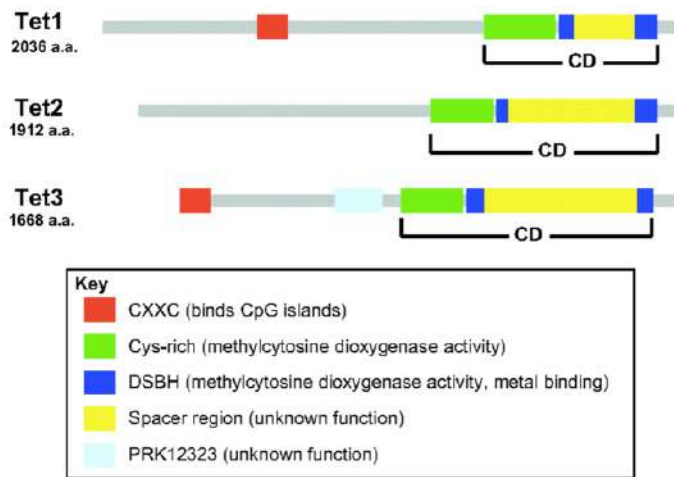


Figure 9 : Représentation schématique de la structure des protéines TET (221).

Les protéines TET1, TET2 et TET3 partagent des régions hautement conservées : une région riche en cystéine et des domaines catalytiques DSBH (*Double-Stranded β -Helix*). En revanche, TET1 et TET3 présentent un domaine CXXC de liaison à l'ADN, absent de TET2.

TET1 a été découvert à l'origine comme un gène qui, à la suite d'une translocation chromosomique, se trouvait fusionné au gène *MLL* chez certains patients atteints de leucémie aigue myéloïde (LAM) (221). Cependant, le rôle de TET1 dans l'hématopoïèse et la leucémogénèse reste ambigu. Jusqu'à présent, les mutations *TET1* et *TET3* ont été très peu décrites dans les hémopathies malignes (222). A contrario, *TET2* est le membre le plus étudié des gènes de la famille *TET* et a été signalé comme l'un des gènes les plus fréquemment mutés dans les hémopathies myéloïdes (222–226). Outre les hémopathies, les niveaux de 5hmC sont réduits dans un large éventail de tumeurs solides, notamment les gliomes, le cancer du colon, le cancer du sein et les mélanomes (227–229).

3.2 Expression et fonction

TET3 est exprimée dans l'oocyte et le zygote ; les trois protéines TET sont exprimées dans les blastocystes ; *TET1* et *TET2* sont exprimées dans les cellules souches embryonnaires (SCE) ; et *TET2* et *TET3* sont exprimées de manière ubiquitaire dans les cellules différenciées (208). Au cours des premières divisions cellulaires survenant après la fécondation, les cytosines méthylées disparaissent rapidement du pronucléus paternel au stade unicellulaire (230,231). Cette élimination coïncide avec l'apparition des 5hmC, liée à l'activité de *TET3*, fortement exprimée dans les oocytes et le zygote (232,233). Les hydroxyméthyl-cytosines sont ensuite progressivement diluées lors de la réplication de l'ADN au cours des trois divisions cellulaires suivantes, ce qui permet finalement une déméthylation globale du génome paternel (234). La réduction de l'expression du gène *Tet3* par ARN interférence (233) et la délétion du gène dans les cellules germinales femelles (235) ont permis de confirmer son rôle dans l'oxydation des 5mC du génome paternel. L'altération de la fonction de *Tet3* chez la souris conduit à une augmentation de la fréquence des anomalies de développement embryonnaire (235).

Les trois enzymes TET semblent avoir des cibles qui se chevauchent et d'autres qui sont distinctes dans le génome. Par exemple, dans les CSE de souris, *TET2* plutôt que *TET1* est

responsable de la grande majorité de la génération de 5hmC, et TET1 facilite préférentiellement la déméthylation des promoteurs tandis que TET2 et TET3 agissent sur les enhancers (236,237).

La distribution de 5hmC à l'échelle du génome reflète la forte association des enzymes TET avec la transcription des gènes. Le 5hmC est enrichi au niveau des enhancers les plus actifs et des corps des gènes les plus fortement transcrits (238). En outre, il a été démontré que de multiples FT importants pour la différenciation cellulaire et l'engagement vers une lignée, notamment NANOG, SALL4A, WT1, EBF1, SPI1 et E2A, recrutent les protéines TET sur des loci génomiques spécifiques (promoteurs et enhancers) (239–241). Par conséquent, les défauts d'expression ou de l'activité des protéines TET entraînent une différenciation cellulaire biaisée ou arrêtée dans de multiples lignées, en particulier celles du système hématopoïétique (242–244).

3.3 Régulation de l'expression et de l'activité

3.3.1 Contrôle transcriptionnel et post-transcriptionnel

L'expression des gènes *TET* peut être directement régulée au niveau de la transcription. Dans les CSE, les facteurs de pluripotence tels que *Oct4* et *Sox2* activent directement l'expression de *Tet1* en se liant à des régions non codantes conservées qui ressemblent à des sites composites Oct4-Sox2 consensus (245). De multiples microARN (miARN) contrôlent l'expression de TET après la transcription et plusieurs miARN régulent de manière coordonnée l'expression de tous les gènes *TET*. Par exemple, il a été signalé que le miR-22 oncogène réprime l'expression de TET2 en se liant directement à sa région 3' non traduite (3'UTR) dans les cancers du sein et les cellules souches hématopoïétiques (246). Les gènes *TET2* de l'homme et de la souris ont conservé des éléments qui sont censés lier miR-22 à leurs 3'UTR. L'expression de miR-22 chez les souris transgéniques imite les phénotypes observés chez les souris déficientes en *Tet2* : auto-renouvellement et transformation accrues des CSH, qui peuvent être rétablis par l'expression ectopique de protéine TET2 portant des mutations dans les sites de liaison de miR-22.

3.3.2 Protéolyse par les caspases et les calpaïnes

Les niveaux de méthylcytosines oxydées génomiques peuvent également être maintenus en régulant le turnover de TET au niveau protéique. En plus de son rôle dans le recrutement de TET2 à la chromatine, des niveaux élevés d'IDAX induisent la dégradation de la protéine TET2 en provoquant l'activation des caspases (247).

Les calpaïnes, une famille de protéases dépendantes du calcium, contrôlent les niveaux de toutes les protéines TET à l'état d'équilibre (248). Plus précisément, la calpaïne 1 régule

les niveaux des protéines TET1 et TET2 dans les CSE indifférenciées, tandis que la calpaïne 2 régule les niveaux de TET3 dans les cellules différenciées.

3.3.3 Interactions avec d'autres facteurs

Il a été démontré que la vitamine C stimule l'activité enzymatique des protéines TET, *in vivo* et *in vitro* (249–251). Cela peut être détecté par une augmentation des niveaux des produits d'oxydation de la cytosine 5hmC, 5fC, et 5caC ainsi que par une légère réduction de la méthylation globale de l'ADN en l'absence de changements dans les niveaux d'expression de TET (249). Bien que le mécanisme précis soit inconnu, il est probable que la vitamine C interagisse directement avec le domaine catalytique des protéines TET et fournisse un environnement réducteur local qui augmente l'efficacité du recyclage du cofacteur Fe(II) (249,252).

La supplémentation en vitamine C des CSE et des MEF (mouse embryonic fibroblasts) de souris entraîne une augmentation rapide et globale de 5hmC avec une diminution concomitante de 5mC (250,251) et la déméthylation de l'ADN au niveau des régions promotrices des gènes. Les changements induits par la vitamine C sur le profil de 5hmC et 5mC sont dépendants de TET1 et/ou TET2, car ces processus de déméthylation sont absents dans les ESC DKO *Tet1/Tet2* (250). Les régions génomiques résistantes à la déméthylation de l'ADN médiée par la vitamine C présentent des niveaux plus élevés de H3K9me3 (Blaschke 2013), ce qui suggère que cette marque d'histone pourrait empêcher la déméthylation de l'ADN médiée par TET2. Park *et al.* (253) ont montré que la vitamine C, à certaines concentrations, induit une inhibition de la prolifération des cellules dans les lignées cellulaires LAM et les cellules primaires LAM, qui dépend de la dose et du temps. Il y a déjà plus de 30 ans, des études avaient rapporté que la croissance des cellules progénitrices de patients atteints de LAM et de SMD est significativement modulée par la vitamine C (254,255). Deux articles récents (256,257) décrivent le rôle de la vitamine C dans la biologie des cellules souches. Ils montrent que les niveaux de vitamine C régulent le nombre et la fonction des cellules souches hématopoïétiques sanguines, en grande partie grâce à des effets sur TET2. Cimmino *et al.* ont utilisé des cellules primaires de patients LAM, qu'ils ont étudiées *in vitro* ou transplantées chez des souris. Dans les deux cas, la supplémentation en vitamine C a induit la différenciation et la mort des cellules leucémiques. Chez les souris, ces changements ont entraîné une diminution des taux de progression de la LAM (257).

N'ayant pas de domaine de liaison à l'ADN identifié, TET2 peut être recruté par les FT sur les éléments de régulation des gènes. Un certain nombre de FT interagissant avec TET2 ont été décrits, notamment SPI1 (239), WT1 (258), NANOG (241) et SALL4 (259).

TET2 est physiquement associé à la O-GlcNAc transférase (OGT), une N-acétylglucosamine transférase liée à l'O, qui ajoute un GlcNAc aux résidus Ser/Thr des protéines cibles (260,261). L'OGT se lie au domaine catalytique C-terminal de TET2, et TET2 et l'OGT ont montré une co-localisation à l'échelle du génome, principalement au niveau des promoteurs riches en GC et des sites d'initiation de la transcription (TSS, transcription starting site) qui contiennent des îlots CpG. La perte de TET2 dans les CSE empêche l'association de l'OGT avec la chromatine, et la surexpression de TET2 augmente le niveau de l'OGT liée à la chromatine. Cependant, la délétion de l'OGT n'affecte pas l'association TET2- chromatine et son activité enzymatique (260).

Enfin, un lien fonctionnel entre TET2 et CREBBP est soutenu par plusieurs études, dont celle de Zhang *et al.* qui montre que CREBBP et son paralogue p300 peuvent acétyler TET2, augmentant ainsi son activité enzymatique et sa stabilité (128). De plus, Ichiyama *et al.* ont rapporté peu auparavant que la marque 5hmC co-localise avec les enhanceurs liés par p300 et que la déficience en TET2 dans les LT diminue le recrutement de p300 au niveau des promoteurs et des enhanceurs (262).

3.4 Implication dans les mécanismes de transformation tumorale

Quelques mutations *TET1* ont été signalées dans la LLC (263), la LAM (264) et la leucémie lymphoblastique aiguë à cellules T (LAL-T) (265). De même, des mutations *TET3* sont occasionnellement identifiées dans les lymphomes périphériques à cellules T (PTCL) (266) et la LLC (263).

Parmi les gènes de la famille *TET*, *TET2* est le plus souvent muté dans les cancers hématopoïétiques. Les mutations somatiques de *TET2* ont d'abord été identifiées chez des patients atteints de LAM, syndromes myéloprolifératifs (SMP) et de SMD (267,268). Leur fréquence est d'environ 13 % dans les SMP classiques (Polyglobulie de Vaquez (PV), Thrombocytémie Essentielle (TE), (269)), 25 à 35 % dans les SMD (219,270), 50 % dans la Leucémie Myélomonocytaire Chronique (LMMC, (271)) et 12 à 34 % dans la LAM (225,272). Les mutations *TET2* sont considérées comme un événement génétique précoce dans les SMD et les SMP, et il est démontré qu'elles n'ont pas d'impact pronostique évident (273,274). En revanche, la mutation *TET2* est un marqueur de pronostic défavorable dans la LMMC (275).

Les mutations somatiques de *TET2* se produisent également dans les hémopathies B humaines (lymphomes à cellules du manteau, ~4% ; certains DLBCL, ~12%, (276,277)) et les hémopathies T (AITL, ~75%, (278,279)).

Il n'y a pas de hotspot de mutation dans *TET2*, les mutations sont réparties dans toutes ses régions codantes. Divers types de mutations sont observées : des mutations faux-sens, non-sens, des insertions/délétions entraînant un décalage du cadre de lecture et des

mutations de site d'épissage (280). Ces altérations du gène *TET2* entraînent une réduction significative des niveaux globaux de 5hmC dans les progéniteurs hématopoïétiques, ce qui suggère que l'activité enzymatique de *TET2* est affectée (281).

Les mutations *TET2* coexistent souvent avec d'autres mutations génétiques dans les hémopathies malignes humaines. Par exemple, il a été démontré que des mutations de *SRSF2*, *EZH2* et *ASXL1* coexistent avec les mutations *TET2* dans le SMD (282–284) tandis que chez les patients atteints d'AITL, les mutations *TET2* s'associent à une mutation récurrente de *RHOA* (*RHOAG17V*) et/ou des mutations de *DNMT3A* (279,285).

3.5 Perte de fonction de *TET2* et *TET3*

3.5.1 Dans l'hématopoïèse globale

Les CSH et progéniteurs murins présentent une forte expression des ARNm de *Tet2* et *Tet3* par rapport à l'ARNm de *Tet1*, et les niveaux d'expression de TET sont contrôlés dynamiquement au cours de la différenciation hématopoïétique en aval (281,286). Bien que les trois gènes *Tet* soient exprimés dans les progéniteurs hématopoïétiques, la perte de *Tet2* n'entraîne pas d'augmentation de la transcription de *Tet1* ou *Tet3* (286–288).

Le gène *TET2* est fortement exprimé dans les cellules souches et progénitrices hématopoïétiques, et est régulé à la baisse pendant la différenciation (281,289). Le knockdown (kd) de *Tet2* par ARN interférence dans les cellules souches hématopoïétiques de souris et les cellules CD34+ du sang de cordon humain entraîne une diminution des niveaux de 5hmC et favorise l'expansion de la lignée monocyttaire (287,290). De nombreux modèles murins ont permis d'étudier les effets d'une perte de *Tet2* dans les cellules hématopoïétiques. La délétion de *Tet2* dans les cellules de la MO de souris entraîne amplification du compartiment des progéniteurs multipotents hématopoïétiques. En outre, grâce à leur capacité accrue d'auto-renouvellement et de prolifération, les CSH *Tet2*^{-/-} ont un avantage compétitif sur les CSH sauvage pour le repeuplement des lignées hématopoïétiques (291). Plusieurs modèles de souris KO pour *Tet2* ont été générés en utilisant différentes stratégies de ciblage. Une fraction des souris invalidées pour *Tet2* développe une maladie ressemblant à la LMMC humaine (monocytose persistante, splénomégalie, dysplasie de plusieurs lignées myéloïdes). Bien que ce soit la pathologie prédominante, d'autres maladies ont été détectées comme des syndromes myélodysplasiques à prédominance érythroïde ou des syndromes myéloprolifératifs (291). Les souris hétérozygotes pour *Tet2* développent les mêmes maladies, (286–288) suggérant un effet d'haplo-insuffisance, ce qui est cohérent avec le fait que les patients sont souvent porteurs d'une seule mutation *TET2*. Ces études confirment le rôle suppresseur de tumeur de *TET2* et indiquent que sa délétion est suffisante pour permettre

la transformation myéloïde. Cependant la latence de développement (2 à 12 mois) suggère également que des lésions génétiques additionnelles coopèrent probablement avec la perte de Tet2 (286–288,291).

Les modèles murins rapportant les effets d'un knockdown ou d'un KO de *Tet3* sont à ce jour rares. La déficience constitutive en TET3 induit une létalité périnatale (235). Les conséquences hématopoïétiques de la déficience en TET3 ont été étudiées par Ko *et al.* (243), dont les données indiquent que le déficit en TET3 ne modifie pas de manière significative la fréquence et le nombre de cellules myéloïdes, lymphoïdes B ou érythroïdes dans la MO, mais entraîne une légère augmentation de la fréquence des cellules Lin-Sca-1+c-Kit+ (LSK) et une diminution de la fréquence et du nombre absolu de CSH. Malgré la diminution des CSH, le déficit en TET3 augmente la capacité de repeuplement des CSH dans les essais de greffe compétitive.

Une autre étude (242) rapporte l'effet du DKO Tet2-Tet3 sur l'hématopoïèse globale chez la souris. La double déficience en TET2 et TET3 entraîne la perte presque complète de la marque 5hmC dans les cellules hématopoïétiques de la MO et de la rate des animaux. Les souris ont présenté une leucocytose rapide et progressive avec neutrophilie, monocytose, thrombocytopénie et anémie sévère, qui s'est transformée en quelques semaines en une LAM très agressive chez 100% des animaux. Ce phénotype n'est pas observé chez les souris déficientes pour TET2 ou TET3, ce qui indique que TET2 et TET3 pourraient se compenser mutuellement et fonctionner de manière redondante pour supprimer l'oncogenèse.

Deux équipes ont rapporté dernièrement l'existence de familles présentant des mutations hétérozygotes constitutives de *TET2*, permettant ainsi de déterminer les conséquences d'une haploinsuffisance en *TET2* chez l'Homme. Kaasinen *et al.* (292) ont décrit une famille finlandaise avec une mutation hétérozygote constitutive de *TET2* entraînant un décalage du cadre de lecture et la perte hétérozygote de TET2 au niveau protéique. Parmi les porteurs de cette mutation, 3/6 ont développé un lymphome de Hodgkin, les trois autres membres étant des générations suivantes et n'ayant pas d'hémopathie signalée au moment de l'étude. Les porteurs de la mutation ont un degré significativement plus élevé d'hyperméthylation globale et une diminution de l'hypométhylation. Le déficit en TET2 augmente les niveaux de méthylation en particulier au niveau des enhanceurs actifs et des régions régulatrices possédant des séquences de liaison pour des FT clés tels que RUNX2, GATA1 ou SPI1.

La deuxième étude vient de Duployez *et al.* (293) qui décrivent une famille française porteuse d'un autre variant germinal de *TET2*, aboutissant également à un décalage du cadre de lecture, dans laquelle trois frères et soeurs ont développé des hémopathies

myéloïdes (LMMC, PV et LAM). Alors que les différences entre les hémopathies au sein de cette famille peuvent s'expliquer assez aisément par la présence individuelle de mutations additionnelles (*JAK2* pour la PV, *SRSF2* pour la LMMC, *BRAF* et *ZRSR2* pour la LAM), il est moins facile d'expliquer la différence de pathologie entre les deux familles (myéloïdes vs Hodgkin), mais les mutations additionnelles présentes dans la famille finlandaise ne sont pas disponibles.

Enfin, une troisième équipe a très récemment rapporté l'existence de mutations homozygotes constitutives de *TET2* (une mutation faux-sens et une non-sens) au sein de deux familles consanguines indépendantes (294). Les trois enfants concernés sont atteints d'immunodéficience sévère et ont développé respectivement un lymphome de Hodgkin, un PTCL TFH-like et un DLBCL. De façon intéressante, les auteurs ont constaté une absence de cellules B-mémoires switchées chez deux enfants et l'un d'entre eux n'a pas répondu à la vaccination anti-pneumocoque. Un modèle de différenciation *in vitro* à partir de leurs lymphocytes B primaires montre l'apparition de plasmablastes mais une incapacité à se différencier en plasmocytes matures et à sécréter des IgG.

3.5.2 Dans la lignée lymphoïde B

Comme dans de nombreux autres types cellulaires, on observe des changements importants dans la modification des cytosines de l'ADN à l'échelle du génome au cours du développement des cellules B (107). Des progéniteurs de la MO aux cellules B mémoire et plasmocytes, le génome des cellules B subit une déméthylation progressive au cours de la différenciation. Les signatures transcriptionnelles qui distinguent les progéniteurs B des cellules B matures naïves, et les cellules B matures naïves de leur descendance sont établies grâce à une réorganisation ordonnée et progressive du méthylome (107,111,115). Ainsi, chez l'homme, on estime que 30 % du méthylome est modifié au cours du développement des cellules B, affectant plusieurs millions de sites CpG (107,109).

Les ARNm *Tet2* et *Tet3* sont exprimés à tous les stades du développement des cellules B (**Figure 10**), tandis que l'ARNm *Tet1* est exprimé à des niveaux beaucoup plus faibles (281,295). Plusieurs études ont permis d'établir que les protéines TET, en particulier TET2 et TET3, jouaient un rôle essentiel dans la différenciation des cellules B et la tumorigénèse (296).

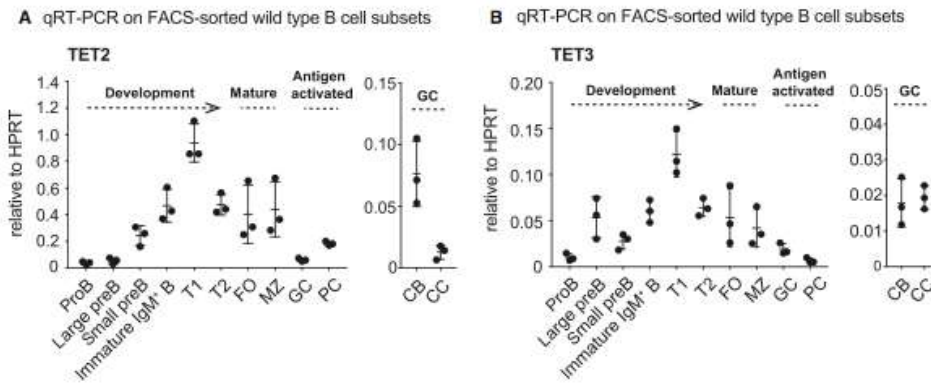


Figure 10 : Expression de l'ARNm de TET2 et TET3 en fonction du stade de développement des cellules B (295).

Analyse par qRT-PCR de l'expression de TET2 (A) et TET3 (B), normalisée par rapport à l'expression de HPRT. Cellules B transitionnelles 1 (T1), cellules B transitionnelles 2 (T2), cellules B folliculaires matures (FO), cellules B de la zone marginale (MZ), cellules B du centre germinatif (GC) et plasmablastes/plasmocytes (PC). Les cellules B du GC ont ensuite été divisées en centroblastes (CB) de la zone sombre et en centrocytes (CC) de la zone claire.

Les principaux modèles murins d'inactivation de *Tet2* et *Tet3* sont présentés dans la

Table 2.

Genotype	Phénotype observé	Anomalies spécifiques de la lignée B	Références
Tet2/Tet3 Mb1-Cre	Expansion massive de cellules B immatures ressemblant à un lymphome lymphoblastique aigu (pénétrance 100%)	Différenciation des cellules B partiellement arrêtée lors du passage du stade pro-B au stade pré-B. Pas de réponse à l'immunisation T-dépendante, absence de centre germinatif et de cellules B switchées.	(112,240)
Tet2 CD19-Cre	Développement d'hémopathies B LLC-like après une longue latence.	Accumulation d'une sous-population de cellules B1 B220low	(297)
Tet2/Tet3 Ubc-Cre ERT2	Pas de suivi des animaux	Après immunisation, nombre de BGC normal mais diminution du CSR en IgG1 <i>in vivo</i> et <i>in vitro</i>	(244)
Tet2/Tet3 Cg1-Cre	Non décrit	Après immunisation, diminution du CSR en IgG1 et IgE <i>in vitro</i> et en IgG1 <i>in vivo</i> , diminution de la différenciation plasmocytaire.	(295)
Tet2/Tet3 CD19-Cre	Manifestations auto-immune (présence d'ac anti-nucléaires), développement d'une maladie lupus-like	Splénomégalie, adénopathies, augmentation du nombre de B activés et des plasmocytes dans la rate (sans immunisation).	(298)

Table 2 : principaux modèles murins d'inactivation de *Tet2* et *Tet3*

L'une des fonctions des protéines TET au cours du développement précoce des cellules B est de réguler le réarrangement des gènes codant pour la chaîne légère Ig kappa (Ig κ), qui s'associe à la chaîne lourde déjà réarrangée pour former le récepteur complet des cellules B. Les protéines TET régulent le réarrangement du locus Ig κ en oxydant la 5mC des enhancers de Ig κ et en facilitant la déméthylation de l'ADN ainsi que l'accessibilité de la chromatine.

Les résultats de Lio *et al.* (240) ont mis en évidence une forte interaction fonctionnelle entre les protéines TET et deux FT clés exprimés dans les cellules pro-B, E2A et SPI1. En effet, ils ont observé de l'activité TET (dépôt de 5hmC et déméthylation ultérieure) au niveau des enhancers marqués par la liaison de E2A et SPI1 et des expériences de Co-IP ont confirmé l'interaction physique entre les protéines TET et E2A et/ou SPI1. Ces résultats suggèrent que E2A et SPI1 recrutent les protéines TET sur les enhancers, où elles

augmentent l'accessibilité de la chromatine en déposant 5hmC et en facilitant la déméthylation de l'ADN. En effet, le recrutement des protéines TET au niveau de l'enhancer distal du locus *Igk* était significativement diminué dans les cellules déletées pour E2A ou SPI1.

SPI1 a été proposé comme un FT "pionnier" capable de s'associer à la chromatine "fermée" (299,300). La liaison de SPI1 à l'échelle du génome était relativement similaire entre les cellules pro-B WT et Tet2/3 DKO, ce qui suggère que SPI1 se trouve en amont de TET et qu'il fait appel à l'activité de TET pour moduler l'accessibilité des enhancers cibles. Ces données suggèrent une relation fonctionnelle étroite entre les protéines TET et les FT clés au cours du développement des cellules B.

Un scénario envisageable serait une boucle de rétroaction positive dans laquelle E2A, SPI1 et d'autres FT (239,258,301) recrutent des protéines TET sur les enhancers, facilitant ainsi la modification de l'ADN médiée par le TET ; les protéines TET recrutées, via la déméthylation de l'ADN, maintiennent l'accessibilité de la chromatine à ces régions, favorisant une liaison accrue de E2A et d'autres FT tels que SPI1.

Il convient de noter que dans l'étude de Orlanski *et al.* (112), comme dans celle de Lio *et al.* (240), aucun phénotype cellulaire B frappant n'a été observé chez les animaux KO pour *Tet2* ou *Tet3* séparément, ce qui suggère qu'une déméthylation partielle des régions régulatrices clés des cellules B *in vivo* pourrait permettre un développement et un fonctionnement normaux des cellules B (**Table 2**).

Chez l'homme, des mutations de *TET2* sont fréquemment observées dans DLBCL, une hémopathie dérivée des cellules B du CG (122,302), ce qui suggère que les protéines TET peuvent réguler la fonction des cellules B matures.

Dans le modèle de Schoeler *et al.* (295), il est intéressant de noter que la double déficience en TET2/TET3 ne nuit pas à la croissance cellulaire *in vitro* bien qu'elle empêche la différenciation plasmocytaire (**Table 2**). Leurs résultats suggèrent que la fonction de TET pendant la génération de plasmablastes ne dépend pas de la prolifération. Ceci est cohérent avec un rapport de Caron *et al.* (114) qui conclut que la déméthylation de l'ADN pendant la différenciation plasmocytaire *in vitro* est associée au dépôt de 5hmC dans les régions subissant une déméthylation ciblée, indépendamment de la prolifération.

Ces résultats laissent à penser que la régulation à la baisse des protéines TET dans les cellules BGC prévient une différenciation terminale prématurée, et qu'une régulation à la hausse de TET2 est nécessaire pour une différenciation plasmocytaire optimale. Dès lors, de faibles niveaux de TET2, comme chez les patients présentant une hématoïèse clonale de potentiel indéterminé (CHIP) avec des mutations perte de fonction de TET2 (303), empêcheraient la différenciation terminale et faciliteraient ainsi la lymphomagenèse B.

4 Protéines ETS

4.1 Généralités

La famille de protéines ETS (E-Twenty Six) se compose de 28 facteurs de transcription chez l'Homme (27 chez la souris), tous contenant un domaine ETS de liaison à l'ADN, hautement conservé (304,305). Les domaines ETS sont longs d'environ 90 acides aminés et se replient en une structure en hélice-tour-hélice qui se lie à un motif commun 5'**GGA(A/T)**3' (306). En plus du domaine ETS, environ la moitié des membres de la famille contiennent un domaine SAM ou "pointed" (PNT) N-terminal qui facilite les interactions protéine-protéine et la dimérisation (305). Un alignement multiple des séquences des 28 facteurs ETS a permis de les regrouper en 12 sous-familles : 1) ETS, 2) ERG, 3) PEA3, 4) ETV2, 5) TCF, 6) ERF, 7) PDEF, 8) ELF, 9) ESE, 10) TEL, 11) SPI, et 12) ELG (**Figure 11**). Ces facteurs ETS régulent de nombreux processus essentiels, notamment le développement embryonnaire et adulte humain, le contrôle du cycle cellulaire, la prolifération et la différenciation cellulaire, l'apoptose, la migration cellulaire et l'angiogenèse (304,305,307,308). Comme ces processus sont communs à l'embryogenèse et à la tumorigenèse (309), l'expression aberrante des facteurs ETS en dehors des conditions physiologiques habituelles a été fortement associée à l'initiation et à la progression tumorale, notamment dans le cancer de la prostate (310), le sarcome d'Ewing (311), le cancer du sein (312) et les hémopathies (313).

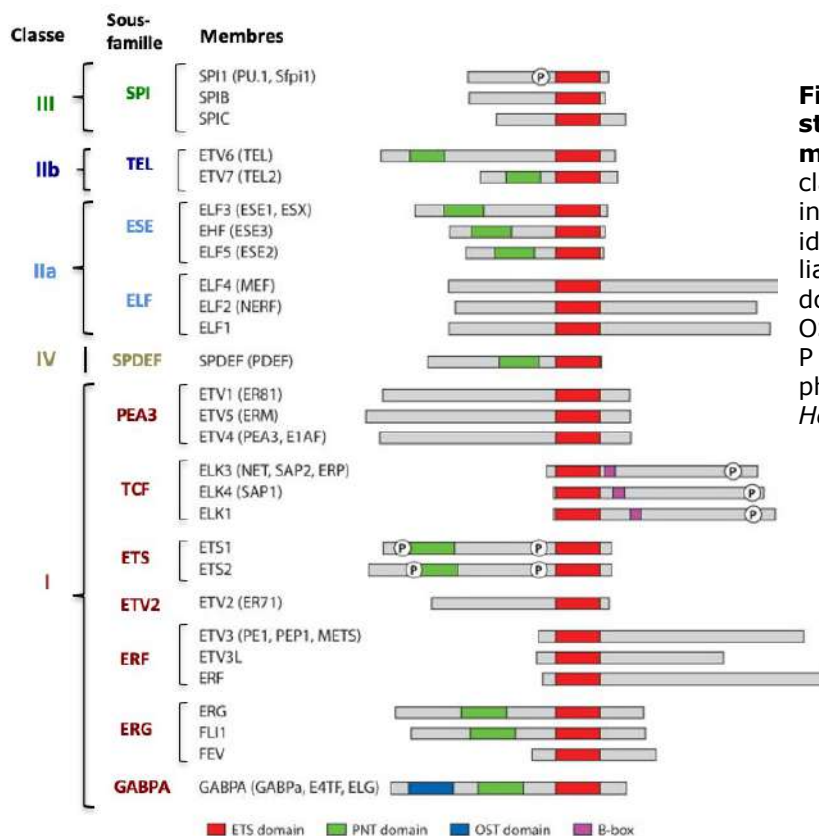


Figure 11 : Domaines structuraux et fonctionnels des membres de la famille ETS. Les classes et les sous-familles sont indiquées à gauche. Les couleurs identifient le domaine ETS de liaison à l'ADN (rouge), le domaine PNT (vert), le domaine OST (bleu) et la B-box (rose). Le P encadré symbolise une région phosphorylée. (*adapté de Hollenhorst et al, (306)*)

Les différentes sous-familles de protéines ETS ont en commun des séquences bien conservées, des fonctions moléculaires et biologiques le plus souvent relativement similaires, et fonctionnent généralement en collaboration avec d'autres FT (308). Les mécanismes de signalisation, les partenaires protéiques et les variations d'expression peuvent différer pour chaque membre, expliquant en partie une spécificité d'action. L'expression des différents facteurs ETS a été recherchée dans divers tissus et des lignées cellulaires, révélant qu'en moyenne plus de la moitié des membres de la famille ETS y était exprimée (314). La spécificité de chaque membre reste néanmoins encore mal connue.

4.2 Liaison à l'ADN

La sélectivité de certains membres de la famille ETS pour une séquence d'ADN particulière a été l'objet de plusieurs études *in vitro* (315–317). L'ensemble des protéines ETS murines et humaines a ainsi été soumis à des études de spécificité de liaison protéique à l'ADN *in vitro* (318).

Les préférences de liaison observées grâce à ces travaux ont permis de séparer les protéines ETS en quatre classes (318) (**Figure 12**). La classe I contient plus de la moitié des membres de la famille ETS (ETS1, ETS2, ETV1, ETV2, ETV3, ETV4, ETV5, ELK1, ELK3, ELK4, ERF, ERG, FEV, FLI1, GABPA ; comprenant les sous-familles ETV2, PEA3, TCF, ETS, ERF, ELG et ERG), et présente une séquence consensus ACC**GG**AAGT. La classe II comprend huit membres (classe IIa : EHF, ELF1, ELF2, ELF3, ELF4, ELF5 ; classe IIb : ETV6, ETV7 ; comprenant les sous-familles TEL, ESE et ELF) et a une séquence consensus qui diffère principalement au niveau du premier nucléotide (CCC**GG**AAGT). La classe III est composée uniquement de la sous-famille SPI (SPI1/PU.1, SPIB, SPIC), dont les membres se lient préférentiellement à une séquence riche en adénine en 5' du noyau GGA (AGAG**GG**AAGT) (316). Cette famille est la plus éloignée des autres familles d'un point de vue structural et de sa séquence cible. La classe IV ne comprend que PDEF, qui a une préférence unique pour une séquence de noyau GGAT au lieu de GGAA (319).

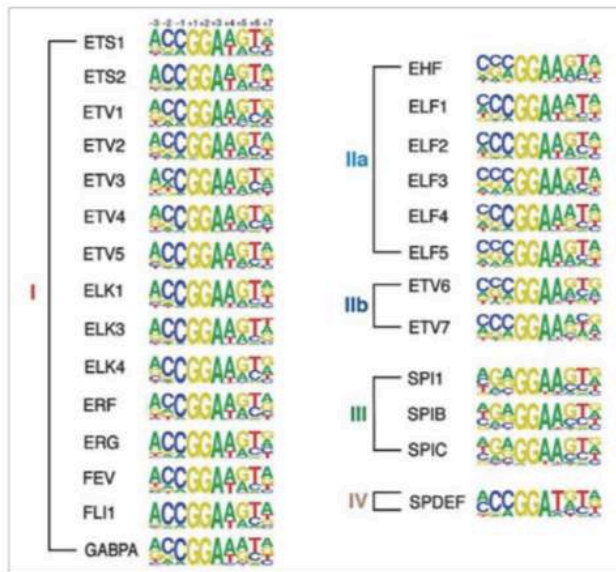


Figure 12 : Spécificités de liaison à l'ADN des différentes classes de facteurs ETS (318)

Selon les études quantitatives visant à évaluer la liaison des différentes protéines ETS à l'ADN, l'affinité pour une séquence donnée varierait d'un facteur 10 environ, mais l'affinité de la même protéine pour différentes séquences fluctuerait de 5 à 80 fois (320). Les différences de préférence de liaison à l'ADN entre les quatre classes de protéines ETS sont dues en partie à des divergences d'acides aminés au niveau de résidus précis en contact avec l'ADN. Pour citer un exemple qui nous intéressera par la suite, la glutamine (Q) en position 226 chez les protéines de classe III est responsable de la liaison à un G en position -2 du motif commun GGA(A/T) tandis que l'acide glutamique (E) présent à cette même position chez les protéines de classe I/IIa favorise la liaison à un C en position -2 (318) (**Figure 13**).

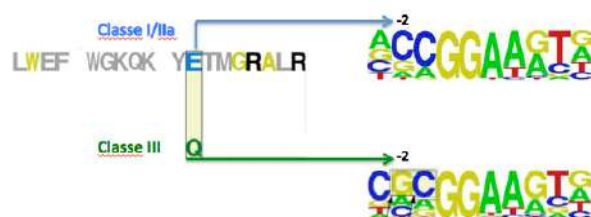


Figure 13 : Spécificité de liaison à l'ADN des classes de facteurs ETS en fonction des résidus d'acides aminés. Le changement de l'acide glutamine (E, en bleu) pour une glutamine (Q, en vert) change la préférence de fixation en -2 du cœur GGAA (motifs à droite). (adapté de Wei et al (318))

Les affinités de liaison des différentes classes de protéines ETS à leur séquence cible et leur influence sur l'activité transcriptionnelle de ces FT restent assez largement méconnues *in vivo*.

La liaison des protéines ETS à l'ADN est influencée par l'interaction avec d'autres FT (320), la liaison de séquences palindromiques sous forme de dimères (321), et la présence de régions auto-inhibitrices autour du domaine ETS (322,323). L'auto-inhibition de la liaison à l'ADN de certaines protéines ETS a pu être identifiée suite à l'observation d'une affinité pour l'ADN plus importante de certaines formes tronquées comparée à celle de la protéine entière. L'exemple le mieux caractérisé de ce phénomène est celui d'ETS1. Il a été découvert par des analyses de délétion (324,325) et sa pertinence biologique confirmée par l'existence de deux variants naturels.

4.3 Interactions protéiques

Une grande partie des FT ETS sont capables d'interagir avec des partenaires de co-régulation (326). Ceci facilite un contrôle combiné de l'expression des gènes et augmente la spécificité d'action des protéines ETS. La place de ces interactions dans l'induction de programmes transcriptionnels spécifiques a été soulignée grâce à l'étude de SPI1 et de son partenaire IRF4 dans la lignée lymphoïde B. Il a ainsi été rapporté que les gènes régulés par SPI1 seul étaient différents de ceux régulés par le complexe SPI1/IRF4 (327).

Les interactions peuvent impliquer d'autres FT, liés ou non à l'ADN. Les interactions avec des FT liés à l'ADN implique le plus souvent le domaine ETS et le domaine de liaison à l'ADN correspondant sur le FT partenaire, rapprochant ainsi leurs sites de liaison à l'ADN respectifs. On peut citer l'exemple de Ets1 et Runx1, dans lequel les interactions réciproques entre leurs domaines respectifs de liaison à l'ADN favorisent la formation d'un complexe tertiaire (328). Dans le cas de la sous-famille TCF et de SRF, les interactions entre les domaines de liaison à l'ADN sont minimales. En effet, l'interaction principale se situe entre une région B-box des TCF et le domaine de liaison à l'ADN de SRF (329). La séquence B-box étant réservée aux membres de la sous-famille TCF, cela implique que ces protéines sont spécifiquement recrutées sur les sites de liaison composites SRF-ETS. Les interactions peuvent également se faire avec des FT ayant peu ou pas de contact avec l'ADN. Dans ce cas, d'autres régions peuvent être impliquées, à part le domaine ETS, comme les régions répétées riches en ankyrine de GABPB qui se lient au domaine ETS et à une hélice α C-terminale de GABPA (330).

Outre les FT spécifiques, les protéines ETS peuvent également recruter des complexes co-répresseurs et co-activateurs épigénétiques. C'est le cas par exemple de CREB-binding protein (CBP), qui a été décrit comme co-activateur de plusieurs protéines ETS, comme SPI1, ETS1, et ETV1 (308,331).

4.4 Sites génomiques de liaisons à l'ADN

Les préférences de fixation des différents facteurs ETS ont pu être étudiées à l'échelle du génome grâce à des techniques d'immunoprécipitation (IP) de la chromatine couplée à l'utilisation de puces ou de séquençage (332–334). Parmi les cibles des protéines ETS, on trouve d'une part les sites de liaison dits « redondants », situés dans les promoteurs proximaux des gènes ubiquitaires (320). En effet, les régions ADN liées de façon redondante se trouvent le plus souvent à une courte distance en amont des TSS et présentent des caractéristiques de promoteurs actifs, comme par exemple la présence d'une marque triméthyle sur l'histone H3K4. Les sites de liaison redondants sont caractérisés par une séquence consensuelle ETS (CCGGAAGT), suggérant que la plupart, voire la totalité des protéines ETS ont la capacité de s'y fixer avec une affinité relativement élevée. La fréquence de cette séquence (retrouvée dans près d'un quart des promoteurs humains), l'association avec des gènes fréquemment exprimés et sa présence à proximité des TSS suggèrent un rôle commun des facteurs ETS pour différentes étapes de la transcription ou pour le maintien d'une chromatine ouverte. D'autre part, il existe des sites de liaison dits "spécifiques" que l'on trouve le plus souvent dans les régions enhancers, associées à des gènes pouvant avoir des fonctions biologiques spécifiques d'un membre de la famille ETS (320). Ces sites spécifiques sont caractérisés par une séquence ETS à faible affinité de liaison, souvent AGGAA, parfois entourée de sites de liaison pour d'autres FT. Ces sites de faible affinité liés par les ETS serviraient de médiateurs pour des fonctions spécifiques des protéines ETS, nécessitant des interactions avec leurs FT partenaires.

4.5 Implication dans les mécanismes de transformation tumorale

Au cours des trois dernières décennies, le rôle des facteurs ETS dans les cancers a été de plus en plus reconnu (305). De nombreux facteurs ETS ont fait l'objet d'études approfondies pour déterminer leur rôle dans le cancer du sein, le cancer colorectal, le cancer gastrique, le cancer des poumons, le cancer des ovaires, le cancer du pancréas, le cancer de la prostate, le sarcome d'Ewing et le glioblastome. Un nombre important d'études ont indiqué que les fonctions aberrantes des ETS peuvent être provoquées par divers mécanismes, notamment les réarrangements chromosomiques, les variations du nombre de copies des gènes, les mutations gain de fonction, la stabilisation des protéines ETS et les mutations intervenant dans les régions régulatrices d'autres gènes. L'exemple le plus éloquent de ce dernier point est la découverte dans certains cancers solides (mélanome, carcinome hépatocellulaire) de mutations au sein du promoteur du gène *TERT*, qui créent des sites de liaison de novo pour les FT ETS, entraînant la réactivation de la télomérase dans les cellules tumorales (335).

Un certain nombre d'études montre que les facteurs ETS, notamment FLI1 (311,336), ERG (337-339) et ETV6 (313,340), sont impliqués dans les hémopathies malignes par le biais d'une expression aberrante et de translocations chromosomiques.

4.6 SPI1/PU.1

4.6.1 Structure et expression

Un des facteurs ETS impliqué dans l'hématopoïèse est codé par le gène *SPI1*, intervenant dans le contrôle de différentes étapes de différenciation, via son niveau d'expression et son interaction avec d'autres FT ou avec des facteurs épigénétiques. Le locus de *Spi1* a été identifié comme un site d'insertion récurrent pour le SFFV (spleen focus-forming virus) dans les érythroleucémies murines de Friend (341). De fait, l'intégration du SFFV dans des régions en amont du gène *Spi1* entraîne l'expression continue de SPI1 dans les proérythroblastes et bloque leur différenciation.

La séquence murine du gène *Spi1* code pour une protéine de 272 ou 266 acides aminés (aa), selon le codon d'initiation utilisé pour la traduction. Elle présente une homologie de séquence de 85% avec la séquence humaine, située sur le chromosome 11, comprenant cinq exons et quatre introns et codant pour une protéine de 264 aa pour un poids moléculaire de 31kDa (342).

SPI1 constitue une famille avec les protéines SPIB et SPIC (43% et 40% d'homologie globale avec SPI1, respectivement) et contient plusieurs domaines fonctionnels. Le domaine ETS correspond au domaine de liaison à l'ADN, et reconnaît les séquences portant le motif central GGAA. Des études thermodynamiques ont révélé une variation de l'affinité de SPI1 pour l'ADN *in vitro* d'une amplitude de 400, selon les séquences entourant le noyau GGA (343).

L'extrémité N-terminale abrite un domaine de transactivation (TAD) qui contient trois régions riches en aa acides et une région riche en glutamine. Le domaine TAD est la cible d'évènements de phosphorylation, intervenant sur de nombreux résidus serines et régulés par p38, JNK et la caséine kinase (344). Ces phosphorylations peuvent moduler la liaison à l'ADN et l'interaction avec les protéines partenaires. Par exemple, la phosphorylation de la sérine 148 de SPI1 par la caséine kinase II favorise les interactions avec son partenaire co-régulateur IRF4 (345).

Le domaine PEST, situé entre les deux domaines précédents, est essentiel pour de nombreuses interactions protéine-protéine. Tous les domaines fonctionnels de SPI1 sont impliqués dans des interactions avec des protéines partenaires mais dans le cadre de la différenciation lymphoïde B, l'interaction entre le domaine PEST et les FT IRF4/8 revêt une importance particulière.

L'expression de SPI1 semble principalement réservée au système hématopoïétique. Au sein des cellules hématopoïétiques, SPI1 présente un modèle d'expression très dynamique (**Figure 14**). Une analyse détaillée de l'activité du promoteur de SPI1 a été rendue possible grâce à plusieurs modèles de souris avec des systèmes rapporteurs dans lesquels une protéine verte fluorescente (GFP) a été introduite (346,347). SPI1 est fortement exprimé dans les cellules hématopoïétiques précoces, y compris les CSH, les CLP, les progéniteurs de granulocytes/macrophages (GMP) et les progéniteurs thymiques précoces (ETP). Les progéniteurs érythroïdes/mégacaryocytes expriment SPI1 à des niveaux élevés dans le foie fœtal mais à des niveaux faibles dans la MO osseuse adulte. C'est dans la lignée myéloïde que la concentration de SPI1 est la plus importante, elle y joue un rôle dans l'induction de son programme transcriptionnel spécifique (299). SPI1 est un arbitre majeur de l'engagement entre les lignées myéloïde et lymphoïde (15,348). L'expression de SPI1 est abaissée au début de la lymphopoïèse B, mais doit être maintenue tout au long de la différenciation des LB (346,347). SPI1 n'est pas exprimée dans les cellules T matures (exception faite de certaines cellules auxiliaires T de type Th2) et les cellules Natural Killer (NK).

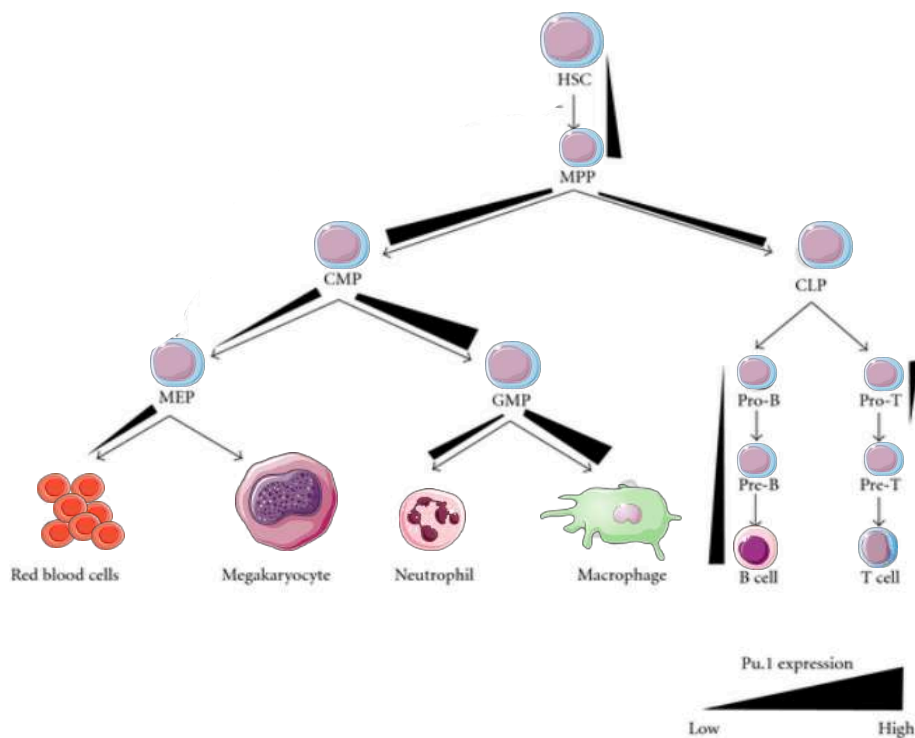


Figure 14 : Expression de SPI1 dans l'hématopoïèse.

Les niveaux de SPI1, lorsqu'ils sont connus, sont représentés par des gradients de couleur noire. Les voies de différenciation sont indiquées par des flèches. Abréviations : HSC, cellule souche hématopoïétique ; MPP, progéniteur multipotent ; CMP, progéniteur myéloïde commun ; CLP, progéniteur lymphoïde commun ; MEP, progéniteur mégacaryocyte-érythroïde ; GMP, progéniteur granulocyte-macrophage. (*Adapté de Mak et al. (349)*).

SPI1 est un régulateur de différenciation clé capable de moduler l'expression de protéines de surface (CD11b, CD16, CD18 et CD64), de cytokines et de leurs récepteurs respectifs (G-CSF, GM-CSF et M-CSF) dans les cellules hématopoïétiques (350). La différence de niveau d'expression de SPI1 entre les progéniteurs myéloïdes et lymphoïdes est responsable d'une activation distincte d'un programme transcriptionnel incluant SPI1, IKAROS mais aussi EGR1 et GFI1 (351–353). L'expression d'IKAROS et de GFI1 entraînent une régulation à la baisse de SPI1, ce qui entraîne à son tour l'activation d'EBF et du programme transcriptionnel de la lignée B. D'autres FT dont E2A, MYB et MEF2C participent aussi à cette régulation (354–356). SPI1 régule également l'expression de FLT3 et IL7R, deux récepteurs cytokiniques clés pour la différenciation myéloïde et lymphoïde (15,357). Le contrôle du niveau d'expression de SPI1 est essentiel au processus d'engagement vers une lignée cellulaire, sa dérégulation pouvant entraîner un blocage de différenciation favorisant l'apparition de leucémie aiguë (358–360).

De nombreuses études ont décrit le contrôle moléculaire de l'expression de SPI1. Un éclairage essentiel sur la régulation de SPI1 a été apporté suite à l'identification d'un élément régulateur en amont (URE, Upstream Regulatory Element) hautement conservé, situé à 14 kb en amont du TSS (361). La suppression de cet élément *in vivo* entraîne des changements marqués dans l'expression de SPI1, avec une forte diminution dans les cellules myéloïdes, mais une expression accrue dans les cellules T CD4-CD8-DN1-DN3 immatures (358,359). Ainsi, l'URE de -14 kb transmet des signaux de régulation positifs ou négatifs, bien que les facteurs impliqués dans cette régulation soient encore largement inconnus. Les protéines TCF (facteur des cellules T), médiateurs de la signalisation Wnt, se lient à cette région et régulent à la baisse l'expression de SPI1 dans les cellules T immatures (359). SPI1 elle-même se lie également à cet élément, car des expériences de mutagenèse sur le site de liaison de SPI1 dans des lignées cellulaires myéloïdes entraînent une diminution drastique de l'activité transcriptionnelle provoquée par l'URE (346,362).

4.6.2 Fonctions

4.6.2.1 Dans les cellules souches et progéniteurs

Iwasaki *et al.* ont montré que SPI1 est exprimée dans les CSH à long terme et à court terme. Ces auteurs ont rapporté que les CSH provenant de souris dont les allèles SPI1 ont été supprimés 2j après la naissance ne pouvaient pas concurrencer efficacement des CSH sauvages pour reconstituer les lignées hématopoïétiques après greffe, ce qui suggère un rôle important de SPI1 dans le maintien des CSH à long terme. De même, les CSH du foie foetal SPI1KO ne contribuent que de façon transitoire au pool de CSH lorsqu'elles sont transplantées chez des receveurs adultes (363). Dans ce sens, une étude de Staber *et al.* montre que les facteurs de la famille RUNX préviennent l'épuisement des CSH grâce au

maintien des niveaux de SPI1 dans ces cellules. La liaison de RUNX au niveau de l'URE - 14kb de SPI1 faciliterait la formation d'une boucle chromosomique entre l'enhancer et le promoteur proximal de SPI1, favorisant ainsi sa transcription (364).

Les transferts adoptifs de CSH qui portent un allèle inductible SPI1 nul, ne contribuent pas aux lignées de lymphocytes B et T lorsque SPI1 est délété, mais permettent en revanche la génération durable de neutrophiles (362). La délétion de SPI1 dans les progéniteurs hématopoïétiques a été décrite comme responsable de la disparition phénotypique des populations des progéniteurs myéloïdes et lymphoïdes communs (362,363), ces derniers résultats doivent toutefois être interprétés avec prudence, car l'expression des marqueurs de surface utilisés pour définir ces populations précoces peut elle-même être régulée par SPI1.

4.6.2.2 Dans la lignée myéloïde

Les macrophages comptent parmi les cellules présentant les niveaux les plus élevés de SPI1, et celui-ci contrôle directement l'expression de gènes impliqués dans la différenciation et la fonction des macrophages (par exemple, l'intégrine CD11b, les récepteurs M-CSF et GM-CSF) (365). En effet, les macrophages, ainsi que les cellules apparentées comme les ostéoclastes et les cellules gliales, ne se développent pas dans les mutants SPI1-nuls. Il a récemment été démontré que la différenciation des macrophages promue par SPI1 dépend de l'induction des régulateurs de transcription Egr-1/2 (Early Growth Response-1/2) et Nab-2 (NGFI-A Binding Protein 2) (352), qui nécessitent des niveaux élevés de SPI1 pour leur expression. Ces facteurs agissent pour renforcer le programme d'expression génique des macrophages et en parallèle réprimer le programme des neutrophiles via la répression de Gfi1 (Growth Factor Independent-1). Les granulocytes sont également absents des fœtus KO pour SPI1, bien que la délétion de SPI1 chez les souris adultes entraîne une augmentation surprenante et spectaculaire du nombre de neutrophiles. Ainsi, SPI1 joue des rôles distincts dans la granulopoïèse fœtale et adulte. Une activité élevée du promoteur de SPI1 semble être associée à l'engagement myéloïde, mais n'est pas nécessaire pour celui-ci. Au cours de l'hématopoïèse fœtale, les progéniteurs myéloïdes sont présents dans les foies foetaux KO pour SPI1, et peuvent être maintenus *ex vivo* avec l'IL3 (366). En outre, les fœtus SPI1-nuls où l'expression de la GFP est sous le contrôle du promoteur de SPI1 présentent des cellules GFP+ (c'est-à-dire qui possèdent une activité transcriptionnelle SPI1) et expriment sélectivement des FT spécifiques de la lignée myéloïde tels que C-EBP α et C-EBP ϵ (346).

4.6.2.3 Dans la lignée lymphoïde B

Le KO constitutif de *Spi1* n'est pas viable au delà de la naissance, les animaux meurent de septicémie en 48h et l'on constate une absence de progéniteurs B chez les fœtus (363,367) ce qui suggère que SPI1 est, entre autres, un régulateur essentiel de la différenciation B.

En revanche, ce rôle semble limité à la lymphopoïèse précoce car la délétion conditionnelle de SPI1 dans les cellules B exprimant CD19 (stade pro-B) est compatible avec un développement normal (363,367). Ces données suggèrent que l'expression de SPI1 est indispensable pour l'initiation de la lymphopoïèse B, entre les stades CSH et CLP et sont confortées par une étude récente (368).

Les principaux modèles murins d'inactivation de *Spi1* et *SpiB* sont rapportés dans la

Table 3.

genotype	Phénotype <i>Spi1</i> -KO	Phénotype DKO <i>Spi1/SpiB</i>	Références
SPI1 CD19-Cre	Pas d'effet sur la différenciation B Réponse à la stimulation LPS et CD40 (production de cellules sécrétrices d'ac de haute affinité et formation de CG) Nombre de cellules sécrétrices d'ac diminué	-	(369)
SPI1 CD19-Cre	Nombre normal de cellules pré-B et de B matures dans la moelle osseuse Réponse à la stimulation IgM et LPS (prolifération)	-	(363)
SPI1 CD19-Cre	Développement des lymphocytes B normal Avec l'âge, augmentation de la proportion de lymphocytes B1 et diminution des lymphocytes B2 (NB : Une augmentation similaire des cellules B1 est observée chez les souris dépourvues de l'enhancer SPI1 -14 kb, qui expriment de faibles niveaux de SPI1 (359))	-	(370)
SPI1/SPIB CD23-Cre	Nombre de cellules sécrétrices d'ac normal Réponse réduite à l'antigène Présence de CG	Augmentation de la proportion et du nombre de plasmocytes dans la rate et dans la moelle osseuse. Absence de réponse à l'antigène Absence de CG	(371)
SPI1/SPIB CD19-Cre	Nombre de lymphocytes B normal Légère diminution de la proportion des B folliculaires Légère augmentation des B marginaux	Diminution du nombre de lymphocytes B Blocage de différenciation au stade pré-B avec le développement de LAL pré-B Forte diminution de la proportion des B folliculaires Forte augmentation des B marginaux	(372)

Table 3 : principaux modèles murins d'inactivation de *Spi1* et *SpiB*

Au total, les effets relativement faibles de la délétion de SPI1 dans les cellules B sont quelque peu surprenants, étant donné le grand nombre de preuves suggérant un rôle critique de ce facteur dans la régulation de l'expression des gènes des cellules B, en association avec les FT IRF4/8. Il est donc possible que le facteur apparenté SPIB, également exprimé dans les cellules B, puisse compenser une partie de la fonction de SPI1. Dans la lignée lymphoïde B, SPI1 et SPIB régulent l'expression de nombreux composants de la voie de signalisation du BCR (Lyn, Syk, Btk, Ptpn6, Card11) et des récepteurs CD40, BAFF-R et TLR (371,373). Un certain nombre d'études se sont donc concentrées sur les effets de l'absence combinée de SPI1 et de SPIB dans la lignée B (**Table 3**).

Le KO de *Spi1* seul dans la lignée B entraîne donc un phénotype modéré, contrairement au DKO *Spi1-SpiB*. Pourtant, les données de Polli *et al.* (369) fournissent des éléments

plaidant contre une redondance fonctionnelle complète entre SPIB et SPI1 suggérée dans d'autres études (374). Premièrement, l'expression de *Ebf1*, un gène cible connu de SPI1, responsable avec E2A de l'engagement dans la lignée B, est réduite dans les cellules pré-B déficientes en SPI1 malgré l'expression continue de SPIB. Deuxièmement, des expériences de band shift réalisées avec une séquence possédant une grande affinité pour SPI1 (oligonucléotide du promoteur SV40), donc potentiellement pour tout facteur capable de s'y substituer fonctionnellement, suggèrent qu'il n'y a pas de liaison de SPIB sur cette séquence dans les cellules B SPI1/CD19. En outre, l'expression de SPIB n'a pas été modifiée en l'absence de SPI1, ce qui suggère que la différenciation ASC accrue observée chez des souris doublement déficientes en SPI1 et IRF8 s'est produite malgré la présence de SPIB (78).

4.6.3 Partenaires protéiques de SPI1

4.6.3.1 Autres facteurs de transcription

La nécessité d'une coopération avec d'autre FT pour la liaison à l'ADN est inversement corrélée à l'affinité des motifs : les motifs de haute affinité sont souvent liés par SPI1 seul, alors que les motifs de faible affinité ne sont liés que lorsque d'autres FT se fixent à proximité.

Dans ce paragraphe, nous nous intéresserons aux partenaires impliqués dans la différenciation lymphoïde B.

Deux des partenaires majeurs de SPI1, IRF4 et IRF8, jouent un rôle essentiel dans différentes étapes de la différenciation des LB. SPI1 peut se lier seul à l'ADN sur les sites canoniques des Ets, ou conjointement avec les FT IRF4 et IRF8 sur des sites composites appelés EICE (Ets-interferon consensus elements) (299,327). Des éléments composites ETS-IRF ont été identifiés dans plusieurs éléments régulateurs majeurs des LB, notamment les gènes IGH et IGL, ainsi qu'IRF4, MEF2C et PAX5 (78,375). Des blocages de différenciation au stade pré-B sont observés dans les cellules B dépourvues de SPI1/IRF4 ou IRF4/IRF8, suggérant qu'un complexe ETS/IRF contrôle la lymphopoïèse B précoce (376).

A un stade plus tardif, SPI1 joue également un rôle dans la différenciation des LB en ASC. Le complexe IRF8/SPI1 contrôle la susceptibilité des cellules B à subir la commutation de classe et la différenciation en plasmocytes en coordonnant l'expression de BCL6 et PAX5 tout en réprimant AID et BLIMP1 (22,73,78). SPI1 est exprimé dans les LB activés avant d'être progressivement régulé à la baisse, sans être complètement éteint dans les plasmocytes (78) (**Figure 4**). Cette régulation à la baisse de SPI1 semble nécessaire sur le plan fonctionnel. En effet, des souris déficientes en miR-155, un microARN ciblant directement la région 3'UTR de *Spi1*, ont une expression augmentée de SPI1. Chez ces

animaux, des défauts de différenciation en ASC ainsi qu'une diminution de la sécrétion d'IgG et d'IgM sont observées (377,378). En temps normal, miR-155 diminue l'expression de SPI1, entraînant une diminution de celle de PAX5 et une augmentation de celle de BLIMP1, permettant ainsi l'initiation du programme des ASC. Dans des modèles murins, le KO de *Spi1* et *Irf8* dans la lignée B entraînait une augmentation de la commutation de classe en IgG1 et IgE et de la proportion d'ASC, suggérant l'importance des éléments composites Ets-Irf dans la différenciation tardive des LB. Etant donné qu'IRF4 joue un rôle de régulateur positif de la différenciation en ASC et de la démonstration qu'IRF8 et IRF4 liaient les mêmes sites dans des gènes cibles primordiaux de la différenciation lymphoïde B chez la souris, un modèle a été proposé dans lequel le niveau de différenciation en ASC serait contrôlé par la concentration de ces deux facteurs (78).

IRF4 se lie faiblement à l'ADN lorsqu'il est seul, mais sa liaison est plus forte en présence de SPI1. L'interaction entre IRF4 et SPI1 a d'abord été caractérisée sur le locus génomique de l'enhancer en 3' du gène Ig kappa (379). Par son domaine ETS, SPI1 se lie à la séquence d'ADN et recrute IRF4 (cet enhancer contient un site composite EICE pour SPI1 et IRF4). Un processus similaire a été signalé avec le site Lambda B1 pour le gène Ig lambda. Cette interaction est cruciale pour le développement correct des cellules B, étant donné qu'elle régule l'expression des chaînes légères des Ig. En l'absence de SPI1, le domaine auto-inhibiteur de l'IRF4 interagit avec son domaine de liaison à l'ADN, l'empêchant de se lier au site. En présence de SPI1, le domaine auto-inhibiteur de IRF4 interagit avec le domaine PEST de SPI1, permettant à IRF4 de se lier au site composite EICE. Il en résulte une interaction coopérative entre les deux protéines (380). Une structure cristalline du complexe SPI1/IRF4/ADN a été obtenue, et il a été proposé que les deux protéines interagissent également par leurs sites respectifs de liaison à l'ADN (381,382).

4.6.3.2 Acteurs épigénétiques

SPI1 peut former des complexes avec un certain nombre de protéines modifiant directement la chromatine ou l'ADN. Il pourrait ainsi agir comme une plateforme recrutant, au niveau de ses gènes cibles, différentes protéines modifiant l'état de la chromatine, ayant des fonctions activatrices ou répressives de la transcription.

La capacité de SPI1 à moduler la structure de la chromatine et modifier le programme transcriptionnel des cellules font qu'il est souvent qualifié de facteur "pionnier" (383). Cependant, il n'est pas un facteur pionnier tout à fait classique (définis par leur capacité à reconnaître leurs sites de liaison dans l'ADN nucléosomal) car il est incapable de se lier à l'ADN méthylé ou lié aux nucléosomes. Malgré cela, l'introduction de SPI1 dans des lignées cellulaires dépourvues d'expression endogène de SPI1 entraîne un remodelage *de novo* important de la chromatine aux sites de liaison de SPI1 et l'induction rapide d'un

programme d'expression de gènes myéloïdes (384). Cette faculté dépend de l'interaction du domaine acide N-terminal de SPI1 avec la famille SWI/SNF, complexes de remodelage de la chromatine dépendant de l'ATP (384).

Rao *et al.* (385) rapportent que SPI1 recrute les protéines TET sur les enhancers, TET augmente l'accessibilité de la chromatine via la déméthylation, ce qui permet à SPI1 de se fixer davantage (boucle de rétrocontrôle positive). Des expériences de co-immunoprécipitation (co-IP) ont montré une interaction de TET2 avec SPI1 dans les cellules B en développement et la perte de TET2/TET3 modifie la liaison de SPI1 sur l'ADN. Des données de ChIPSeq suggèrent que l'activité de TET (mesurée par le dépôt de la marque 5hmC) dans les enhancers est corrélée avec la fixation de SPI1. Au cours de la différenciation monocyttaire, il a été rapporté que SPI1 se lie aux promoteurs dont le niveau de méthylation varie et qu'il recrutait DNMT3b et TET2 sur ces mêmes sites (239). Enfin, des études ont montré que SPI1 interagit physiquement avec CBP et l'empêche d'acétyler ses protéines cibles (FT GATA-1 et protéines nucléaires type histone) (331,386).

4.6.4 Rôle dans la tumorigénèse

D'abord identifié dans les erythroleucémies induites chez la souris du fait de son activation par insertion rétrovirale du virus de Friend (341), ce qui aboutit à sa surexpression, son rôle dans les processus de transformation chez l'Homme a longtemps été recherché.

Contrairement à son rôle d'oncogène dans les cellules érythroïdes, *SPI1* fonctionne comme un suppresseur de tumeur dans les LAM. L'irradiation induit des LAM chez les souris, qui présentent fréquemment une perte de la région du chromosome 2 abritant le gène *Spi1*, ainsi qu'une inactivation fréquente de l'allèle *Spi1* restant par des mutations ponctuelles dans le domaine de liaison à l'ADN (387). La perte d'une copie de *Spi1* est également un événement secondaire récurrent dans les leucémies survenant chez les souris transgéniques exprimant l'oncoprotéine PML-RARA (388), et augmente considérablement la survenue des leucémies dans ce modèle. Une fonction suppresseur de tumeur pour *Spi1* a également été trouvée chez des souris avec des mutations ciblées qui conduisent à une perte complète ou partielle de la fonction de SPI1. Les souris possédant un allèle knockdown (kd) supprimant l'enhancer -14 kb expriment SPI1 à hauteur de 20% des taux sauvages et développent des LAM en 3 à 8 mois (358). De même, des leucémies myéloïdes se développent lorsque *Spi1* est délété chez les souris adultes (362). Globalement, ces études suggèrent que la perte des taux optimaux de SPI1 dans les progéniteurs myéloïdes adultes ou dans les CSH entraîne une leucémogénèse myéloïde chez les souris. Steidl *et al.* (389) ont récemment identifié les régulateurs de transcription c-Jun et JunB (qui appartiennent au complexe AP-1) comme des cibles directes de SPI1 qui sont régulées à la baisse dans les CSH SPI1^{kd/kd}, et ont montré que ces gènes sont essentiels à la fonction

de suppresseur de tumeur de *SPI1*. Alors que leur présence était controversée chez l'homme (390–392), une étude récente a confirmé l'existence de mutations acquises (bien qu'à très faible fréquence), perte de fonction de *SPI1* dans le sous-groupe des LAM avec réarrangement de *MLL* (393). Ces mutations se situent dans le domaine de liaison à l'ADN et sont associées à une diminution du niveau d'expression de gènes cibles connus de *SPI1* (*CSF1R*, *CSF3R* et *AIF1*). Bien que les mutations dans la région codante *SPI1* semblent rares dans la LAM, plusieurs études ont démontré qu'une diminution de l'expression ou de l'activité de *SPI1* par d'autres mécanismes est impliquée dans la pathogenèse. Par exemple, dans les LAM exprimant la protéine de fusion AML1-ETO codée par la translocation t(8;21), un rôle de la protéine de fusion AML1-ETO dans la fixation à *SPI1* et l'inhibition de son activité de transactivation a été démontré (394).

En dehors des LAM, des transcrits de fusion impliquant *SPI1* (STMN1-*SPI1* et TCF7-*SPI1*) ont été identifiés dans les LAL-T pédiatriques de haut risque. Les protéines de fusion en résultant conservaient leur capacité de liaison à l'ADN, ainsi qu'une activité transcriptionnelle. Leurs expressions dans des progéniteurs murins augmentaient la prolifération cellulaire et provoquaient un blocage de la maturation T (395). En effet, ces fusions ont pour conséquence un maintien de l'expression de *SPI1* à un stade de la différenciation T où elle est censée s'éteindre, entraînant le blocage de différenciation et la leucémogénèse, suggérant dans ce cas un gain de fonction.

En revanche, notre étude est la première à avoir identifié l'implication de *SPI1* dans le développement d'hémopathies lymphoïdes B (cf article n°2 p 124).

5 Lymphomes non Hodgkiniens B

5.1 Généralités

Les lymphomes sont des hémopathies malignes, traditionnellement divisés entre les lymphomes de Hodgkin (environ 10 % de tous les lymphomes) et les lymphomes non hodgkinien (LNH). Les LNH représentent un large spectre de maladies qui vont des pathologies les plus indolentes aux plus agressives (**Figure 15**). Environ 85 à 90 % des LNH sont dérivés de cellules B, les autres lymphomes sont dérivés de cellules T ou de cellules NK. Nous nous concentrerons sur deux d'entre eux : le Lymphome B diffus à grandes cellules (Diffuse Large B-cell Lymphoma, DLBCL), et la maladie de Waldenstrom (MW).

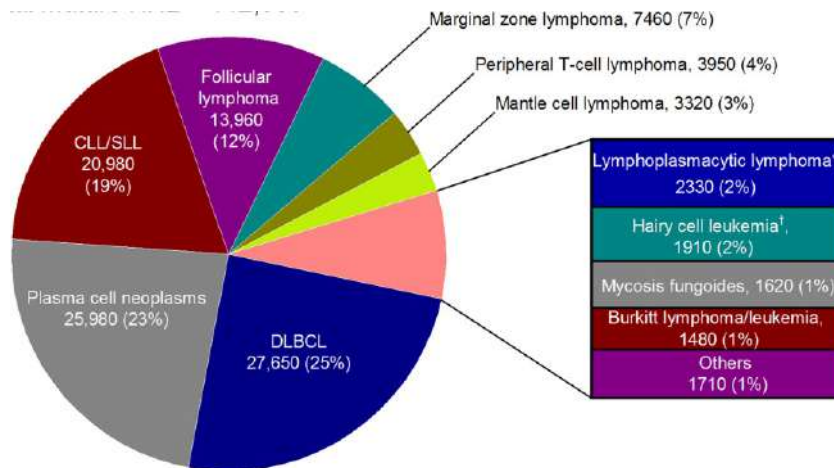


Figure 15 : Fréquences des différents lymphomes non hodgkinien chez l'adulte (133).

CLL : chronic lymphocytic leukaemia
 SLL : small lymphocytic lymphoma
 DLBCL : Diffuse large B-cell lymphoma

5.2 DLBCL

5.2.1 Épidémiologie

Le plus fréquent des LNH est le DLBCL (20-40% des LNH selon les cohortes). L'âge médian des patients au diagnostic est de 70 ans, avec une légère prédominance masculine.

L'étiologie du DLBCL reste inconnue. Ces tumeurs apparaissent généralement *de novo* (DLBCL primaires) mais peuvent également provenir de la transformation d'un lymphome moins agressif (DLBCL secondaire), comme la LLC, le lymphome folliculaire ou le lymphome de la zone marginale. L'immunodéficience constitue un facteur de risque important.

Les patients peuvent présenter une maladie nodale ou extranodale ; jusqu'à 40% des cas sont confinés dans des sites extranodaux au moins au début de la maladie (396). Le site extranodal le plus courant est le tractus gastro-intestinal, mais pratiquement tous les sites extranodaux peuvent être atteints (os, testicules, rate, glandes salivaires, thyroïde, foie, rein, SNC). L'atteinte de la MO par le DLBCL est variable (de 10 à 25% des cas) et sa détection peut être facilitée par l'utilisation de techniques sensibles comme la cytométrie en flux et la biologie moléculaire (397,398). En revanche, l'infiltration du sang périphérique par le DLBCL est rare.

5.2.2 Evaluation morphologique et immunophénotypique

Trois variants morphologiques communs existent : le plus courant est le variant centroblastique. Les centroblastes sont des cellules lymphoïdes de taille moyenne à grande, généralement ovales à rondes, avec des noyaux à chromatine fine. Dans la variante immunoblastique, plus de 90 % des cellules sont des immunoblastes, avec un

seul nucléole central et un cytoplasme basophile abondant. Enfin, le variant anaplasique est caractérisé par des cellules de grande à très grande taille avec des noyaux pléomorphes pouvant ressembler à des cellules de Hodgkin/Reed-Sternberg.

Les cellules néoplasiques expriment généralement des marqueurs pan-B tels que CD19, CD20, CD22, CD79a et PAX5, mais il peut leur manquer un ou plusieurs de ces marqueurs. L'Ig de surface et cytoplasmique (le plus souvent IgM, suivie par IgG et IgA) peut être présente dans 50 à 75% des cas. Le CD30 peut être exprimé dans 10 à 20 % des cas, notamment dans la variante anaplastique (399). Les expressions des CD10, BCL6 et IRF4/MUM1 sont variables. L'indice de prolifération Ki-67 est élevé ; il est généralement bien supérieur à 40% et peut être > 90% dans certains cas (400).

5.2.3 Evaluation clinique et thérapeutique

Les patients présentent généralement une masse tumorale qui grossit rapidement à un ou plusieurs sites nodaux ou extranodaux. De nombreux patients sont asymptomatiques, mais des symptômes B (amaigrissement, asthénie, fièvre, sueurs nocturnes) peuvent être présents. Des symptômes spécifiques de la localisation peuvent être présents, ils dépendent alors fortement du site extranodal concerné (401).

Depuis une trentaine d'année l'utilisation de la combinaison Cyclophosphamide, Doxorubicine, Vincristine et Prednisone (CHOP) constitue la référence du traitement des DLBCL. Une avancée majeure a été apportée par le développement du Rituximab, Ac monoclonal ciblant le CD20. Bien que l'association R-CHOP soit maintenant le traitement standard pour les patients atteints d'un DLBCL, 30% à 50% des patients ne sont pas guéris, selon le stade de la maladie ou l'indice de pronostic. Parmi les patients pour lesquels le traitement R-CHOP échoue, 20% souffrent d'une maladie réfractaire primaire (progression pendant ou juste après le traitement) tandis que 30% rechutent après avoir obtenu une rémission complète (402).

5.2.4 Classification des DLBCL

5.2.4.1 Classification en fonction de la cellule d'origine

On distingue historiquement deux types de DLBCL, basés sur des profils d'expression génique correspondant à différentes étapes de la différenciation des cellules B. Un sous-type exprime des gènes caractéristiques des cellules B du centre germinatif ("germinal center B-like DLBCL", DLBCL-GCB) ; alors que le second exprime des gènes normalement induits lors de l'activation *in vitro* des cellules B du sang périphérique ("activated B-like DLBCL", DLBCL-ABC). (pour revue, (403)) La classification des patients atteints de DLBCL selon la contrepartie physiologique cellulaire s'est révélée prédictive de la survie globale,

l'ABC-DLBCL représentant le sous-groupe de moins bon pronostic (404,405). Avec ce système, 10 à 20 % des DLBCL ne peuvent néanmoins pas être rattachés à une catégorie ou à l'autre et le pronostic clinique demeure hétérogène au sein de chaque sous-type, soulignant les limites de ce mode de classification.

5.2.4.2 Nouvelles classifications en fonction des altérations génétiques et épigénétiques`

En 2018, deux études ont proposé un cadre entièrement nouveau pour la classification des DLBCL en fonction de la génétique des cellules tumorales (302,406). Elles ont chacune suggéré que les DLBCL peuvent être divisées en au moins 4 groupes génétiques, dont certains sont fortement enrichis dans les sous-groupes précédents GC ou ABC. Les quatre sous-types génétiques identifiés sont le MCD (mutations *MYD88L265P* et *CD79B*), BN2 (translocations *BCL6* et les mutations *NOTCH2*), N1 (mutations *NOTCH1*) et EZB (mutations *EZH2* et les translocations *BCL2*) (302). Cette classification est en perpétuelle évolution grâce à l'avènement des techniques de séquençage à haut-débit et d'analyse intégrative de données provenant de multiples plateformes analytiques. Un article récent de Wright *et al.* (407) définit ainsi maintenant 7 sous-types de DLBCL (**Figure 16**). Par rapport aux classifications précédentes, deux nouveaux sous-types sont décrits, ST2 (mutations de *TET2* et *SGK1*) et A53 (mutations/délétions de *TP53*) et le sous-type EZB est divisé en deux en fonction de la présence de réarrangement de *MYC*.

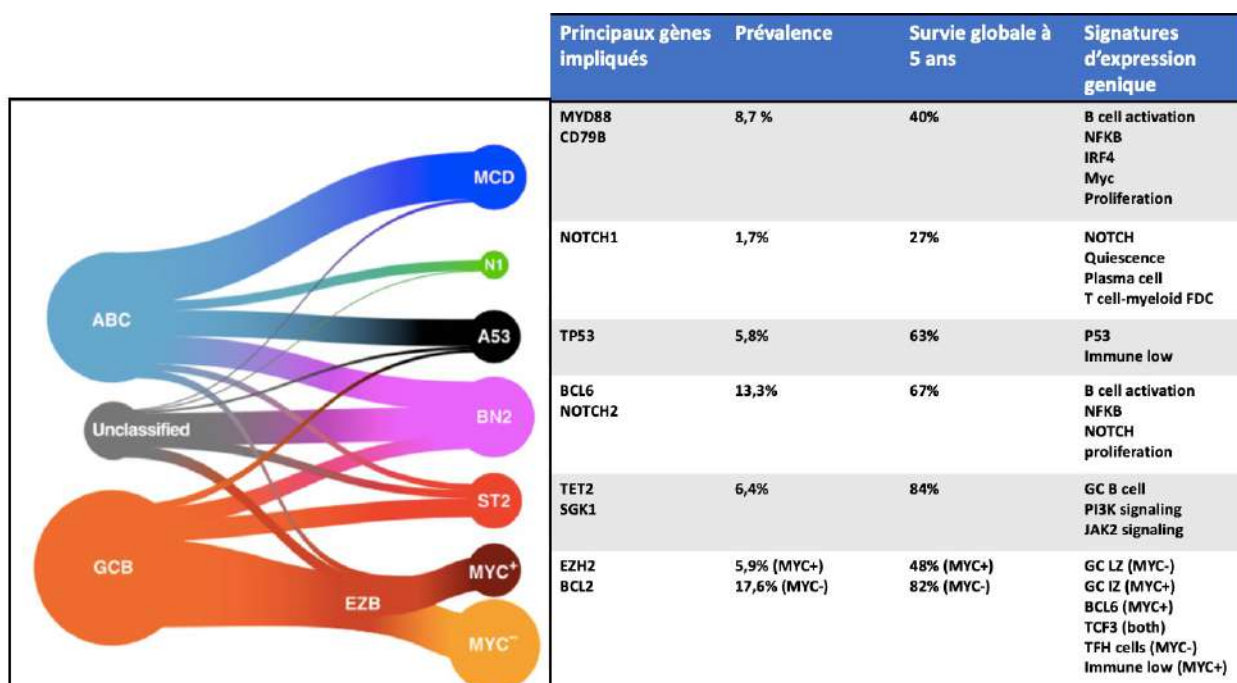


Figure 16 : Classification des DLBCL. Correspondance entre les sous-groupes historiques selon la cellule d'origine et les nouveaux sous-types génétiques (à gauche). Les principaux gènes altérés, les données épidémiologiques et les profils d'expression de chaque sous-type sont présentés (à droite). LZ, light zone ; IZ, intermédiaire zone. (adapté de Wright *et al.* (407))

5.2.5 Principales anomalies cytogénétiques et moléculaires

5.2.5.1 Altérations des modificateurs épigénétiques

Les dernières études s'intéressant à la génomique des DLBCL ont mis en lumière la fréquence élevée des lésions affectant les enzymes de modification des histones et de la chromatine, le plus souvent les acétyltransférases et les méthyltransférases.

Les mutations et délétions inactivant *CREBBP* et, plus rarement, *EP300* sont détectables dans 35% des DLBCL (avec une certaine préférence pour le sous-type GC) et 40% des lymphomes folliculaires (FL) (408,409). Ces phosphoprotéines nucléaires ubiquitaires modifient les résidus de lysine sur les protéines nucléaires histones et non-histones, servant de coactivateurs de transcription pour un grand nombre de FT se liant à l'ADN (410). *CREBBP/EP300* facilitent la transcription par divers mécanismes, notamment l'acétylation ciblée des histones au niveau des résidus lysine (K) (notamment la lysine 27 de l'histone 3) (411,412), l'acétylation des activateurs de transcription (par exemple, le suppresseur de tumeur p53) (413,414) et l'inactivation des répresseurs de transcription par acétylation (par exemple, *BCL6*) (415).

Parmi les candidats nouvellement identifiés, *MLL2* (mixed-lineage leukemia 2, précédemment appelée *MLL4*) est couramment muté dans les LNH, dans 30% des DLBCL et 89% des lymphomes folliculaires (408,409). *MLL2* code pour une histone méthyltransférase hautement conservée et exprimée de manière ubiquitaire, qui contrôle la transcription des gènes en modifiant la lysine 4 de l'histone 3 (H3K4) (416). La triméthylation de H3K4 représente une marque de chromatine active sur le plan de la transcription conservée au cours de l'évolution et est étroitement associée aux promoteurs et aux régions transcrites des gènes actifs, où elle compense l'environnement chromatinien répressif imposé par la méthylation de H3K9 et H3K27 (416).

L'oncogène *EZH2* du groupe polycomb code pour une histone méthyltransférase responsable de la mono-, di- ou triméthylation de la lysine située en position 27 de l'histone H3 (H3K27me1, 2 ou 3). Chez ~22% des patients atteints de GC-DLBCL, des mutations gain de fonction entraînent le remplacement ponctuel d'une tyrosine conservée au cours de l'évolution (Tyr641) (417), ce qui augmente les capacités catalytiques d'*EZH2* et aboutit *in fine* à une augmentation des niveaux de H3K27me3 (418).

5.2.5.2 Lésions génétiques conduisant à une activité constitutive NF-κB et à une signalisation BCR active chronique

Le rôle essentiel de NF-κB dans la pathogénie de l'ABC-DLBCL a été révélé en observant que la signature transcriptionnelle de ces tumeurs est considérablement enrichie en gènes

cibles de NF- κ B, et que l'activité de NF- κ B est nécessaire à leur prolifération et à leur survie (419). Par la suite, un certain nombre d'études ont montré la présence d'altérations structurales qui affectent une variété de voies dont le dénominateur commun est la capacité à induire une activation constitutive de NF- κ B, y compris celles en aval du BCR, des TLR et du CD40.

- Mutations de *CD79A/B*

Une signalisation BCR tonique est nécessaire pour la survie des cellules B (420) (voir paragraphe 2.3.1), cependant, les ABC-DLBCL dépendent d'une forme chroniquement active de la signalisation BCR, qui est soutenue dans plus de 20 % des cas par des mutations de *CD79B* et *CD79A* (154). Le knockdown de plusieurs sous-unités proximales et distales du BCR est spécifiquement délétère pour l'ABC-DLBCL, ce qui justifie le développement de thérapies ciblées visant à inhiber la signalisation du BCR dans ce sous-type de lymphome (154,155). En effet, l'utilisation de l'ibrutinib, un inhibiteur de BTK, s'avère très efficace chez les patients atteints d'ABC-DLBCL porteurs de mutations *CD79A/B* (421).

- Mutations de *MYD88*

Environ 30% des patients atteints d'ABC-DLBCL présentent une mutation récurrente L265P dans le domaine intracellulaire de cette protéine adaptatrice en aval des récepteurs de la famille des TLR et du récepteur à l'IL1 (IL1-R), qui a le potentiel d'activer les voies NF- κ B et JAK/STAT3 (165). Le besoin de MYD88 pour la survie des ABC-DLBCL implique un rôle pathogène des TLR dans ce sous-type de lymphome (voir paragraphe 5.3.4).

- Altérations des régulateurs négatifs et positifs de NF- κ B

Dans près de 30% des cas de ABC-DLBCL (et une fraction plus faible de GCB-DLBCL), des mutations et/ou délétions inactivent de façon biallélique le gène *TNFAIP3* codant pour la protéine A20 qui intervient dans la régulation négative de la voie de signalisation du NF- κ B. La perte d'A20 pourrait donc favoriser la lymphomagenèse, au moins en partie, en induisant une signalisation NF- κ B prolongée de manière inappropriée (422,423).

5.2.5.3 Dérégulation de l'activité de *BCL6*

BCL6 est un répresseur transcriptionnel qui se lie à des séquences spécifiques dans le promoteur de ses gènes cibles et module leur expression via l'interaction avec des complexes de co-répresseurs. *BCL6* contrôle le programme biologique des cellules BGC en

réprimant un large ensemble de gènes impliqués dans de multiples voies de signalisation, y compris la signalisation BCR et CD40, l'activation des cellules B par les LT, la détection et la réponse aux dommages de l'ADN (par la répression directe de *TP53* et de la kinase ATR), l'induction de l'apoptose (par la répression de *BCL2*), et la différenciation en plasmocytes (par la répression de *BLIMP1*) (47,62,85). Ces activités maintiennent le statut prolifératif des cellules BGC tout en leur permettant de tolérer les cassures de l'ADN entraînées par les SHM et le CSR. De plus, *BCL6* empêche l'activation prématurée des cellules B et leur sortie du CG avant la sélection des clones produisant des Ac de haute affinité (pour revue, (61)).

Des réarrangements chromosomiques du locus *BCL6* sont présents chez ~35% des patients atteints de DLBCL, avec des fréquences plus élevées dans le sous-type ABC-DLBCL (61,403). Ces événements de recombinaison juxtaposent le domaine codant intact de *BCL6* à des séquences hétérologues dérivées de plus de 20 partenaires chromosomiques alternatifs. Les translocations impliquant *BCL6* empêchent la régulation à la baisse de son expression, nécessaire à la différenciation en cellules B post-CG (61).

En plus des translocations, les séquences en 5' de *BCL6* peuvent être les cibles d'hypermuation dans près de 75% des DLBCL (424,425). Ce phénomène reflète en grande partie l'activité du mécanisme physiologique SHM opérant dans les cellules BGC (425,426).

5.2.5.4 Perturbation de la voie de différenciation terminale

Dans les cellules BGC en phase de différenciation terminale, *BCL6* doit être réprimé afin de permettre l'expression de son gène cible, régulateur clé de la différenciation plasmocytaire, *BLIMP1* (65). Cet axe de régulation est perturbé dans la majorité des ABC-DLBCL en raison de lésions génétiques mutuellement exclusives qui inactivent *BLIMP1* soit directement - par des mutations et délétions (25 % des cas) - soit indirectement - par sa répression suite à la dérégulation de *BCL6* (25 % des cas) (427,428). Les translocations de *BCL6* et l'inactivation de *BLIMP1* représentent des mécanismes oncogéniques convergeant sur la même voie pour favoriser la lymphomagenèse en bloquant la différenciation terminale. En cohérence avec cela, la délétion conditionnelle de *Blimp1* dans le CG conduit à des maladies lymphoprolifératives apparentées à l'ABC-DLBCL humain (427).

5.3 Maladie de Waldenstrom

5.3.1 Épidémiologie

La MW est un lymphome indolent rare représentant 1 à 2% des hémopathies malignes. Elle affecte plus souvent les hommes que les femmes avec un sex ratio 2:1 et un âge médian au diagnostic de 65 ans. La MW est un syndrome lymphoprolifératif B, caractérisé

par une gammopathie monoclonale IgM et une infiltration de la MO par un lymphome lymphoplasmocytaire (LPL) (429). Ce dernier est défini par la présence de lymphocytes, de plasmocytes et de lymphoplasmocytes tumoraux (430). Une protéine monoclonale IgM, quelle que soit sa concentration, ne peut pas être considérée comme indicative d'un diagnostic de MW, car elle peut être trouvée chez des patients atteints d'autres syndromes lymphoprolifératifs à cellules B. Une évaluation morphologique et immunophénotypique détaillée de la MO en parallèle d'une évaluation clinique approfondie sont donc nécessaires si l'on veut établir un diagnostic définitif de MW.

5.3.2 Evaluation morphologique et immunophénotypique

Le myélogramme montre en général une MO de richesse normale à diminuée, infiltrée par un clone tumoral constitué de cellules B présentant un continuum de différenciation allant du lymphocyte au plasmocyte en passant par les cellules lymphoplasmocytaires (431). Les mastocytes sont augmentés dans environ un quart des cas (432).

Une étude récente a montré que la majorité des patients atteints de MW ont un immunophénotype caractéristique. En plus de l'expression des marqueurs pan-B CD19, CD20 et CD79, environ 90 % des cas ont un phénotype CD22faible CD25+ CD27+ IgM+ mais sans expression de CD5, CD10, CD11c, CD23 et CD103 (433). On retrouve également une monotypie dans l'expression de la chaîne légère κ ou λ , avec une large prédominance des cas IgM+ κ +. Le nombre de plasmocytes chez les patients atteints de MW est généralement normal, mais l'immunophénotype de ces cellules (CD38++CD19+/-CD56-CD45++CD20+) ne correspond ni tout à fait à des plasmocytes normaux (CD20-) ni à ceux retrouvés dans le myélome (CD19-CD56+) (434).

5.3.3 Evaluation clinique et thérapeutique

La présentation clinique des patients est variable, allant d'une situation indolente, chez environ un tiers des patients, à des symptômes plus ou moins sévères ou invalidants. Ces symptômes sont associés d'une part aux conséquences de l'hypersecretion de l'IgM monoclonale (notamment hyperviscosité, organomégalie, manifestations auto-immunes de type neuropathie périphérique, cryoglobulinémie et maladie des agglutinines froides) et d'autre part à l'infiltration tumorale systématiquement médullaire, responsable de cytopénies (la plus fréquente étant l'anémie). L'envahissement extramédullaire symptomatique est rare au moment du diagnostic dans la MW. Les adénopathies sont présentes chez environ 15 % des patients au diagnostic mais peuvent être plus fréquentes lorsque la maladie progresse. La médiane de survie des patients varie de 5 à 10 ans. Un tiers d'entre eux présentent une maladie plus agressive avec une médiane de survie de

moins de 4 ans. Une minorité de patients (10% à 15%) peuvent avoir une transformation en lymphome agressif (DLBCL) (435,436).

Il n'y a pas de thérapie standard pour la MW symptomatique. Plusieurs essais prospectifs ont démontré l'efficacité de divers médicaments seuls ou en association, mais il est difficile de les comparer sur la base des taux de réponse. Néanmoins, l'immunothérapie est aujourd'hui à la base du traitement pour la majorité des patients, des progrès significatifs ayant été obtenus depuis l'avènement du rituximab, Ac monoclonal chimérique qui cible l'Ag CD20, exprimé sur les cellules B, y compris les cellules MW (437). En combinaison avec de la chimiothérapie standard, il constitue l'un des traitements les plus efficaces dans la MW et est uniformément recommandé dans les directives nationales et internationales pour le traitement de première ligne comme de rattrapage (438). Des thérapies plus récentes ont également leur application dans cette pathologie, comme les inhibiteurs de BTK (ibrutinib) ou les inhibiteurs de Bcl2 (venetoclax).

5.3.4 Principales anomalies cytogénétiques et moléculaires

L'analyse du statut mutationnel des chaînes lourdes des Ig (IGHV) a rapporté une hypermutation des gènes VH dans environ 90 % des cas de MW (439,440). L'étude du répertoire montre un biais d'utilisation avec une surexpression de la famille VH3 et en particulier du VH3-23. La mise en évidence d'un CDR3 court et de marqueurs d'une pression de sélection a été aussi décrite (439). La diversité intra-clonale n'est pas fréquente dans cette pathologie (441,442). Bien que rare, la CSR a néanmoins été décrite dans les cellules MW (443). Plusieurs théories pourraient expliquer l'apparente incapacité des cellules de MW à exécuter la CSR (442,444). L'une d'entre elles pourrait être des délétions au niveau de la région switch μ , comme cela semble se produire dans LLC (445). Une autre possibilité résiderait dans un déclenchement défectueux du processus de commutation, mais avec une machinerie de CSR apparemment normale (446).

L'anomalie cytogénétique la plus fréquente dans la MW est la délétion du bras long du chromosome 6, de 31 à 54% selon les séries. Cette anomalie n'est pas spécifique de la MW, on retrouve en effet des délétions du 6q dans d'autres hémopathies B comme la LLC, les lymphomes de la zone marginale et le myélome multiple, mais avec une fréquence moindre (447,448). Il existe 4 zones minimales de délétion sur le bras long du chromosome 6. Les 2 principales régions contiennent toutes deux un gène suppresseur de tumeur, *BLIMP1* en 6q21 et *TNFAIP3* en 6q23, qui pourraient être impliqués dans la physiopathologie de ce lymphome (449). Les études de CGH-array ont montré que dans 17% des cas, il existe un gain du 6p associé à la perte du 6q (449). Une trisomie 4 est retrouvée dans 10 à 20% des cas. Elle semble être l'anomalie récurrente la plus spécifique. Elle n'est pas décrite dans les lymphomes de la zone marginale. Cette trisomie est associée

à un taux élevé d'IgM, de CRP et de β 2 microglobuline mais n'a pas d'impact connu sur l'évolution de la maladie (450,451). On retrouve également des anomalies décrites classiquement dans la LLC : une délétion du 11q22 impliquant *ATM* dans 8% des cas, une trisomie 12 dans 3% des cas. La délétion du 13q14, décrite dans 50% des LLC et dans le myélome multiple, est retrouvée dans 10% des cas de LPL/MW par CGH array (449). La délétion du 17p, impliquant *TP53* est observée dans 5 à 10% des cas, elle est associée à une progression de la maladie avec une survie sans progression plus courte (451).

- Mutation de MYD88

Localisé chez l'Homme en 3p22.2, le gène *MYD88* code pour une protéine adaptatrice impliquée dans la réponse de la voie des TLR et de l'IL1-R. La structure de MYD88 comprend un domaine DD (Death Domain), dans sa partie N-terminale, impliqué dans la signalisation en aval, un domaine intermédiaire de liaison, impliqué dans la signalisation intracellulaire avec le DD et le domaine TIR en C terminal qui interagit avec les TLR ou le récepteur à IL1.

Lors de l'activation d'un TLR ou de l'IL1-R par son ligand, MYD88 est recrutée et se dimérise. La dimérisation de MYD88 est nécessaire à l'assemblage du myddosome, structure composée de dimères de MYD88 qui recrutent et activent IRAK4, qui à son tour recrute et active IRAK1/2 (452).

L'effecteur final du myddosome est l'activation du FT NF- κ B via l'implication de TRAF6 (Tumor Necrosis Factor Receptor-Associated factor 6), TAK1 (TGF β Activated Kinase 1) et le complexe hétérotrimérique I κ K (Inhibitor of Kappa-B Kinase, composé des sous-unités I κ K α , I κ K β et I κ K γ alias NEMO). NF- κ B régule entre autres le développement, la maturation et la survie des LB. Dans la cellule au repos, les sous-unités p50 et p65 de NF- κ B sont séquestrées dans le cytoplasme par la protéine I κ B (Inhibitor of kappa B). Lorsque le myddosome est activé, I κ B est phosphorylé puis dégradé par le protéasome libérant les sous-unités p50 et p65 du complexe NF κ B. Celles-ci vont pénétrer le noyau, activant ainsi la transcription de gènes cibles (**Figure 17**).

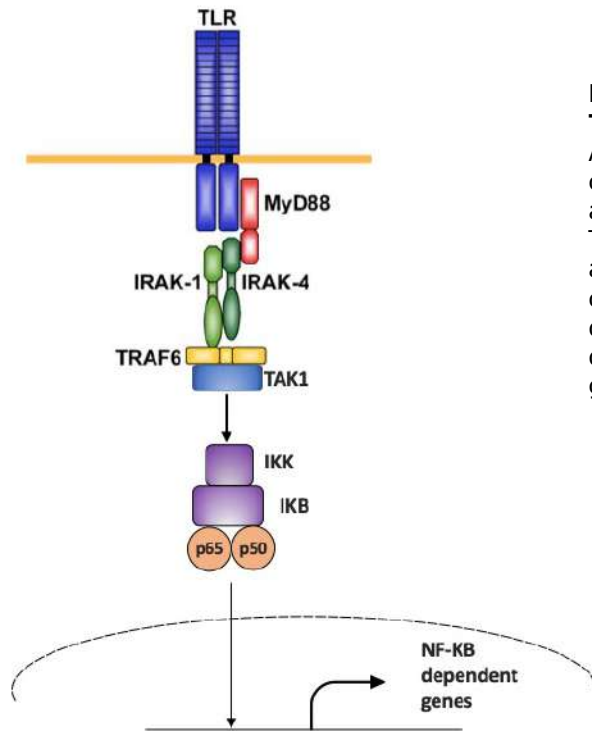


Figure 17 : Voie de signalisation en aval du TLR, faisant intervenir MYD88.

Après stimulation des TLR, MYD88 est recruté comme homodimère et forme un complexe avec IRAK4, conduisant à l'activation d'IRAK1. TRAF6 est ensuite activé par IRAK1, et interagit avec TAK1, qui à son tour phosphoryle le complexe IKK, permettant de libérer les dimères p65 et p50 de NF-κB, qui s'accumulent dans le noyau où ils activent la transcription de gènes cibles (adapté de Takeda et al (164)).

Il existe plusieurs mutations de *MYD88*, toutes affectant le domaine TIR de la protéine. En 2012, les travaux de Treon *et al.* ont permis d'identifier une mutation somatique récurrente L265P du gène *MYD88* chez des patients atteints de MW. Sa fréquence s'élève à 95% des cas dans cette pathologie (453,454). Elle est également détectée chez près de la moitié des patients atteints de gammopathie monoclonale IgM de signification indéterminée (MGUS à IgM), et sa présence, ainsi que le niveau d'expression, semblent associés à la progression maligne vers la MW (455). En outre, la mutation *MYD88* L265P a également été rapportée dans les DLBCL de type ABC (14-29%), le lymphome primitif du système nerveux central (33%), le lymphome du MALT (9%), et la LLC (3%) (167).

Cette mutation acquise récurrente est donc un évènement moléculaire majeur dans la pathogenèse de la MW mais n'en est pas spécifique. Bien que MYD88 soit impliqué dans la signalisation de la voie des TLR de nombreuses cellules hématopoïétiques (granulocytes, macrophages, lymphocytes B et T), la mutation n'est retrouvée, parmi les hémopathies, que dans les LNH B. Une expression augmentée de MYD88 est associée à un mauvais pronostic dans divers cancers solides comme le cancer colorectal, de l'ovaire, ou dans le carcinome hépatocellulaire (456-458).

La mutation L265P se situe sur le domaine TIR de MYD88 et modifie la structure tertiaire de la protéine (459). La conséquence est la formation du myddosome indépendamment de la présence de ligands pour les TLR. L'impact des mutations de *MYD88* dans la survie des cellules tumorales de DLBCL-ABC et de MW a été établi par des expériences de knockdown et/ou de surexpression de MYD88.

Ngo *et al.* (165) avaient constaté qu'en réprimant la signalisation de MYD88 dans les DLBCL-ABC, la voie NF- κ B était inhibée. La survie des cellules portant la mutation L265P était favorisée par le mutant par rapport à *MYD88* sauvage, ce qui démontre que le caractère gain de fonction de la mutation L265P. De plus, l'inhibition du signal de MYD88 par un inhibiteur de kinase IRAK1/4 entraînait une diminution de sécrétion des IL6 et IL10 ainsi que de la phosphorylation de STAT3.

Treon *et al.* (166) obtenaient des résultats similaires sur des cellules de MW mutées *MYD88* L265P et cultivées en présence d'un inhibiteur d'homodimérisation ou d'un inhibiteur IRAK1/4 kinase. Ainsi, une activation constitutive de BTK, IRAK1/4, TAK1 et de la voie JAK/STAT3 est observée dans des modèles cellulaires de MW *MYD88L265P* (166,167,453,460,461).

Le mutant L265P favorise donc la survie des cellules tumorales en assemblant spontanément et en stabilisant le myddosome, activant ainsi la signalisation NF- κ B, la phosphorylation de STAT3 par JAK et la sécrétion d'IL6, d'IL10 et d'interféron- β (165-167).

Un modèle murin avec l'expression conditionnelle de la mutation de *Myd88* L252P (homologue murin de la mutation *MYD88L265P* retrouvée chez l'Homme) a été développé par Knittel *et al.* (462). Les souris développent un syndrome lymphoprolifératif B avec une transformation en lymphome agressif dans environ un tiers des cas. L'expression de la forme mutée *MYD88L265P* entraîne un blocage de la différenciation B après le passage dans le CG. Le phénotype observé montre principalement une infiltration tumorale des ganglions et de la rate, la MO ne présentant pas de signes d'envahissement. C'est l'association avec une hyperexpression de BCL2 qui conduit à un phénotype systématique de DLBCL-ABC. La mutation activatrice *MYD88L265P* pourrait donc être considérée un évènement initiateur mais pas un évènement transformant dans la MW. Ceci a été confirmé par une autre étude (463), montrant que chez la souris, la mutation *Myd88L265P* favorise le développement d'un syndrome lymphoprolifératif B de bas grade, non clonal, présentant des ressemblances clinico-pathologiques avec la MW humaine. La transformation histologique en DLBCL est observée chez une minorité des animaux vieillissants (18%), et est associée à l'acquisition d'altérations génétiques secondaires fréquemment retrouvées dans les DLBCL humains de novo ainsi que dans les cas de transformation de MW (*Malt1, Klf2, Gna13, Dusp2, Pik3c2g, Pdgfrb et Pim1*).

- Mutation de *CXCR4*

Des mutations constitutionnelles de ce gène avaient été précédemment découvertes dans le déficit immunitaire de WHIM (Warts, Hypogammaglobulinemia, Infections and Myelokathexis). Les mutations de *CXCR4* à l'origine de la pathogénèse de ce syndrome entraînent un blocage de son internalisation après stimulation par le ligand (CXCL12) dans

les cellules myéloïdes. Le récepteur CXCR4 est alors constamment activé, retardant ainsi la libération des polynucléaires neutrophiles de la MO vers le sang et favorisant le «homing» vers la MO. Cette mutation est retrouvée dans environ un tiers des MW et quasiment systématiquement associée à la mutation *MYD88* L265P (454,464).

5.3.5 Physiopathologie / Origine cellulaire

Définir la contrepartie physiologique d'une prolifération tumorale est un élément important pour sa classification. Il semble important de préciser ici pour la suite du propos la nuance entre « cellule d'origine » et « contrepartie physiologique ou équivalent normal ». En effet, la véritable cellule d'origine d'un lymphome (celle dans laquelle va survenir le premier événement oncogénique) peut être une cellule à un stade de différenciation plus précoce que son équivalent normal. Cela peut se produire si l'événement oncogénique initiateur se produit tôt dans l'hématopoïèse, mais permet à la cellule de continuer sa différenciation avant que des événements secondaires ne se produisent et aboutissent à la transformation tumorale.

La contrepartie physiologique est désormais connue pour beaucoup d'hémopathies malignes, mais certaines, comme la MW, ont une origine encore controversée. Or cela revêt une importante capitale si l'on veut comprendre les mécanismes physiopathologiques aboutissant au développement de l'hémopathie et envisager des thérapeutiques ciblées. Un certain nombre d'éléments peuvent être considérés pour appréhender cette question : les caractéristiques morphologiques, immunophénotypiques, le statut mutationnel IGHV, le transcriptome et le profil épigénétique des cellules tumorales.

Bien qu'il soit clair que la cellule MW provienne d'une cellule B hautement différenciée, la population tumorale est morphologiquement hétérogène, composée de petits LB, de lymphoplasmocytes et de plasmocytes, ce qui permet d'envisager que plusieurs stades de différenciation des LB puissent correspondre à la cellule initiatrice de la MW.

Le profil d'expression immunophénotypique des cellules MW est assez comparable à l'expression des cellules B mémoires normales, sauf pour l'expression du CD200 (positive dans la MW et négative dans les B mémoires). Les cellules B mémoires ont traditionnellement été identifiées dans le sang périphérique sur la base de l'expression de CD27, un marqueur acquis au cours de la réaction du CG. En plus du CD27, différents profils d'expression des molécules impliquées dans les interactions entre cellules B-T sont détectés dans les cellules B mémoires par rapport aux cellules B naïves, comme par exemple des niveaux élevés de récepteurs à l'IL2 (CD25), de molécules pro-apoptotiques Fas (CD95), et une diminution de l'expression du CD23 et CD200. Le CD27 associé aux B mémoires est exprimé de manière hétérogène (de négatif à dim-positif) chez plus de la

moitié des patients atteints de MW (434), suggérant une origine différente pour cette catégorie de patients, peut être issu d'un autre sous-type de B mémoire. En effet, des B mémoires CD27⁻ ont été décrits chez l'Homme, ils sont porteurs de peu de mutations somatiques dans leur région V et sont augmentés chez les personnes âgées, ce qui a conduit à proposer qu'ils puissent représenter un pool de B mémoires "épuisés" (465–467).

Concernant le statut mutationnel IGHV, la MW est associée à une surreprésentation de la famille IGHV3-23 et à un taux de mutation somatique élevé, avec des caractéristiques de pression de sélection et des régions CDR3 courtes (439). Ces caractéristiques semblent indiquer que les cellules B de MW sont d'origine post-CG, ayant été soumises à une sélection antigénique T-dépendante, activées pour être transformées en plasmocytes. L'autre possibilité est de les considérer un peu plus précoces, comme des cellules du CG. Cependant, la MW est caractérisée par l'absence de CSR (443,446) (les cellules expriment une IgM de surface) et l'absence de variation intraclonale (440,444), ce qui les situerait juste entre ces deux origines potentielles.

Cependant, certaines cellules MW sont CD27⁻, dépourvues d'hypermutation somatique, ou ont subi le CSR. Ainsi, la plupart des données soutiennent une origine de cellule B mémoire pour la MW, mais une fraction des cas pourrait avoir une origine différente.

OBJECTIFS DU TRAVAIL

Le développement des cellules B est un processus multi-étapes étroitement régulé au niveau transcriptionnel. Un certain nombre de FT indispensables à la différenciation des cellules B sont maintenant connus. D'une importance capitale pour assurer les programmes transcriptionnels du CG et des plasmocytes, ces FT doivent avant toute chose pouvoir se lier à leurs loci cibles. Cette liaison est dépendante de l'accessibilité de la chromatine, elle-même régulée par des modificateurs épigénétiques. Ainsi, la différenciation des LB repose sur une combinaison de régulateurs moléculaires, modulés de façon dynamique au cours de la réponse immunitaire. Depuis la cellule souche jusqu'au plasmocyte, l'acquisition de mutations somatiques par la cellule B peut venir perturber les programmes transcriptionnels physiologiques ou impacter les régulateurs épigénétiques de la structure chromatinienne et de l'expression génique. Les lymphomes B sont les cancers hématopoiétiques les plus fréquents et correspondent à un clone de LB ayant échappé au contrôle physiologique de sa différenciation. Une étude récente a identifié 150 gènes mutés de façon récurrente dans les DLBCL et indique que plus de la moitié d'entre eux codent pour des modificateurs épigénétiques ou des FT (122).

Le but de ce travail a été d'étudier la différenciation terminale des LB et l'impact de mutations altérant sa régulation transcriptionnelle ou épigénétique. Dans un premier temps, je me suis intéressée à la dioxygénase Tet2, enzyme responsable de l'oxydation des cytosines méthylées, dont des mutations perte de fonction sont décrites dans près de 10% des DLBCL de type ABC. De ce projet a découlé l'étude du rôle de Tet3 dans la différenciation B terminale, autre protéine de la famille Tet exprimée dans la lignée B, pour laquelle un certain nombre de données laissent penser qu'elle pouvait également jouer un rôle dans la différenciation B. La deuxième partie du travail s'articule autour de la caractérisation d'une mutation du FT SPI1, détectée chez 6% des patients ayant une maladie de Waldenstrom (MW) et engendrant un pronostic moins favorable. Ces malades étant par ailleurs quasiment tous porteur d'une mutation activatrice du gène *MYD88*, j'ai donc par la suite cherché à identifier les bases moléculaires d'une coopération entre les altérations de *SPI1* et de *MYD88*. Les différents acteurs étudiés au sein de ce projet de thèse interviennent soit dans la régulation épigénétique, soit dans la régulation transcriptionnelle de la différenciation B, les deux étant étroitement liés. De fait, SPI1 a été rapporté comme capable de recruter Tet2 au niveau d'*enhancers* spécifiques dans les progéniteurs B. Ces acteurs sont intéressants car ils sont l'objet de mutations dans les lymphomes B, pathologies dérivées de la réaction du CG. Bien qu'il ne s'agisse pas de la même tumeur, le DLBCL type ABC et la MW proviennent de cellules B activées ayant subi la réaction du CG. Les cellules d'origine présumées de la MW et de la DLBCL ABC devraient

donc être proches dans la différenciation des cellules B. De plus, il existe un chevauchement significatif entre les mutations trouvées dans ces deux pathologies (*CD79A/B*, *TNFAIP3*, *MYD88*, *TP53*, *ARID1A/B*), ce qui laisse penser qu'elles reposent - au moins partiellement - sur les mêmes voies oncogéniques.

Grâce à des modélisations *in vivo* et *in vitro* de la réaction du CG, ou de l'activation des LB suite à une stimulation antigénique de façon plus générale, nous espérons apporter des pistes permettant de mieux comprendre les étapes clés de cette différenciation. Cela est important d'un point de vue thérapeutique dans la mesure où leur ciblage peut conduire au développement de nouveaux axes de traitement pour les patients.

PARTIE 1 : Enzymes TET

ARTICLE n°1 : TET2 Deficiency Causes Germinal Center Hyperplasia, Impairs Plasma Cell Differentiation, and Promotes B-cell Lymphomagenesis.

1. Résumé de l'article

Le gène *TET2* code pour une dioxygénase capable d'hydroxyler les cytosines méthylées (mC) en hydroxylméthylcytosines (hmC), participant ainsi à la deméthylation de l'ADN. Les hmC sont associées aux éléments cis-régulateur de la transcription (enhancers) et peuvent affecter l'expression des gènes. Des mutations acquises inactivant *TET2* sont observées dans une grande partie des hématopoïèses clonales (CHIP), 15 % des hémopathies myéloïdes et environ 10% des proliférations lymphoïdes matures de type T et B. L'absence de *TET2* dans les étapes précoces de la différenciation hématopoïétique confère un avantage compétitif aux cellules mutées. Néanmoins, l'inactivation de *TET2* ne suffit pas à la transformation cellulaire, des événements oncogéniques additionnels coopérateurs étant nécessaires pour obtenir le phénotype tumoral dans les hémopathies myéloïdes. Il n'était pas connu si la présence de mutations de *TET2* dans les hémopathies lymphoïdes matures était due à un avantage des progéniteurs mutés ou si *TET2* remplissait aussi un rôle spécifique au cours de la différenciation lymphoïde. Dans ce travail, nous avons analysé dans les lymphocytes B (LB), le rôle de *TET2* dans la réaction du centre germinatif (CG) afin de comprendre l'impact de son inactivation sur la réponse immunitaire humorale et sur la transformation des LB matures.

La réponse à l'immunisation par SRBC (sheep red blood cells) ou par un antigène spécifique (NP-CGG) a été analysée dans plusieurs modèles de souris conditionnelles pour *Tet2*. Nos résultats montrent que les souris inactivées pour *Tet2* présentent une hyperplasie des CG, associée à une augmentation du pourcentage et du nombre absolu des LB du CG (BGC), ainsi qu'un blocage de la différenciation plasmocytaire. L'analyse de la réponse au NP-CGG révèle une diminution du taux des IgG1 de haute et de faible affinité dans le sérum et une baisse de la commutation isotopique vers l'IgG1 chez les LB déficients en *Tet2*. L'analyse par ELISPOT confirme une forte baisse de la production et de la sécrétion d'anticorps. Un blocage comparable de la maturation B en contexte TET2KO est observé dans un modèle de CG *in vitro*. Ces résultats établissent le rôle important de *TET2* dans la réponse immunitaire humorale et dans la différenciation lymphoïde B terminale.

L'étude des profils transcriptionnels des BGC déficients pour *Tet2* montre une répression des gènes impliqués dans le contrôle de la sortie du CG, comme le facteur de transcription *PRDM1/BLIMP1*. L'expression ectopique de *PRDM1* dans les LB *Tet2*-déficients permet de

corriger en grande partie le défaut de la réaction du CG *in vitro*, confirmant ainsi le rôle de TET2 dans le programme transcriptionnel de sortie du CG.

La répression transcriptionnelle induite par l'absence de TET2 est associée à la perte focale d'hydroxyméthylation au niveau des enhanceurs de ces gènes cibles, mais aussi une diminution de l'acétylation de la lysine K27 de l'histone 3 (H3K27Ac), marque d'activation transcriptionnelle déposée par l'histone acétyltransférase CBP/CREBBP.

Une interaction physique entre les protéines CBP et TET2 a été rapportée dans d'autres modèles. Une grande partie des enhanceurs des gènes réprimés par la perte de TET2 sont également réprimés lorsque CBP/CREBBP est inactivée, dans des modèles murins ou dans des lymphomes B. Dans ce travail, l'analyse d'une cohorte de 128 patients atteints de lymphome B diffus à grandes cellules (DLBCL) a montré que les mutations *CREBBP* et *TET2* sont mutuellement exclusives, confirmant que ces deux protéines ont des fonctions chevauchantes au cours de la réaction du CG.

Enfin, pour confirmer que la perte de TET2 prédispose à la transformation des LB matures, nous avons croisé les souris TET2KO avec les souris I μ BCL6 (un modèle établi de lymphome) et montré une coopération oncogénique entre les deux mutations : 100% des souris TET2KO/I μ BCL6 développent un lymphome B comparé à seulement 50% des souris du groupe I μ BCL6.

En conclusion, nos résultats montrent que TET2 joue un rôle essentiel dans la régulation de l'expression des gènes permettant la différenciation plasmocytaire ainsi que la sortie du GC. Sa perte de fonction prédispose à la transformation et au développement d'hémopathies lymphoïdes B matures, établissant ainsi son rôle de suppresseur de tumeur dans les LB.

TET2 Deficiency Causes Germinal Center Hyperplasia, Impairs Plasma Cell Differentiation, and Promotes B-cell Lymphomagenesis

Pilar M. Dominguez¹, Hussein Ghamlouch², Wojciech Roskiewicz³, Parveen Kumar³, Wendy Béguélin¹, Lorena Fontán¹, Martín A. Rivas¹, Patrycja Pawlikowska⁴, Marine Armand^{2,4}, Enguerran Mouly², Miguel Torres-Martín⁵, Ashley S. Doane^{1,6}, María T. Calvo Fernandez¹, Matt Durant¹, Veronique Della Valle², Matt Teater^{1,6}, Luisa Cimmino⁷, Nathalie Droin², Saber Tadros⁸, Samaneh Motanagh⁹, Alan H. Shin¹⁰, Mark A. Rubin⁹, Wayne Tam¹¹, Iannis Aifantis⁷, Ross L. Levine¹⁰, Olivier Elemento⁶, Giorgio Inghirami¹¹, Michael R. Green⁸, María E. Figueroa⁵, Olivier A. Bernard², Said Aoufouchi⁴, Sheng Li^{3,12,13}, Rita Shaknovich^{1,4}, and Ari M. Melnick¹

ABSTRACT

TET2 somatic mutations occur in ~10% of diffuse large B-cell lymphomas (DLBCL) but are of unknown significance. Herein, we show that *TET2* is required for the humoral immune response and is a DLBCL tumor suppressor. *TET2* loss of function disrupts transit of B cells through germinal centers (GC), causing GC hyperplasia, impaired class switch recombination, blockade of plasma cell differentiation, and a preneoplastic phenotype. *TET2* loss was linked to focal loss of enhancer hydroxymethylation and transcriptional repression of genes that mediate GC exit, such as *PRDM1*. Notably, these enhancers and genes are also repressed in *CREBBP*-mutant DLBCLs. Accordingly, *TET2* mutation in patients yields a *CREBBP*-mutant gene-expression signature, *CREBBP* and *TET2* mutations are generally mutually exclusive, and hydroxymethylation loss caused by *TET2* deficiency impairs enhancer H3K27 acetylation. Hence, *TET2* plays a critical role in the GC reaction, and its loss of function results in lymphomagenesis through failure to activate genes linked to GC exit signals.

SIGNIFICANCE: We show that *TET2* is required for exit of the GC, B-cell differentiation, and is a tumor suppressor for mature B cells. Loss of *TET2* phenocopies *CREBBP* somatic mutation. These results advocate for sequencing *TET2* in patients with lymphoma and for the testing of epigenetic therapies to treat these tumors. *Cancer Discov* 8(12): 1632-53. ©2018 AACR.

See related commentary by Shingleton and Dave, p. 1515.

INTRODUCTION

Diffuse large B-cell lymphomas (DLBCL) are aggressive tumors and the most common form of non-Hodgkin lymphoma (1). These tumors arise from fully mature and differentiated B cells transiting the germinal center (GC) reaction. GCs form transiently when resting B cells are activated by T cells reacting to an antigenic challenge. GC B cells replicate at an accelerated rate and diversify their immunoglobulin (Ig) genes through somatic hypermutation mediated by the enzyme activation-induced cytosine deaminase (AICDA; ref. 2). This clonal expansion and mutagenesis enables the emergence of B cells with *de novo* high-affinity Ig against specific antigens. Only those GC B cells expressing high-affinity B-cell receptors (BCR; i.e., Ig) for specific antigen are selected by T cells to survive and exit the GC reaction, after which they subsequently differentiate into either memory or long-lived

antibody secreting plasma cells (PC; ref. 3). This proliferative and mutagenic phenotype makes GC B cells especially prone to undergo malignant transformation (4).

Epigenetic mechanisms play a central role in the entry and exit of B cells to the GC reaction and are commonly disrupted in DLBCL (5). Indeed, somatic mutations of histone modifier genes have emerged as a genetic hallmark of this disease (6). Recent studies have shed light on the mechanism through which the commonly mutated histone modifiers *KMT2D*, *EZH2*, *CREBBP*, and *EP300* contribute to pathogenesis. In the case of *EZH2*, gain-of-function mutations result in increased H3K27me3-mediated silencing of gene promoters for cell checkpoint and PC differentiation genes (7, 8). *KMT2D* loss-of-function mutation results in focal reduction of H3K4me1 and hence in impaired activation of gene enhancers that normally respond to CD40 signaling and control expression of tumor suppressor genes such as *TNFAIP3*

¹Department of Medicine, Division of Hematology and Medical Oncology, Weill Cornell Medicine, New York, New York. ²INSERM U1170, équipe labellisée Ligue Nationale Contre le Cancer, Gustave Roussy, Université Paris-Saclay, Villejuif, France. ³The Jackson Laboratory for Genomic Medicine, Farmington, Connecticut. ⁴CNRSUMR8200, équipe labellisée Ligue Nationale Contre le Cancer, Gustave Roussy, Université Paris-Saclay, Villejuif, France. ⁵Sylvester Comprehensive Cancer Center, Department of Human Genetics, University of Miami, Miller School of Medicine, Miami, Florida. ⁶Institute for Computational Biomedicine, Department of Physiology and Biophysics, Weill Cornell Medicine, New York, New York. ⁷Department of Pathology, Laura and Isaac Perlmutter Cancer Center, and The Helen L. and Martin S. Kimmel Center for Stem Cell Biology, NYU School of Medicine, New York, New York. ⁸Department of Lymphoma/Myeloma and Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas. ⁹Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, New York. ¹⁰Center for Epigenetics Research, Memorial Sloan Kettering Cancer Center, New York, New York. ¹¹Pathology and Laboratory Medicine Department, Weill Cornell Medicine, New York, New York. ¹²The Jackson Laboratory Cancer Center, Bar Harbor, Maine. ¹³Department of Genetics and Genome Sciences, University of Connecticut School of Medicine, Farmington, Connecticut. ¹⁴Cancer Genetics, Inc., Rutherford, New Jersey.

Note: Supplementary data for this article are available at Cancer Discovery Online (<http://cancerdiscovery.aacrjournals.org/>).

P.M. Dominguez, H. Ghamlouch, and W. Rosikiewicz contributed equally to this article.

Current address for P.M. Dominguez: Peter MacCallum Cancer Centre, Melbourne, Australia.

Corresponding Authors: Ari M. Melnick, Weill Cornell Medicine, 413 East 69th Street, BB-1430, New York, NY 10021. Phone: 646-962-6725; Fax: 212-746-8866; E-mail: amm2014@med.cornell.edu; Olivier A. Bernard, INSERM U1170, Institut Gustave Roussy, 39 rue Camille Desmoulins, 94805 Villejuif, France; E-mail: olivier.bernard@inserm.fr; Said Aoufouchi, CNRS, UMR8200, Institut Gustave Roussy, 39 rue Camille Desmoulins, 94805 Villejuif, France; E-mail: Said.AOUFOUCHI@gustaveroussy.fr; Sheng Li, 10 Discovery Drive, Farmington, CT 06032; E-mail: sheng.li@jax.org; and Rita Shaknovich, 201 Route 17 N, 2nd Floor, Rutherford, NJ 07070; E-mail: rshaknovich@gmail.com

doi: 10.1158/2159-8290.CD-18-0657

©2018 American Association for Cancer Research.

(9, 10). *CREBBP* loss-of-function mutation causes focal loss of H3K27 acetylation of enhancers and hence transcriptional repression of genes involved in GC exit signals including the BCR, cytokine, CD40, and antigen presentation pathways (11, 12). Hence, disruption of histone-modifying complexes is well documented as a driving force in malignant transformation of GC B cells.

In contrast, not much is known in DLBCL regarding the contribution and relevance of the other major arm of epigenetic modifiers, in particular the enzymes that mediate covalent modifications of DNA. It is known that GC B cells acquire a cytosine hypomethylation signature that is highly dependent on the actions of AICDA, possibly due to replacement of methylated cytosines with unmethylated nucleotides during somatic hypermutation (13). Cytosine methylation profiling studies have shown that DLBCLs feature aberrant cytosine methylation signatures and increased intertumor and intratumor epigenetic DNA methylation heterogeneity, which is also directly linked to the presence of AICDA (14–17). Virtually nothing is known about cytosine hydroxymethylation in the humoral immune response and DLBCL.

Intriguingly, genome-sequencing studies show that ~10% of DLBCLs display somatic mutations of *TET2* (6, 18), which encodes a dioxygenase that converts 5-methylcytosine (5mC) to 5-hydroxymethylcytosine (5hmC; ref. 19). *TET2* is well known to be frequently mutated in myeloid disorders including acute myeloid leukemia and myelodysplasia (20–22). Somatic mutations of *TET2* in hematopoietic stem cells (HSC) are also linked to the development of angioimmunoblastic T-cell lymphoma (23). 5hmC has been described as a necessary intermediate step in DNA demethylation in cooperation with the base excision repair machinery, but is also an independent epigenetic mark linked to transcriptional activation (24–26). Previous studies in hematopoietic stem/progenitor cells (HSPC) and embryonic stem cells have shown that 5hmC is enriched in gene bodies and enhancers and can modulate enhancer activity, affecting gene expression (26–30).

Notably, it has been shown that patients with DLBCL with *TET2* mutations harbor the same mutation in their HSPCs (31). Because *TET2* mutations are typically linked to HSCs and may be present in individuals with healthy hematopoiesis, it is not clear whether the presence of *TET2* mutations in DLBCL is a passenger event or can be specifically shown to function as a tumor suppressor in this disease. Here, we investigate whether and how *TET2* loss of function is linked to malignant transformation of GC B cells by performing functional experiments in mouse models and linking these effects to observations from human patients with *TET2*-mutant DLBCLs.

RESULTS

Tet2 Deletion in Hematopoietic Cells Induces GC Hyperplasia

Because DLBCLs arise from GC B cells, we studied whether *TET2* loss affected GC B-cell formation. We first examined *Tet2* expression in the humoral immune response based on published RNA sequencing (RNA-seq) profiles (32) and observed that *Tet2/TET2* was expressed in both naïve and

GC B cells in mouse and human cells, with reduced levels in further differentiated cells (Supplementary Fig. S1A). Expression was validated using qPCR in mouse GC B cells (Supplementary Fig. S1B). To reflect the situation in human patients with DLBCL where *TET2* mutations are observed in HSCs, we performed conditional deletion of *Tet2* in HSCs in mice under the control of *Vav-Cre* to delete this gene in the HSC compartment. To determine if HSC knockout of *Tet2* could manifest effects in mature B cells undergoing the GC reaction, we induced T cell-dependent antigen response with sheep red blood cell injections (SRBC) in wild-type (WT; *Vav-Cre/Tet2^{+/+}*) and *Tet2*-deficient (*Vav-Cre/Tet2^{-/-}*) mice. Animals were sacrificed 10 days later, and spleens were examined by flow cytometry and IHC. We observed a significant increase in the percentage of GC B cells (B220⁺GL7⁺CD95⁺) from *Vav-Cre/Tet2^{-/-}* mice compared with *Vav-Cre/Tet2^{+/+}* mice ($P < 0.001$; Fig. 1A). *Tet2* deletion in *Vav-Cre/Tet2^{-/-}* GC B cells was confirmed by qPCR (Supplementary Fig. S1B). In contrast, there was no difference in the percentage of total B cells (B220⁺ cells) or the other B-cell subsets in the spleen between *Vav-Cre/Tet2^{+/+}* and *Vav-Cre/Tet2^{-/-}* mice, including mature B cells (B220⁺IgD⁺IgM⁺), transitional B cells (B220⁺IgD^{int}IgM⁺), follicular B cells (B220⁺CD23⁺CD21⁺), marginal zone B cells (B220⁺CD23^bCD21⁺), and plasmablasts/PC (CD138⁺; Fig. 1B; Supplementary Fig. S1C). IHC analysis of the spleens using peanut agglutinin (PNA) and Ki-67 revealed an increase in the size of the GCs of *Vav-Cre/Tet2^{-/-}* compared with *Vav-Cre/Tet2^{+/+}* mice (Fig. 1C and D), but not in the number of GCs between *Vav-Cre/Tet2^{-/-}* and *Vav-Cre/Tet2^{+/+}* animals (Fig. 1C and E).

GC Hyperplasia Induced by *Tet2* Loss of Function Is B-cell Autonomous

Because the *Vav-Cre* allele knocks out *Tet2* in all hematopoietic lineages, we cannot exclude that this GC is caused by *Tet2* loss of function in some other cell type. Therefore, we generated *CD19-Cre/Tet2^{-/-}* mice, in which *Tet2* loss of function is induced specifically during B-cell development. These mice were immunized with SRBC as described above to examine GC formation. Similar to *Vav-Cre/Tet2^{-/-}* mice, *CD19-Cre/Tet2^{-/-}* animals also displayed a significant increase in the proportion of GC B cells after immunization (Fig. 2A), whereas other mature B-cell subpopulations were not affected (Supplementary Fig. S2A). As in the *Vav-Cre* model, we also observed GC hyperplasia as shown by IHC with staining for PNA and Ki67 (Fig. 2B). This was due to expansion in the size, not the number, of GCs in *CD19-Cre/Tet2^{-/-}* animals compared with *CD19-Cre/Tet2^{+/+}* mice (Fig. 2C and D). This result indicates that GC hyperplasia is B-cell autonomous and not dependent on *TET2* loss of function in other cell lineages.

We next induced deletion of *Tet2* specifically in already formed GC B cells crossing *Tet2^{fl/fl}* mice with those bearing the *Cγ1-Cre* (GC-specific) allele. In contrast to the *Vav-Cre* and *CD19-Cre* models, the immunized *Cγ1-Cre/Tet2^{-/-}* mice showed no differences in the percentage of GC B cells between *TET2*-deficient (*Cγ1-Cre/Tet2^{-/-}*) and WT (*Cγ1-Cre/Tet2^{+/+}*) mice (Fig. 2E). Likewise, the percentage of total B cells and other B-cell subpopulations was similar in *Cγ1-Cre/Tet2^{-/-}* and *Cγ1-Cre/Tet2^{+/+}* mice (Supplementary Fig. S2B). We also found that the number and size of GCs was comparable in

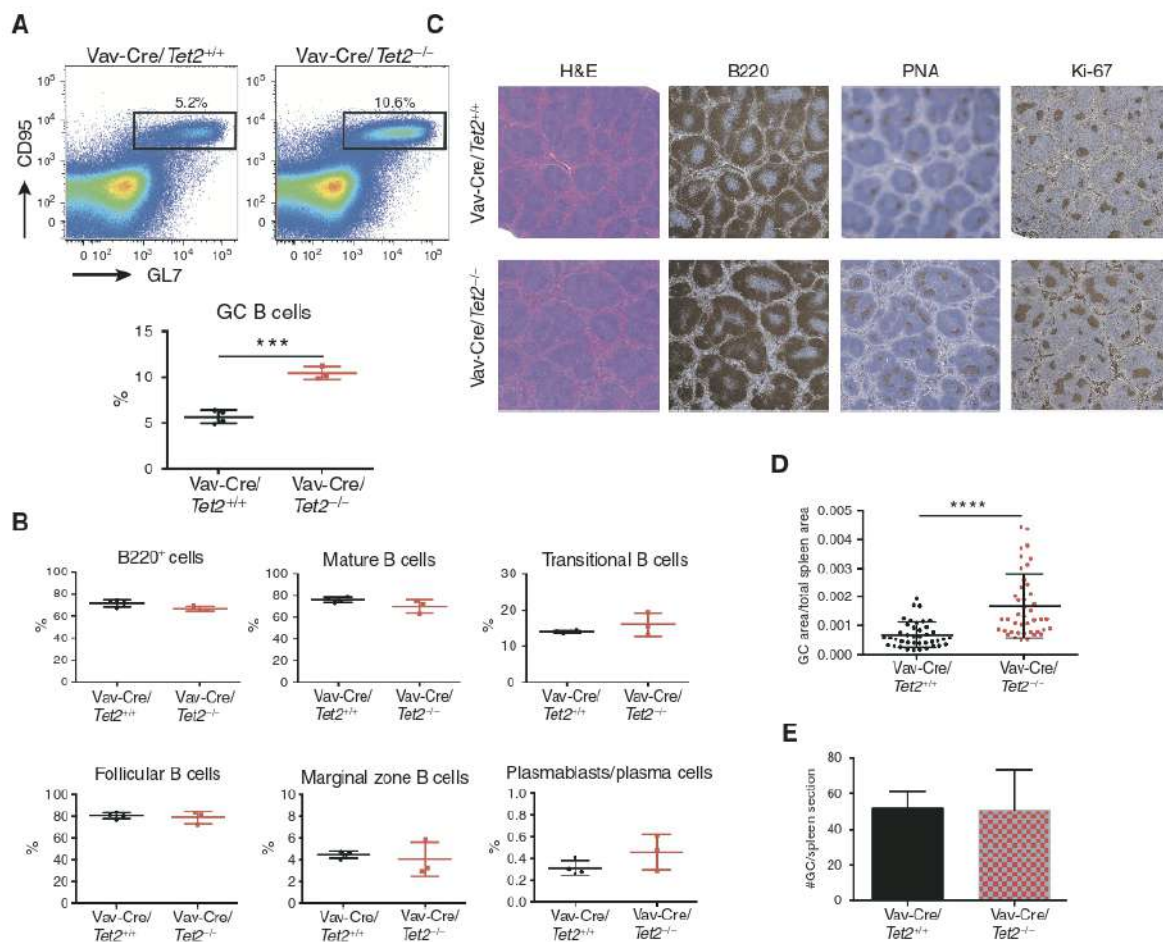


Figure 1. Tet2 deletion in hematopoietic cells induces GC hyperplasia. **A**, Representative flow cytometry plot and quantification of (B220⁺CD95⁺GL7⁺) GC B cells from Vav-Cre/Tet2^{+/+} ($n=5$) and Vav-Cre/Tet2^{-/-} ($n=5$) mice at day 10 after immunization with SRBC. **B**, Quantification of flow cytometry data corresponding to total B cells (B220⁺), mature B cells (B220⁺IgD^{int}IgM⁺), transitional B cells (B220⁺IgD^{int}IgM⁺), follicular B cells (B220⁺CD23^{lo}CD21⁻), marginal zone B cells (B220⁺CD23^{hi}CD21⁺), and plasmablasts/PC (CD138⁺) in the spleens of Vav-Cre/Tet2^{+/+} and Vav-Cre/Tet2^{-/-} mice. **C**, Representative histologic sections of formalin-fixed, paraffin-embedded spleens from Vav-Cre/Tet2^{+/+} and Vav-Cre/Tet2^{-/-} mice. Sections were stained with H&E and antibodies specific for B220, PNA, and Ki-67. **D** and **E**, Quantification of GC area (**D**) and number of GCs (**E**) in the spleens of Vav-Cre/Tet2^{+/+} and Vav-Cre/Tet2^{-/-} mice. Two-tailed t test, ***, $P < 0.001$; ****, $P < 0.0001$.

Cy1-Cre/Tet2^{-/-} and Cy1-Cre/Tet2^{+/+} mice (Fig. 2F-H). We confirmed efficient Tet2 deletion in both CD19-Cre/Tet2^{-/-} and Cy1-Cre/Tet2^{-/-} GC B cells (Supplementary Fig. S2C and S2D), excluding the possibility that the different phenotypes reflected differences in excision efficiency by the two Cre lines. The differential effect in the CD19-Cre versus Cy1-Cre is consistent with what has been observed with other DLBCL epigenetic tumor suppressor models (10, 12).

TET2 Loss of Function Impairs Affinity Maturation and PC Differentiation in a B Cell-Autonomous Manner

The aberrant GC phenotype induced by TET2 loss of function prompted us to also determine whether it is required for immunoglobulin affinity maturation. Therefore, we analyzed the ability of Vav-Cre/Tet2^{-/-} mice to generate specific antibodies against the T cell-dependent antigen NP-Chicken

Gamma Globulin (NP-CGG). Mice were immunized with NP-CGG and then received a booster shot at day 21, after which serum was collected at day 35 to detect Ig and bone marrow at day 60 to identify high-affinity PC (Fig. 3A). ELISA analysis of serum revealed a significant reduction in the titers of both low-affinity (NP26) and high-affinity (NP7) IgG1 antibodies in Vav-Cre/Tet2^{-/-} mice compared with Vav-Cre/Tet2^{+/+} mice ($P < 0.001$; Fig. 3B). The defect disproportionately affected the high-affinity species, consistent with a perturbation in AICDA-driven immunoglobulin affinity maturation during the GC response ($P < 0.01$; Fig. 3C). We also tested lambda light chain maturation (Ig λ) and observed the same pattern of reduced production of both low- and high-affinity Ig λ in Vav-Cre/Tet2^{-/-} mice (NP26, $P < 0.0001$; NP4, $P < 0.001$; Fig. 3B), again with disproportionate effect on the high-affinity species (ratio NP7/NP26, $P < 0.01$; Fig. 3C). Notably, we found a similar reduction in high- and low-affinity

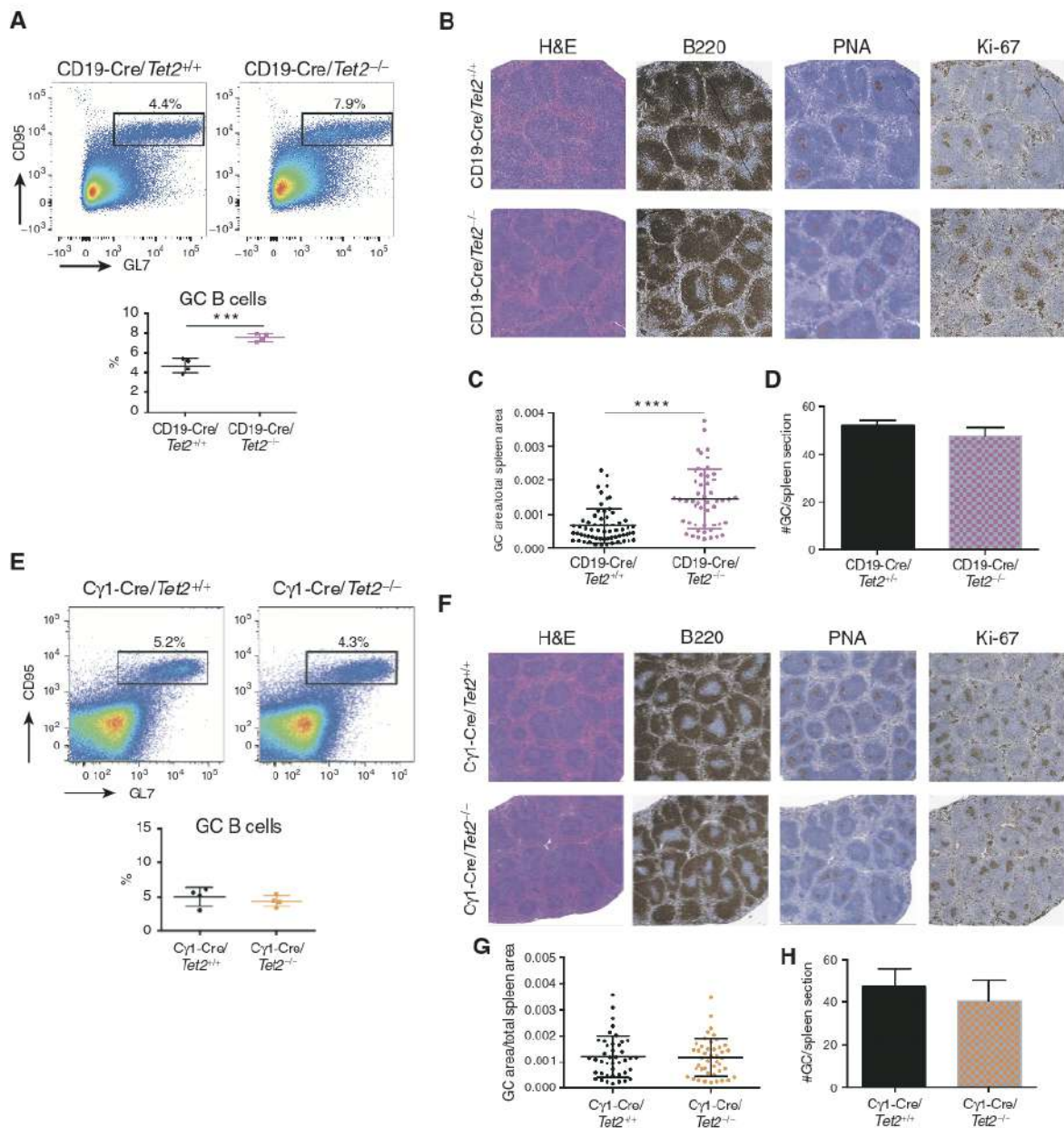


Figure 2. GC hyperplasia induced by *Tet2* loss of function is B-cell autonomous. **A**, Representative flow cytometry plot and quantification of ($B220^+CD95^+GL7^+$) GC B cells from *CD19-Cre/Tet2^{+/+}* ($n = 4$) and *CD19-Cre/Tet2^{-/-}* ($n = 4$) mice at day 10 after immunization with SRBC. **B**, Representative histologic sections of formalin-fixed, paraffin-embedded spleens from *CD19-Cre/Tet2^{+/+}* and *CD19-Cre/Tet2^{-/-}* mice. Sections were stained with H&E and antibodies specific for B220, PNA, and Ki-67. **C** and **D**, Quantification of GC area (**C**) and number of GCs (**D**) in the spleens of *CD19-Cre/Tet2^{+/+}* and *CD19-Cre/Tet2^{-/-}* mice. **E**, Representative flow cytometry plot and quantification of ($B220^+CD95^+GL7^+$) GC B cells from *Cγ1-Cre/Tet2^{+/+}* ($n = 5$) and *Cγ1-Cre/Tet2^{-/-}* ($n = 5$) mice at day 10 after immunization with SRBC. **F**, Representative histologic sections of formalin-fixed, paraffin-embedded spleens from *Cγ1-Cre/Tet2^{+/+}* and *Cγ1-Cre/Tet2^{-/-}* mice. Sections were stained with H&E and antibodies specific for B220, PNA, and Ki-67. **G** and **H**, Quantification of GC area (**G**) and number of GCs (**H**) in the spleens of *Cγ1-Cre/Tet2^{+/+}* and *Cγ1-Cre/Tet2^{-/-}* mice. Two-tailed t test; ***, $P < 0.001$; ****, $P < 0.0001$.

IgG1 and Ig λ , predominantly in high-affinity antibodies, in *Cγ1-Cre/Tet2^{-/-}* mice compared with *Cγ1-Cre/Tet2^{+/+}* mice (Supplementary Fig. S3A and S3B), indicating that *TET2* deletion at the GC stage also impairs affinity maturation. In addition, we performed ELISPOT assays in bone marrow at day 60 to measure the abundance of NP-specific antibody-

secreting cells. This analysis yielded a marked reduction in cells that secrete high- and low-affinity IgG1 NP antibodies in the bone marrow of *Vav-Cre/Tet2^{-/-}* mice ($P < 0.01$; Fig. 3D). In contrast, the abundance of non-class-switched anti-NP IgM-secreting cells was not significantly reduced compared with controls (Fig. 3E). Importantly, this effect was observed

regardless of whether TET2 loss of function was induced in HSCs (Vav-Cre/Tet2^{-/-}), early B cells (CD19-Cre/Tet2^{-/-}), or in GCs (Cγ1-Cre/Tet2^{-/-}; Fig. 3D; Supplementary Fig. S3C–S3F). These findings are suggestive of an impairment in GC exit functions and induction of the PC differentiation program. Accordingly, we observed a reduction in the percentage of PC in the spleen of SRBC-immunized constitutive Tet2-deficient mice (Tet2^{-/-}; n = 5) as compared with controls (Tet2^{+/+}; n = 4; P < 0.01; Fig. 3F). This GC exit/PC differentiation block is consistent with the mechanism of malignant transformation induced by loss of function of epigenetic regulators in DLBCL (9–12). Importantly, this effect is B-cell autonomous

and observed regardless of whether Tet2 deletion was induced in HSCs, pre-B cells, or GC B cells.

To further establish the cell-autonomous nature of this phenotype, we sorted naive B cells from Tet2^{-/-} animals and compared their behavior *in vitro* under culture conditions that mimic the GC reaction (ref. 33; Fig. 3G). This method involves the plating of naive B cells together with 40LB cells and sequential exposure to IL4 for 4 days followed by IL21 for another 4 days (33). We plated equal numbers of Tet2^{-/-} and Tet2^{+/+} cells and observed a significant increase in the amount of B cells both at the 4-day (D4) and 8-day (D8) time points derived from the Tet2^{-/-} mice (P < 0.0005; Fig. 3H). Furthermore,

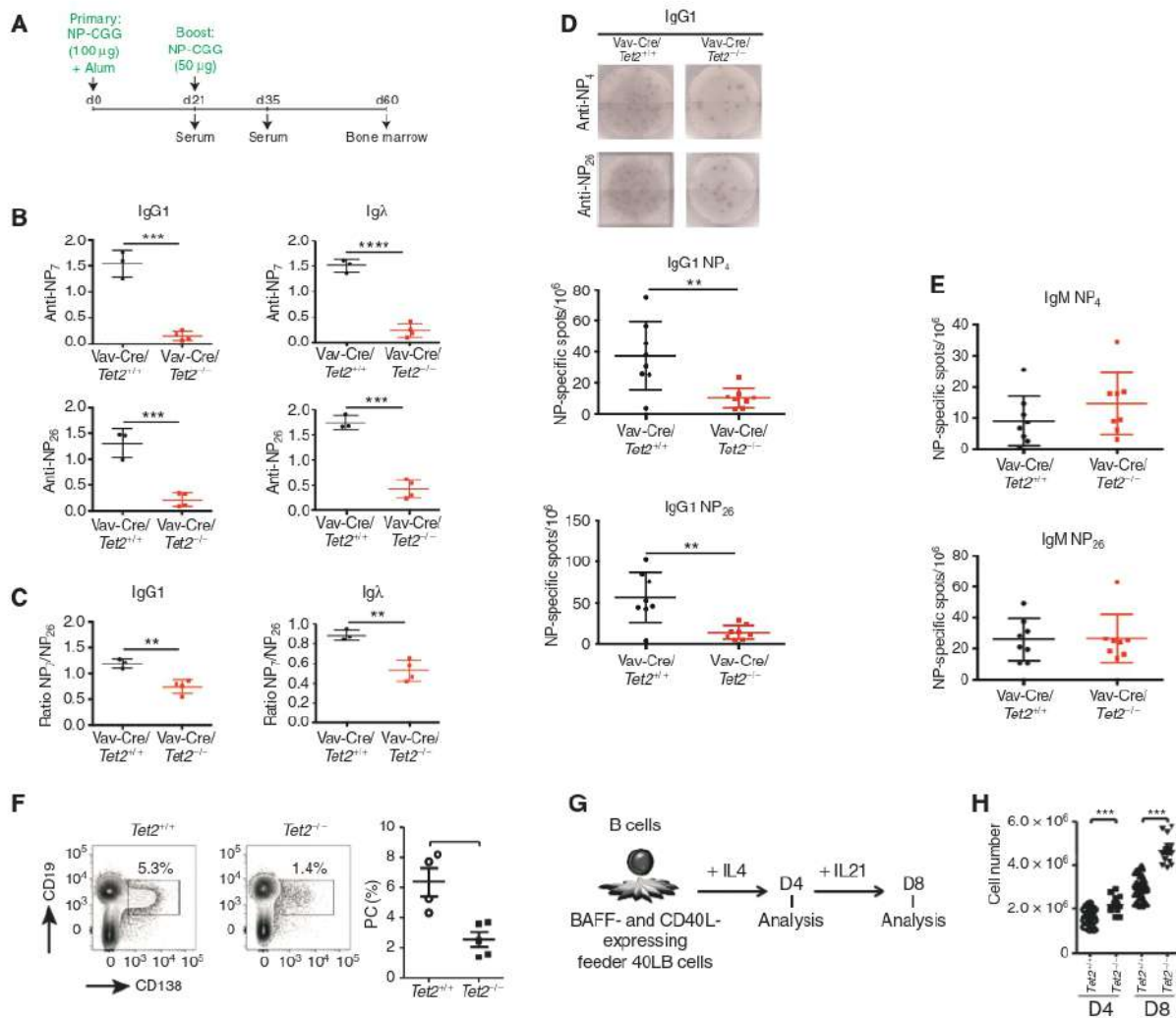


Figure 3. Tet2 loss of function impairs affinity maturation and PC differentiation in a B-cell-autonomous manner. **A**, Schematic diagram of the protocol of primary and secondary immunizations. **B**, Thirty-five days after immunization (14 days after boost), NP-specific antibodies (IgG1 and Igλ) were measured in the sera of Vav-Cre/Tet2^{+/+} and Vav-Cre/Tet2^{-/-} mice by ELISA. **C**, NP₇/NP₂₆ ratio of NP-specific antibodies detected in **B**. **D** and **E**, Sixty days after immunization (40 days after boost), NP-specific IgG1-secreting (**D**) and IgM-secreting (**E**) cells from the bone marrow of Vav-Cre/Tet2^{+/+} and Vav-Cre/Tet2^{-/-} mice were quantified by ELISPOT. Two-tailed t test; *, P < 0.05; **, P < 0.01; ***, P < 0.001; ****, P < 0.0001. **F**, Splenic PC formation after SRBC immunization of Tet2^{+/+} and Tet2^{-/-} mice, shown as FACS profiles (left) and quantitation (right) on day 32; each symbol represents one mouse. **G**, Schematic illustration of the *in vitro* GC B cells (iGC) and PB (iPB) culture system. **H**, Number of live Tet2^{+/+} and Tet2^{-/-} B cells cultured on 40LB with IL4 on D4 and IL21 on D8. (continued on next page)

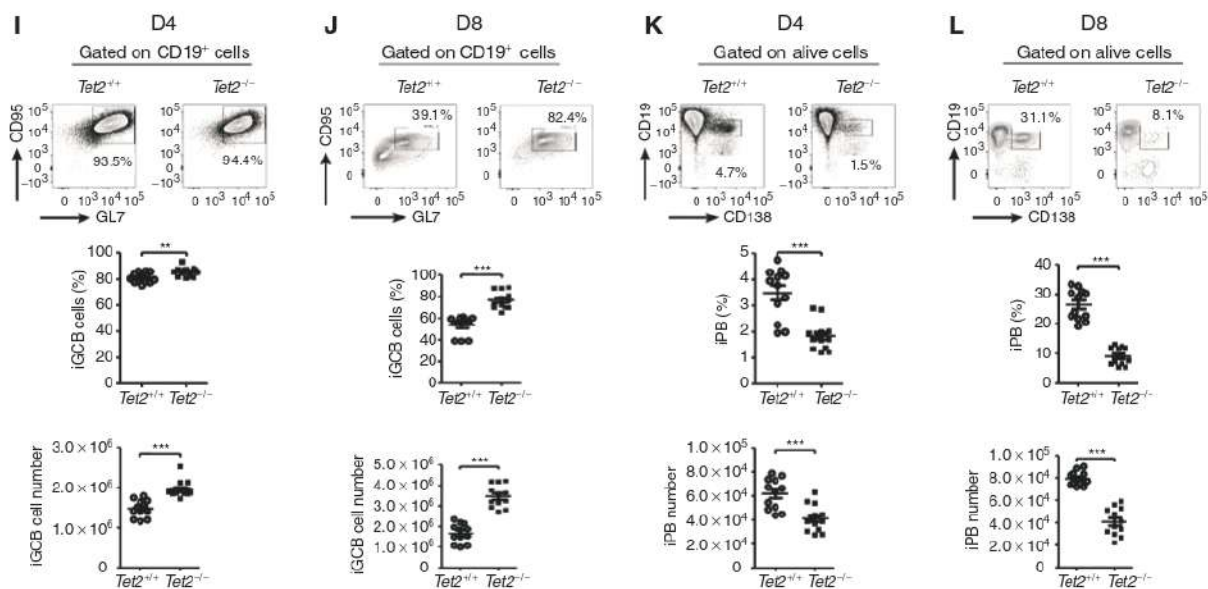


Figure 3. (Continued) I and J, Flow cytometry plots of iGCB cells on D4 (I) and D8 (J); the gated area shows the live iGCB cells (CD19⁺GL7⁺FAS⁻). Histograms represent the percentage (top) and the cell number (bottom) of iGCB cells. K and L, Flow cytometry plots of iPBC cells on D4 (K) and D8 (L); the gated area shows the live iPBC cells (CD19⁺GL7⁺FAS⁻). Histograms represent the percentage (top) and the cell number (bottom) of iPBC cells. For D4, symbols represent triplicate of 5 independent experiments. For D8, symbols represent triplicate of 4 independent experiments. All *P* values were calculated using an unpaired Student *t* test; **, *P* < 0.005; ***, *P* < 0.0005 in all experiments.

flow cytometry analysis showed significantly increased numbers of induced GC B cells (iGC; CD19⁺GL7⁺CD95⁺) derived from *Tet2*^{-/-} versus *Tet2*^{+/+} mice at D4 and D8 (Fig. 3I and J). In marked contrast, the induced plasmablast (iPB) population (CD19⁺CD138⁺) was significantly lower in *Tet2*^{-/-} compared with *Tet2*^{+/+} cells at both time points (Fig. 3K and L).

These *in vitro* data mirror the *in vivo* results and indicate that *Tet2* loss in B cells induces expansion of GC B cells with corresponding blockade of PC differentiation in a B cell-autonomous fashion. Along these lines, analysis of CD19⁺CD138⁻*Tet2*^{-/-} iPB showed significantly higher expression of the GC B cell-associated marker GL7, lower expression of CD138, and higher expression of CD19 compared with *Tet2*^{+/+} cells (Supplementary Fig. S3G). Analysis of IgG1 and IgM expression on D4 iGC B cells showed that the percentage of IgG1-switched *Tet2*^{-/-} iGC B cells was significantly lower compared with *Tet2*^{+/+} cells, suggesting a defect in class switch recombination, similar to what was observed *in vivo* (Supplementary Fig. S3H).

Tet2-Deficient GC B Cells Exhibit Repression of Genes Involved in GC Exit

In order to identify putative mechanisms responsible for the GC disruption and putative tumor suppressor function of TET2, we analyzed the impact of TET2 loss of function on the transcriptional signatures of GC B cells by RNA-seq. Unsupervised hierarchical clustering and principal component analysis showed a clear difference in transcriptional profiles between *Vav-Cre/Tet2*^{+/+}- and *Vav-Cre/Tet2*^{-/-}-sorted GC B cells (Fig. 4A and B). We then performed a supervised analysis and identified 2,100 differentially expressed genes

(DEG) between *Vav-Cre/Tet2*^{-/-} and *Vav-Cre/Tet2*^{+/+} GC B cells [fold change (FC) > 1.2, *P* < 0.05; Fig. 4C]. Notably, the majority of these were skewed toward being more repressed in *Tet2*-deficient GC B cells (*n* = 1,294; 61.6%), suggesting that TET2 is mainly a transcriptional activator in GC B cells. Gene set enrichment analysis (GSEA) was performed against an extensive database of GC and DLBCL relevant gene sets. This analysis indicated significant enrichment for downregulation of genes involved in the exit of GC B cells from the GC reaction, such as NFκB- and IL6-induced genes (Fig. 4D; Supplementary Table S1). Enrichment of those gene sets among downregulated genes in *Vav-Cre/Tet2*^{-/-} was independently confirmed using a hypergeometric test (Supplementary Fig. S4A).

We also noted enrichment for repression of genes whose enhancers are normally repressed during the GC reaction but are induced upon GC exit (Fig. 4D). The genes controlled by these enhancers are repressed in GC B cells by the BCL6 transcriptional repressor with its corepressor SMRT, and reactivated upon GC exit by the CREBBP histone acetyltransferase (11, 34). These genes are accordingly aberrantly repressed in patients with DLBCL and follicular lymphoma (FL) with *CREBBP* somatic mutations (11, 12). Notably, genes repressed in *Tet2*-deficient GC B cells were significantly enriched for genes that are also downregulated in murine and primary human *CREBBP*-deficient lymphomas (Fig. 4D and E; Supplementary Table S1). Likewise, there was significant enrichment for BCL6/SMRT-regulated genes as well as genes that are upregulated upon BCL6 knockdown (Fig. 4D and F; Supplementary Table S1). Hence, TET2 function is associated with regulation of genes whose enhancers

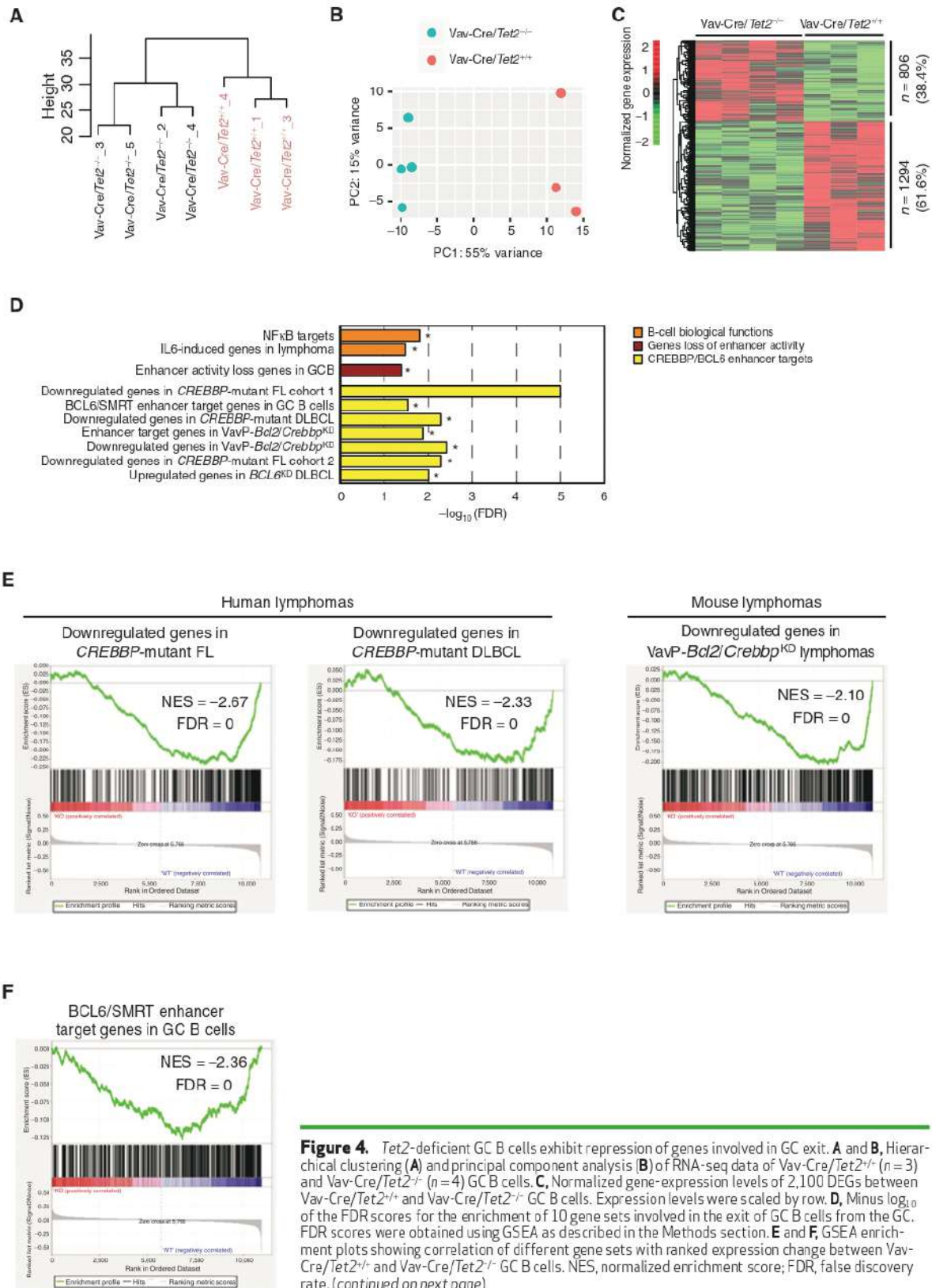


Figure 4. *Tet2*-deficient GC B cells exhibit repression of genes involved in GC exit. **A** and **B**, Hierarchical clustering (**A**) and principal component analysis (**B**) of RNA-seq data of Vav-Cre/*Tet2*^{-/-} (*n* = 3) and Vav-Cre/*Tet2*^{+/+} (*n* = 4) GC B cells. **C**, Normalized gene-expression levels of 2,100 DEGs between Vav-Cre/*Tet2*^{+/+} and Vav-Cre/*Tet2*^{-/-} GC B cells. Expression levels were scaled by row. **D**, Minus log₁₀ of the FDR scores for the enrichment of 10 gene sets involved in the exit of GC B cells from the GC. FDR scores were obtained using GSEA as described in the Methods section. **E** and **F**, GSEA enrichment plots showing correlation of different gene sets with ranked expression change between Vav-Cre/*Tet2*^{+/+} and Vav-Cre/*Tet2*^{-/-} GC B cells. NES, normalized enrichment score; FDR, false discovery rate. (continued on next page)

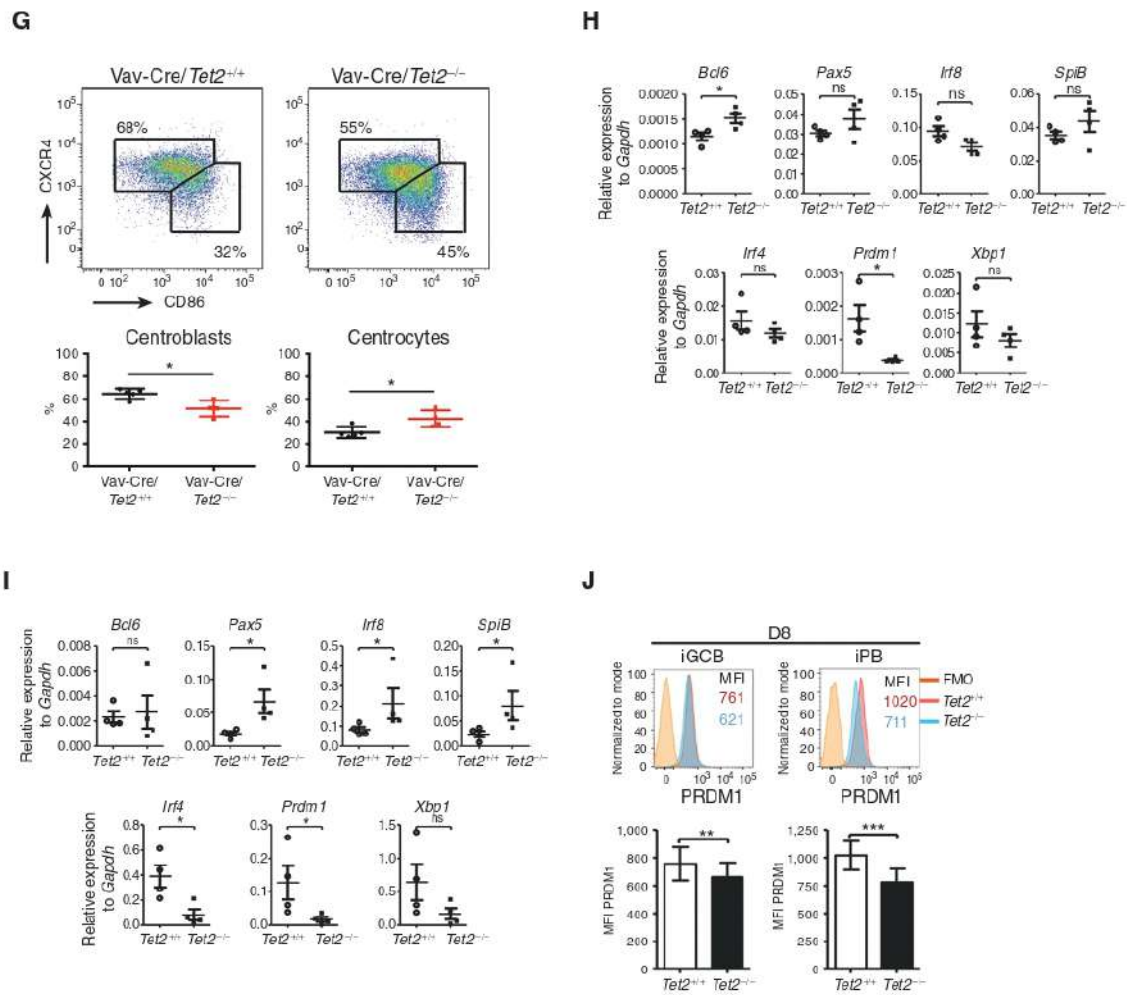


Figure 4. (Continued) **G**, Representative flow cytometry plot and quantification of centroblasts (CXCR4^{hi}CD86^{hi}) and centrocytes (CXCR4^{lo}CD86^{lo}) in the GC of Vav-Cre/*Tet2*^{+/+} ($n = 5$) and Vav-Cre/*Tet2*^{-/-} ($n = 5$) mice at day 10 after immunization with SRBC. **H**, Expression of *Bcl6*, *Pax5*, *Irf8*, *SpiB*, *Irf4*, *Prdm1*, and *Xbp1* by iGCB cells on D8, as measured by quantitative RT-PCR ($n = 4$). **I**, Expression of *Bcl6*, *Pax5*, *Irf8*, *SpiB*, *Irf4*, *Prdm1*, and *Xbp1* by iPB on D8, as measured by quantitative RT-PCR ($n = 4$). **J**, Representative PRDM1 intracellular staining profile of iGCB (CD19⁺CD138⁻) and iPB (CD19⁺CD138⁺) at D8. Numbers indicate the median fluorescence intensity (MFI) of PRDM1. All *P* values were calculated using an unpaired Student *t* test; *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$; ns, not significant, in all experiments.

are switched off in the GC reaction by BCL6 and activated upon GC exit, through the actions of CREBBP.

These data prompted us to examine the impact of TET2 deficiency on B cells exiting the GC reaction. GCs are divided into two regions. The GC dark zone is composed of proliferating and mutating GC B cells called centroblasts, whereas the GC light zone contains noncycling GC B cells that have migrated from the dark zone, which are called centrocytes. Centrocytes can be induced to exit the GC reaction by T follicular helper (TFH) cells (3). Vav-Cre/*Tet2*^{-/-} mice manifested disruption and skewing in the polarity of the GC reaction, with a relative and significant increase in the proportion of centrocytes (Fig. 4G) as compared with control mice. We observed similar skewed proportion toward centrocytes in CD19-Cre/*Tet2*^{-/-} (Supplementary Fig. S4B) but not in *Cy1*-Cre/*Tet2*^{-/-} mice (Supplementary Fig. S4C). These results

suggest that *Tet2* deficiency impairs the ability of GC B cells to exit the GC reaction, leading to their accumulation as centrocytes. To confirm that the changes in gene expression observed in total GC B cells from Vav-Cre/*Tet2*^{-/-} mice were not a consequence of the skewed GC B-cell populations in these animals, we performed RNA-seq in GC B cells from *Cy1*-Cre/*Tet2*^{-/-} and *Cy1*-Cre/*Tet2*^{+/+} mice, which do not exhibit such skewing. This analysis again showed clear separation between TET2-deficient and WT GC B cells (Supplementary Fig. S5A and S5B), with predominant downregulation of gene expression after *Tet2* deletion in *Cy1*-Cre/*Tet2*^{-/-} GC B cells compared with *Cy1*-Cre/*Tet2*^{+/+} cells. More specifically, 1,768 genes were downregulated (59.5%), whereas 1,204 were upregulated (40.5%) as a result of TET2 loss of function in *Cy1*-Cre/*Tet2*^{-/-} GC B cells [FC > 1.2; false discovery rate (FDR) < 0.05; Supplementary Fig. S5C]. Similar to the case of

Vav-Cre/*Tet2*^{-/-} GC B cells, genes downregulated in *Cy1-Cre/Tet2*^{-/-} GC B cells featured enrichment for genes downregulated in human FL and DLBCL with *CREBBP* mutations and mouse *Crebbp* knockdown cells, and NFκB targets (Supplementary Fig. S5D and S5E).

To further explore the centrocyte defect in the absence of *Tet2*, we used qPCR to examine the expression of several key transcription factor (TF) genes that must be switched off (*Pax5*, *Irf8*, *SpiB*, and *Bcl6*) for exit from the GC, or switched on (*Irf4*, *Xbp1*, and *Prdm1*) for GC exit and the transition to the PC phenotype, using our *in vitro* GC culture system. This analysis yielded progressively more significant aberrant expression of these genes as these iGC cells transitioned the *ex vivo* GC reaction and attempted to differentiate to the PC state. Thus, we observed a significant increase of the GC TF *Bcl6* and the corresponding reduction in the GC exit and PC TF *Prdm1* at D8 in *Tet2*^{-/-} iGC B compared with *Tet2*^{+/+} cells ($P < 0.05$; Fig. 4H). Similar results were obtained for iPB at D8, with higher expression of *Pax5*, *Irf8*, and *SpiB* and lower expression of *Irf4* and *Prdm1* in *Tet2*^{-/-} than in *Tet2*^{+/+} cells ($P < 0.05$; Fig. 4I). Flow cytometry for the PC master regulator PRDM1 indicated a reduction in protein expression at D8 in *Tet2*^{-/-} iGC B ($P < 0.01$) and iPB ($P < 0.001$) compared with *Tet2*^{+/+} cells (Fig. 4J). Together, these results show that *Tet2* loss causes failure to induce genes repressed in the GC reaction and which drive terminal differentiation of GC B cells.

TET2 Deficiency Results in Reduction of Enhancer Cytosine Hydroxymethylation and Histone Acetylation

To gain insight into the mechanism through which *Tet2* loss of function perturbs gene expression, we next examined the genomic distribution of 5hmC using hMeDIP-Seq in purified GC B cells from Vav-Cre/*Tet2*^{-/-} versus Vav-Cre/*Tet2*^{+/+} mice. 5hmC peaks were called using MACS2 and were found to be distributed predominantly in intergenic areas and introns in both WT and *Tet2*-deficient GC B cells (Fig. 5A). This result is in line with previous publications showing that TET2 loss of function largely affects DNA hydroxymethylation at enhancers (28, 29, 35–39). We previously mapped the location of gene enhancers in murine GC-derived B cells based on histone modifications (11). When these peaks were overlapped with the 5hmC peaks in Vav-Cre/*Tet2*^{+/+} GC B cells, we observed that at least 70% of enhancers were marked by 5hmC (Fig. 5B). Next, we identified the differentially hydroxymethylated regions (DHMR) between Vav-Cre/*Tet2*^{-/-} and Vav-Cre/*Tet2*^{+/+} GC B cells. This analysis yielded 24,645 DHMRs lost in Vav-Cre/*Tet2*^{-/-} GC B cells (representing 5.5% of the total 5hmC peaks) and only 51 DHMRs gained in Vav-Cre/*Tet2*^{-/-} cells (Fig. 5C). This focal loss of peaks was located predominantly in intergenic and intronic areas in accordance with the general distribution of 5hmC (Fig. 5D). Plotting the read density for these distal and intronic enhancer DHMRs illustrated the reduction in 5hmC reads at these sites (Fig. 5E; Supplementary Fig. S6A).

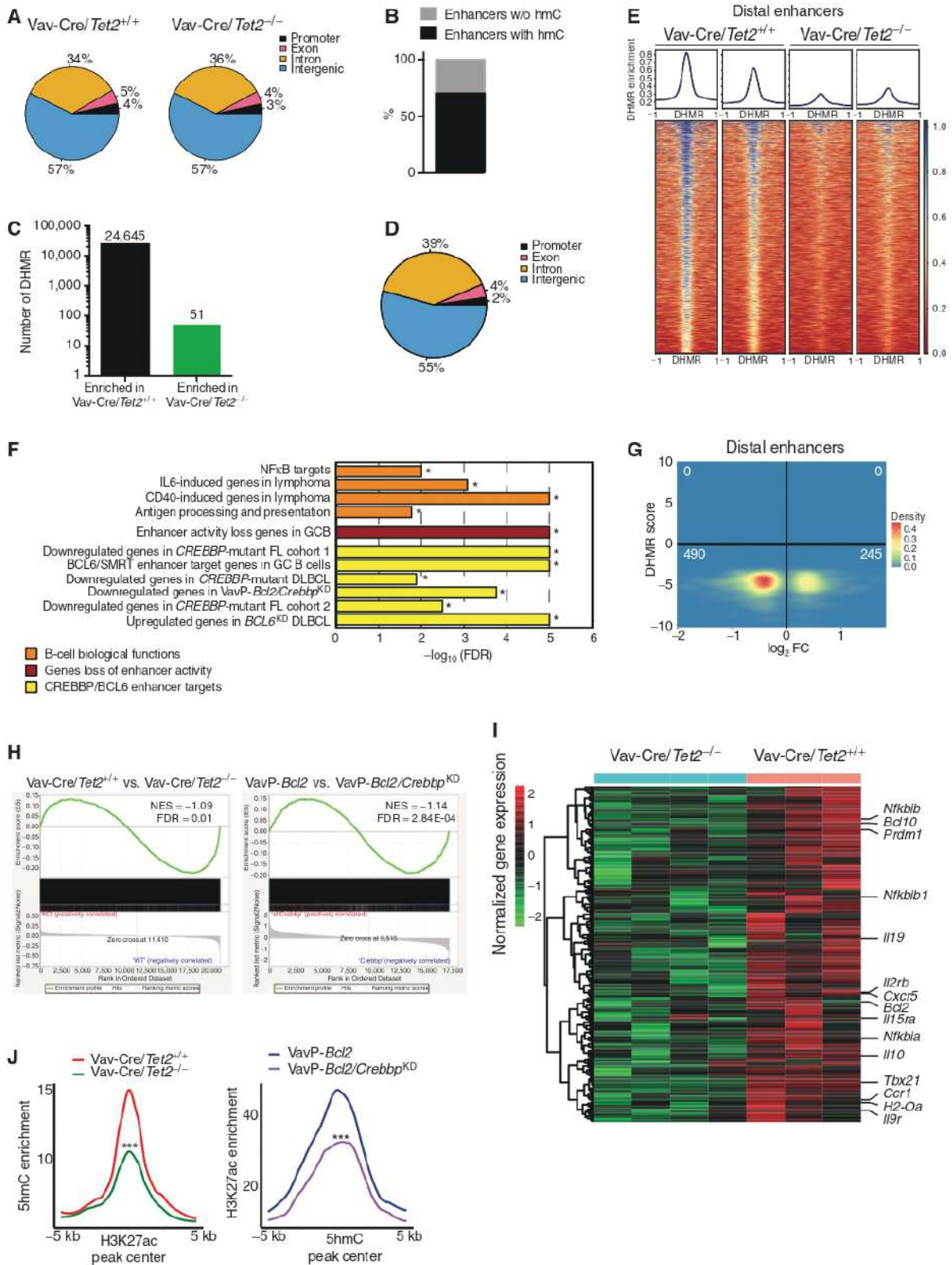
To determine the function of genes affected by reduction of 5hmC, we performed GSEA against our database of gene sets. Similar to what we observed by gene-expression profiling, there was significant enrichment for gene sets involved in GC exit and PC differentiation, including NFκB, IL6, CD40, and

antigen presentation (Fig. 5F). From the mechanistic perspective, hypo-DHMRs were significantly associated with enhancers that are repressed in GC B cells and induced during GC exit (Fig. 5F). Enhancers with 5hmC loss in Vav-Cre/*Tet2*^{-/-} cells were enriched for genes that are repressed in *CREBBP*-deficient lymphomas in mice, as well as *CREBBP*-mutant human patients with FL and DLBCL (Fig. 5F). There was also the reciprocal significant enrichment for BCL6-SMRT target enhancers and genes induced upon BCL6 knockdown (Fig. 5F). Similar results were obtained using a hypergeometric test (Supplementary Fig. S6B). These data suggest that genes downregulated in Vav-Cre/*Tet2*^{-/-} GC B cells are also affected by loss of enhancer 5hmC. To determine how robust this association was, we performed an integrative analysis for 5hmC FC and gene-expression log₂-transformed FC between Vav-Cre/*Tet2*^{-/-} and Vav-Cre/*Tet2*^{+/+} GC B cells. This analysis revealed significant association of distal enhancer (up to 100 Kb) or intronic reduction in 5hmC with transcriptional repression (Fig. 5G, $P = 0.016$; Supplementary Fig. S6C, $P = 0.009$; χ^2 test). Thus, TET2 loss of function results in focal reduction of enhancer hydroxymethylation, which is linked to transcriptional repression of the corresponding genes.

The data also suggest that TET2 and *CREBBP* either cooperate or are mutually dependent on each other to activate gene enhancers that regulate exit from the GC reaction. To address this point, we first analyzed the genes with distal enhancers containing DHMRs shown in Fig. 5E. We found that 60% of those genes (426 of 703) were also associated with H3K27ac signal loss in *CREBBP*-deficient B cells ($P = 5 \times 10^{-31}$; hypergeometric test), indicating that there was a significant correlation between genes containing DHMRs due to TET2 deficiency and those losing H3K27ac in the absence of *CREBBP*. To investigate this putative association further, we identified the peaks that lose both 5hmC and H3K27ac in Vav-Cre/*Tet2*^{-/-} and *Crebbp* knockdown murine B cells (11), respectively. We found 2,619 peaks with at least 35% reduction of both signals. Only 0.43% of all promoters overlapped with both 5hmC and H3K27ac peaks losing signal, whereas 6.1% of all enhancers overlapped with both 5hmC and H3K27ac signal loss ($P = 0$, χ^2 test; Supplementary Fig. S6D). We then performed GSEA for the genes overlapping those peaks and found that they were significantly skewed toward repression in Vav-Cre/*Tet2*^{-/-} and *Crebbp* knockdown murine B cells (FDR < 0.01 and $P = 2.8 \times 10^{-04}$; Fig. 5H). Analysis of the leading edge of those gene sets identified 253 genes downregulated in both TET2- and *CREBBP*-deficient B cells, including *Prdm1*, NFκB-related genes, and genes involved in GC B-cell migration and differentiation such as *Il9r*, *Cor1*, *Cxcr5*, and *Tlr6* (Fig. 5I; Supplementary Table S2). Analysis of 5hmC and H3K27ac distribution in murine TET2 or *CREBBP* loss-of-function GC B cells confirmed the significant reduction of both signals in these 253 genes (Fig. 5J).

Aberrant Repression of *Prdm1* Contributes to the Phenotype of TET2-Deficient B Cells

Given the defect in PC differentiation manifested by *Tet2*-deficient GC B cells, we next focused on the link between this phenotype and aberrant silencing of the *Prdm1* gene. This locus featured two regions with concordant reduction in overlapping 5hmC and H3K27ac peaks in the absence



of TET2 and CREBBP, respectively (Fig. 6A; orange areas). With the selected cutoff of at least 35% loss of both 5hmC and H3K27ac signals, we did not identify additional regions in the proximity of the *Prdm1* gene affected in this manner (data not shown). To confirm the link between 5hmC and H3K27ac, we performed H3K27ac QChIP and observed a significant reduction in this histone mark in Vav-Cre/*Tet2*^{-/-} compared with Vav-Cre/*Tet2*^{+/+} GC B cells (Fig. 6B). As a second validation, we examined H3K27ac by QChIP at the *Il9r* locus where we had observed overlapping 5hmC and H3K27ac (Supplementary Fig. S6E and S6F). TET2 loss of function has also been associated with DNA hypermethylation. We therefore performed bisulfite sequencing of a region within intron 2 described by Barwick and colleagues (40), as becoming hypomethylated in PB as compared with mature B cells. This analysis detected increased methylation in *Tet2*^{-/-} D8 iGC and D8 iPB compared with *Tet2*^{+/+} cells (Fig. 6C). DNA methylation in intron 2 was then quantified using methylation-sensitive restriction enzyme qPCR. Data are represented as the percentage of methylated cytosine in Fig. 6D. No significant differences in *Prdm1* methylation between *Tet2*^{-/-} and *Tet2*^{+/+}-naïve (D0) B cells were observed, whereas significantly higher percentage of methylated cytosine was observed in sorted *Tet2*^{-/-} D4 iGCB ($P < 0.05$), D8 iGCB ($P < 0.01$), and D8 iPB ($P < 0.01$) cells compared with *Tet2*^{+/+} cells (Fig. 6D). In *Tet2*^{+/+} cells, DNA methylation of this region decreased dramatically between naïve (D0) and D4 iGCB and further decreased in D8 iGCB and iPB. In contrast, DNA methylation in *Tet2*^{-/-} cells decreased slightly between naïve (D0) and D4 iGCB and then remained at equivalent levels in D8 iGCB and iPB. These data suggest that reduction in expression of *Prdm1* might also be linked to DNA hypermethylation. Finally, to determine whether aberrant repression of PRDM1 contributes to the *Tet2*-deficient phenotype, we transduced naïve B cells with retrovirus control (MSCV) or retrovirus expressing *Prdm1* (MSCV-*Prdm1*; Fig. 6E). We analyzed GFP⁺ cells 4 days after culture in the presence of IL21 and observed that ectopic *Prdm1* expression rescued the differentiation of TET2-deficient B cells (Fig. 6F).

TET2 Mutations in Human DLBCLs Manifest the Tet2-Deficient GC Gene Signature and Similarities to CREBBP-Mutant Cases

Differentiation blockade such as that caused by *Tet2* deficiency in GC B cells is a canonical hallmark of DLBCL, and hence we predicted that over time, loss of TET2 would lead to progression of GCs toward a more frankly preneoplastic phenotype. To ascertain whether this was the case, we observed Vav-Cre/*Tet2*^{-/-} and Vav-Cre/*Tet2*^{+/+} control mice for a longer

period of time, during which they were immunized on a monthly basis to induce GC formation repeatedly. Animals were sacrificed at 10 months, which is the time point at which Vav-Cre/*Tet2*^{-/-} animals were beginning to manifest signs of morbidity. We observed splenomegaly in Vav-Cre/*Tet2*^{-/-} mice compared with Vav-Cre/*Tet2*^{+/+} mice based on the spleen-to-body-weight ratio (Supplementary Fig. S7A).

We performed a detailed histologic analysis using B220, the GC-specific marker PNA, and the proliferation marker Ki-67. This analysis showed disruption of splenic architecture and nodular expansion of abnormal PNA⁺ and Ki-67-positive GC-derived B cells in the spleens of Vav-Cre/*Tet2*^{-/-} as compared with Vav-Cre/*Tet2*^{+/+} mice (Supplementary Fig. S7B). Vav-Cre/*Tet2*^{-/-} animals also displayed B cells outside of the spleen with involvement of lungs and liver (Supplementary Fig. S7C). In all cases, T cells (CD3⁺) were intermingled with the B cells, suggesting that full transformation might not have been achieved and/or B cells require T cell signals/support. We further validated these findings by flow cytometry, quantifying the proportion of CD95⁺GL7⁺ GC B cells among the total population of B220⁺ cells. We observed a significant increase in GC B cells in the Vav-Cre/*Tet2*^{-/-} mice as compared with Vav-Cre/*Tet2*^{+/+} mice ($P < 0.05$; Supplementary Fig. S7D). Neoplastic B cells in mice manifest reduced staining for IgD, IgM, CD21, and CD23 (41). Along these lines, Vav-Cre/*Tet2*^{-/-} mice displayed significant reduction in B220⁺IgD⁺IgM⁺ cells as compared with Vav-Cre/*Tet2*^{+/+} mice ($P < 0.01$; Supplementary Fig. S7E) and greater proportion of B220⁺IgD^{lo}IgM^{lo/int} B cells ($P < 0.05$; Supplementary Fig. S7E). Likewise, Vav-Cre/*Tet2*^{-/-} mice featured significant expansion of abnormal B220⁺CD23^{lo}CD21^{lo} B cells ($P < 0.01$; Supplementary Fig. S7F).

To investigate whether *Tet2* loss of function manifested a classic tumor suppressor phenotype for B-cell lymphomas, we next explored whether *Tet2* loss of function would cooperate with a canonical DLBCL oncogene. To avoid the confounding effect of *Tet2* loss in other cell lineages, we crossed the GC-specific *Cγ1-Cre/Tet2*^{-/-} mice with the *IμBcl6* mouse model, which directs constitutive expression of the *Bcl6* proto-oncogene in GC B cells, and develops DLBCL after long latency (41). We compared and contrasted the phenotype of *Cγ1-Cre/Tet2*^{-/-};*IμBcl6* to *Cγ1-Cre/Tet2*^{-/-} or *IμBcl6* only. All mice were immunized on a monthly basis to induce GC formation and elicit the DLBCL phenotype and were sacrificed at 7 months, the time point at which *Cγ1-Cre/Tet2*^{-/-};*IμBcl6* animals showed signs of morbidity but prior to the known timing for development of overt lymphoma in *IμBcl6* mice.

Anatomic analysis of these animals revealed significant and much more severe splenomegaly in *Cγ1-Cre/Tet2*^{-/-};*IμBcl6*

Figure 5. TET2 deficiency results in reduction of enhancer cytosine hydroxymethylation and histone acetylation. **A**, Genomic distribution of 5hmC peaks in GC B cells from Vav-Cre/*Tet2*^{+/+} and Vav-Cre/*Tet2*^{-/-} mice. **B**, Overlap of 5hmC peaks with enhancers in Vav-Cre/*Tet2*^{+/+} GC B cells. **C** and **D**, Number **(C)** and genomic distribution **(D)** of DHMR between Vav-Cre/*Tet2*^{+/+} and Vav-Cre/*Tet2*^{-/-} GC B cells. **E**, Heat map of the 5hmC enrichment in all DHMR overlapping with distal enhancers in Vav-Cre/*Tet2*^{+/+} and Vav-Cre/*Tet2*^{-/-} GC B cells. Region spans 1 kb upstream and downstream from the DHMR peak center. Heat map was generated using deeptools (72). **F**, Minus log 10 of the FDR scores for the enrichment of 11 gene sets involved in the exit of GC B cells from the GC. FDR scores were obtained using hypergeometric test for a total of 1,825 genes with enhancers overlapping with DHMR. **G**, Correlation of differential hydroxymethylation in distal enhancers and differential gene expression between Vav-Cre/*Tet2*^{+/+} and Vav-Cre/*Tet2*^{-/-} GC B cells. **H**, GSEA enrichment plots showing correlation of genes associated with peaks losing 5hmC and H3K27ac signals. **I**, Normalized gene-expression level of 253 leading-edge genes associated with peaks losing 5hmC and H3K27ac signals in Vav-Cre/*Tet2*^{-/-} and VavP-Crabbp^{K0} B cells, respectively. **J**, Aggregation plots of 5hmC and H3K27ac enrichment in peaks losing both signals, associated with 253 genes shown in **I**. ***, $P < 0.001$.

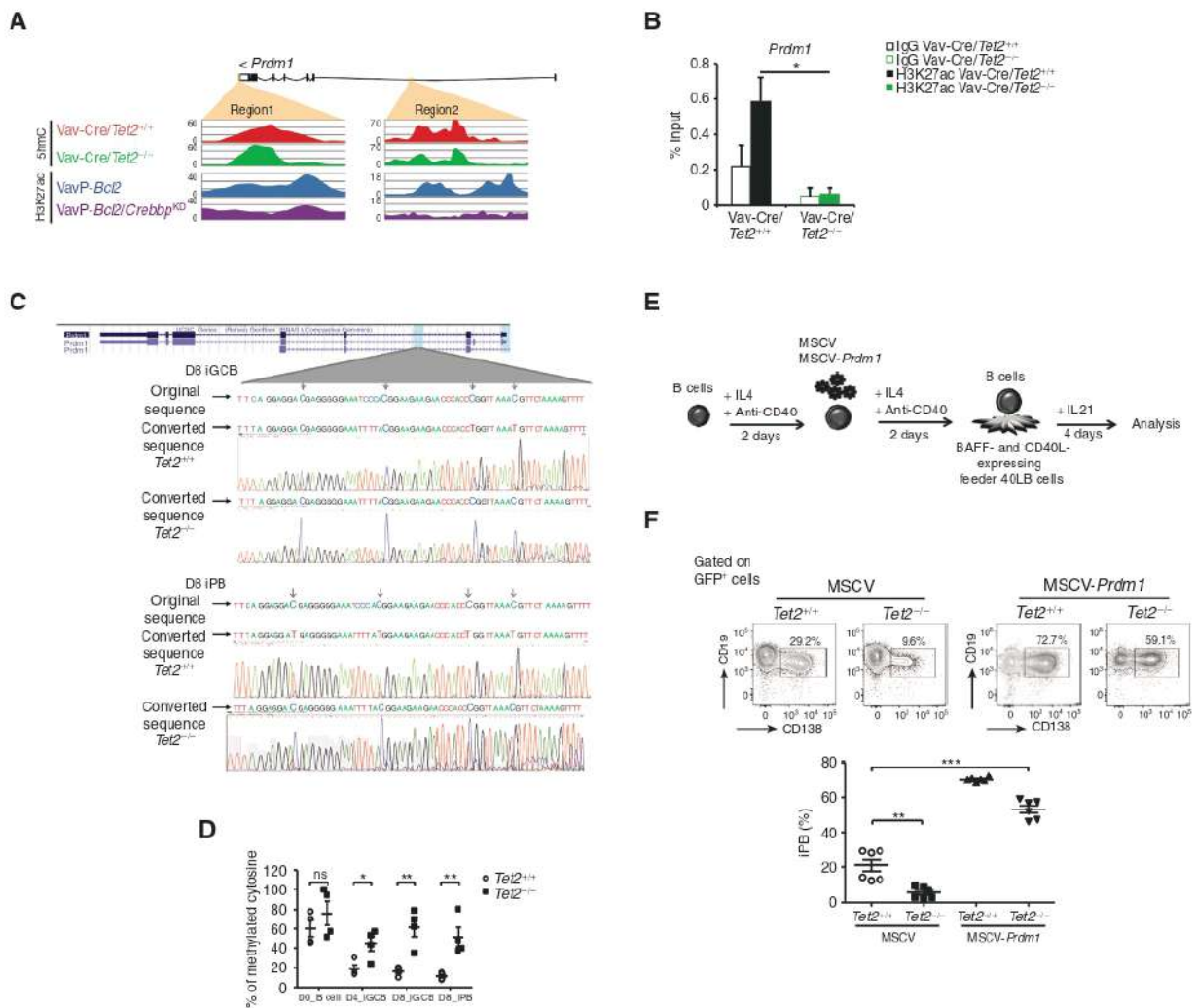


Figure 6. Aberrant repression of *Prdm1* contributes to the phenotype of TET2-deficient B cells. **A**, Read-density tracks of H3K27ac histone modification as well as 5hmC levels in the *Prdm1* locus in *Vav-Cre/Tet2^{-/-}*, *Vav-Cre/Tet2^{+/-}*, *VavP-Bcl2*, and *VavP-Bcl2/Crebbp^{KO}* B cells. Areas marked in orange indicate regions losing 5hmC and H3K27ac signals. **B**, H3K27ac enrichment in region 1 shown in **A**. **C** (top), *Prdm1* DNA methylation status in iGCB cells on D8. Shown are DNA original sequence (top line), representative bisulfite-converted DNA sequence of *Tet2^{+/-}* iGCB cells (middle line), and representative bisulfite-converted DNA sequence of *Tet2^{-/-}* iGCB cells (bottom line). The experiment was performed with three biological replicates. Bottom, *Prdm1* DNA methylation status in iPB on D8. Shown are DNA original sequence (top line), representative bisulfite-converted DNA sequence of *Tet2^{+/-}* iPB (middle line), and representative bisulfite-converted DNA sequence of *Tet2^{-/-}* iPB (bottom line). The experiment was performed with two biological replicates. **D**, Quantification of DNA methylation by methylation-sensitive restriction enzyme qPCR. Differentially methylated regions represented as the percentage of methylated cytosine in the *Prdm1* gene were validated in day 0 (D0, naive) B cells, day 4 (D4) iGCB cells, day 8 (D8) iGCB cells and day 8 (D8) iPB cells ($n = 4$). **E**, Schematic illustration of the ectopic expression of PRDM1 in B cells. Briefly, naive B cells were treated with IL4 and anti-CD40 for 2 days and were infected with a retrovirus expressing GFP or PRDM1-GFP. After 2 more days of culture with IL4 and anti-CD40, cells were harvested and cultured in the presence of 40LB feeder cells and IL21 for 4 days. **F**, GFP-positive cells were analyzed for CD19 and CD138 expression. Representative plots of FACS analysis of the *in vitro* induced CD19⁺CD138⁺ population. Histograms represent the percentage of iPB (CD19⁺CD138⁺); symbols represent triplicate of 2 independent experiments. *, $P < 0.05$; **, $P < 0.005$; ***, $P < 0.0005$.

mice as compared with the *Cy1-Cre/Tet2^{-/-}* or *I μ Bcl6* cohorts (Supplementary Fig. S7G and S7H). We performed a detailed histologic analysis of the spleens of these mice using hematoxylin and eosin (H&E) staining, as well as IHC for B220, CD3, and the proliferation marker Ki-67. Strikingly, 100% of *Cy1-Cre/Tet2^{-/-}*; *I μ Bcl6* mice ($n = 12$) displayed either follicular or diffuse lymphoma, which frequently led to the total effacement of splenic architecture. In these mice, B220⁺

cells displayed a high proliferation index (>50% of cells Ki-67 positive), and CD3⁺ T cells were frequently mixed among neoplastic cells (Fig. 7A; Supplementary Fig. S7I). In marked contrast, only 50% of *I μ Bcl6* mice ($n = 6$) displayed enlarged spleens, with less severe architectural effacement due to enlarged follicles and/or diffuse lymphoid infiltrates of intermediate/large B220⁺ elements. Ki-67⁺ cells ranged from 20% to 40% of the total cells (Fig. 7A; Supplementary

Fig. S7I). Age-matched $C\gamma 1-Cre/Tet2^{-/-}$ mice ($n = 7$) displayed no evidence of lymphomas and had normal splenic architecture, with well-defined primary and secondary follicles (Fig. 7A; Supplementary Fig. S7I). These results were confirmed using flow cytometry. B cells in the spleens of all $C\gamma 1-Cre/Tet2^{-/-};\mu Bcl6$ mice were mostly neoplastic ($B220^{+}CD23^{lo}CD21^{lo}$), compared with $\mu Bcl6$ ($P < 0.05$;

Fig. 7B) or $C\gamma 1-Cre/Tet2^{-/-}$ mice ($P < 0.0001$; Fig. 7B). This result correlated with a decrease in the proportion of normal splenic follicular B cells in $C\gamma 1-Cre/Tet2^{-/-};\mu Bcl6$ mice compared with $\mu Bcl6$ ($P < 0.01$; Fig. 7B) and $C\gamma 1-Cre/Tet2^{-/-}$ mice ($P < 0.0001$; Fig. 7B). All together, these results demonstrate that TET2 loss of function results in accelerated formation of GC-derived lymphomas.

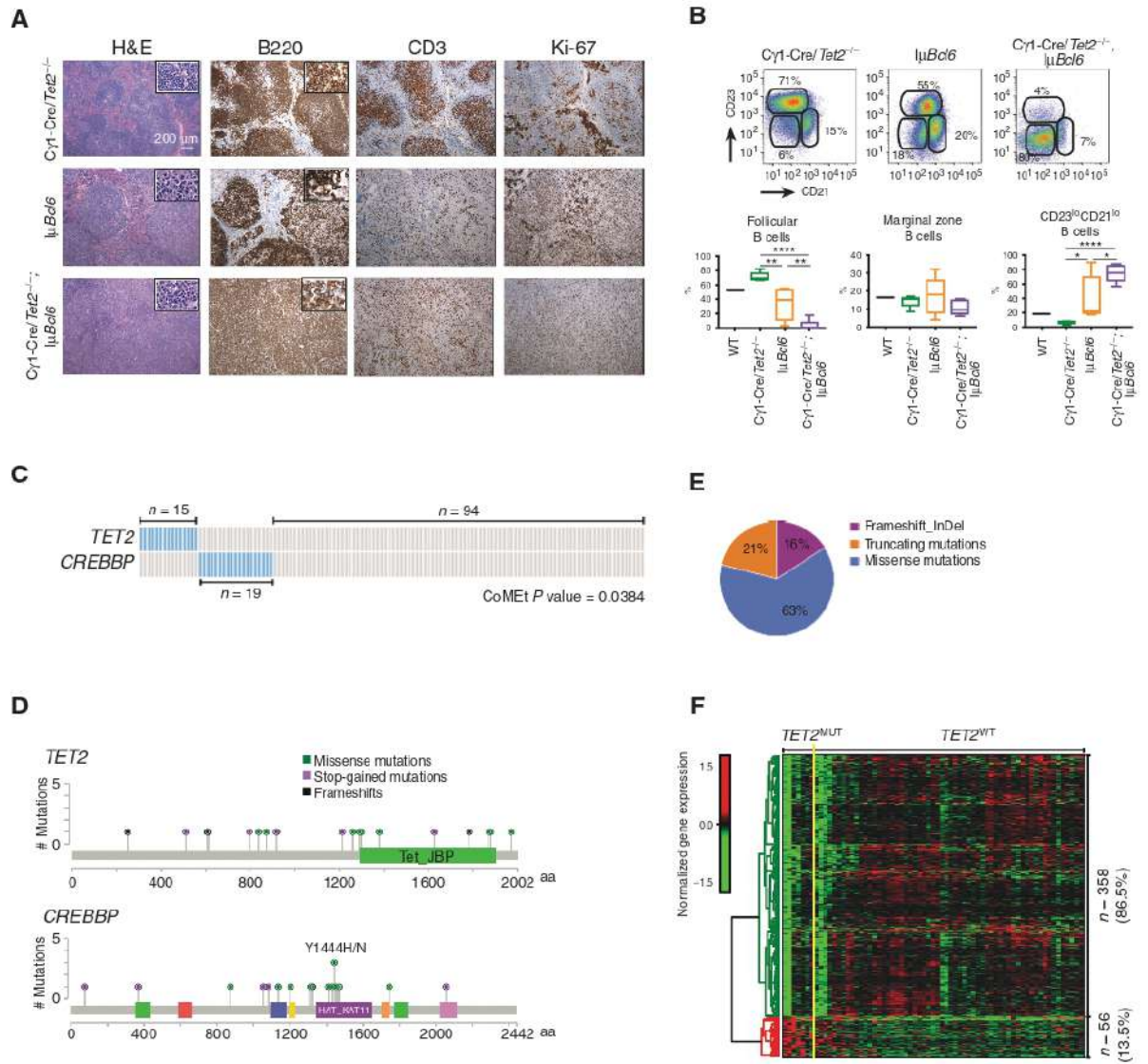


Figure 7. TET2 mutations in human DLBCLs manifest the Tet2-deficient GC gene signature, and similarities to CREBBP-mutant cases. **A**, Representative histologic sections of formalin-fixed, paraffin-embedded spleens from $C\gamma 1-Cre/Tet2^{-/-};\mu Bcl6$, $C\gamma 1-Cre/Tet2^{-/-}$ and $\mu Bcl6$ mice. Sections were stained with H&E and antibodies specific for B220, CD3, and Ki-67. **B**, Representative flow cytometry plot and quantification of CD23⁺CD21⁺ (follicular B cells), CD23⁺CD21⁻ (marginal zone B cells), and CD23^{lo}CD21^{lo} B cells in the spleens of $C\gamma 1-Cre/Tet2^{-/-};\mu Bcl6$, $C\gamma 1-Cre/Tet2^{-/-}$ and $\mu Bcl6$ mice; two-tailed t test; * $P < 0.05$; ** $P < 0.01$; **** $P < 0.0001$. **C**, Mutation matrix for TET2 and CREBBP mutations in a cohort of 128 DLBCL. Blue color indicates any type of mutation, whereas gray indicates no known mutations of the gene. Each of the 128 columns indicates one patient sample. Mutual exclusivity on TET2 and CREBBP mutations was computed using CoMet. **D**, Lollipop of the TET2 and CREBBP mutations, plotted using MutationMapper (73, 74). **E**, Distribution of insertions/deletions and mutations among TET2-mutant DLBCL. **F**, Normalized gene-expression levels of 414 DEGs between TET2-mutant ($TET2^{MUT}$) and nonmutant ($TET2^{WT}$) DLBCL. $TET2^{MUT}$ indicates 7 samples with either stop-codon or frameshift mutation, whereas $TET2^{WT}$ indicates 60 samples without any kind of mutation in TET2 or CREBBP genes. Expression levels were scaled by row. (continued on next page)

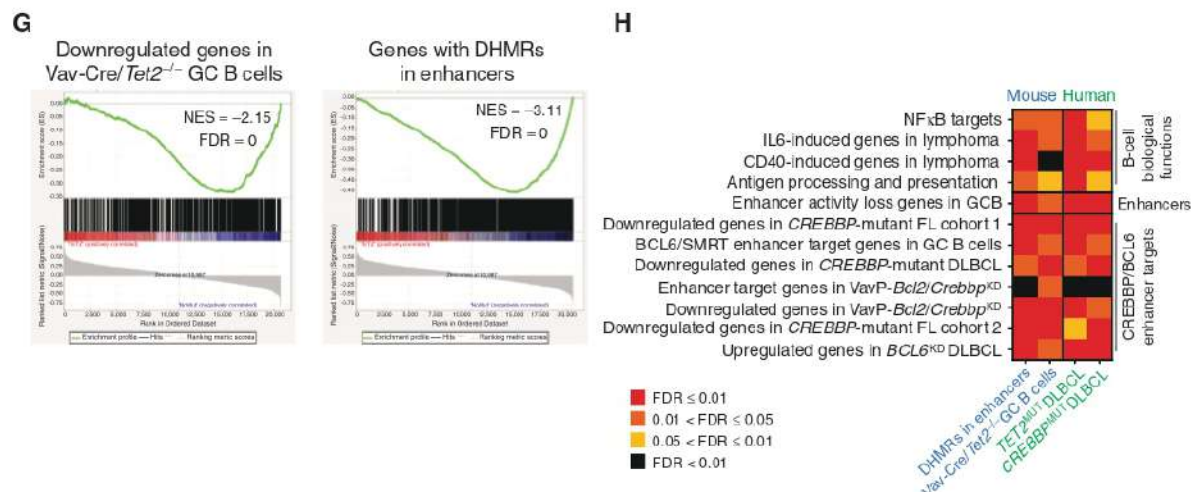


Figure 7. (Continued) G, GSEA enrichment plots showing correlation of different gene sets with ranked expression change between $TET2^{MUT}$ and $TET2^{WT}$ DLBCL. NES, normalized enrichment score; FDR, false discovery rate. **H**, Heat map of the FDR scores for the enrichment of 12 gene sets involved in the exit of GC B cells from the GC. FDR scores were obtained using hypergeometric test for mouse genes with enhancers overlapping DHMR, described as "DHMR in enhancers" and shown previously in Fig. 6F. Column described as "Vav-Cre/ $Tet2^{-/-}$ GC B cells" presents the FDR scores obtained using GSEA, also shown in Fig. 5D. Columns " $TET2^{MUT}$ DLBCL" and " $CREBBP^{MUT}$ DLBCL" indicate GSEA-based FDR scores of the enrichment of the gene sets in DLBCL.

A second link to human lymphomas lies in the gene-expression profile of *Tet2*-deficient GC B cells, which is highly similar to the signature induced by the well-established lymphoma epigenetic tumor suppressor *CREBBP*. This scenario suggests that these two tumor suppressors have overlapping effects, in which case they would most likely not occur together. To address this point, we performed targeted resequencing of *TET2* and *CREBBP* in a cohort of 128 patients with DLBCL, for whom gene-expression microarrays were also available (42). *TET2* mutations were detected in 12% of the cases ($n = 15$), consisting of frameshift, indels, and missense mutations (Fig. 7C-E; Supplementary Table S3), which is in line with previous studies (6, 18). In addition, 15% of these DLBCLs ($n = 19$) had somatic missense mutations of *CREBBP*, 11 of them affecting the HAT domain (Fig. 7C and D; Supplementary Table S3). We focused on HAT domain mutations because these have been reported to induce the most severe phenotype (43). Notably, we observed that *TET2* mutations were significantly mutually exclusive with *CREBBP* mutations (Fig. 7C; $P = 0.0384$). There was also a trend toward mutual exclusivity in the patient cohort reported by Reddy and colleagues (6), although it was not statistically significant ($P = 0.16$). In addition, we found that *TET2*-mutated cases were predominantly classified as GCB-DLBCL subtype (11/15), similar to what has been reported for *CREBBP*-mutated DLBCL (11).

Because gene-expression arrays were available for these patients, we performed a supervised analysis to determine genes differentially expressed between *TET2*-mutant cases and non-*TET2* and non-*CREBBP*-mutant DLBCLs. Even though DLBCLs are extremely heterogeneous, we were still able to identify a set of 358 genes (86.5%) differentially downregulated and 56 genes (13.5%) differentially upregulated in

TET2-mutant DLBCLs ($FC > 1.5$, $FDR < 0.05$; Fig. 7F). This skewing toward transcriptional repression is similar to what was observed in Vav-Cre/ $Tet2^{-/-}$ GC B cells (Fig. 5C). To determine whether the human *TET2*-mutant signature was similar to the Vav-Cre/ $Tet2^{-/-}$ signature from murine GC B cells, we performed GSEA and observed significant enrichment for genes downregulated in Vav-Cre/ $Tet2^{-/-}$ GC B cells ($FDR = 0$; Fig. 7G). Genes affected by DHMRs in mice were also significantly enriched among genes downregulated in human *TET2*-mutant cases ($FDR = 0$; Fig. 7G).

In addition, we compared and contrasted the relative enrichment of gene sets linked to *TET2* loss of function in the gene-expression profiles of Vav-Cre/ $Tet2^{-/-}$ GC B cells, DHMR profiles of Vav-Cre/ $Tet2^{-/-}$ GC B cells, human *TET2*-mutant DLBCLs, and in human *CREBBP*-mutant lymphomas from our new cohort. These included the GC exit gene sets linked to NF κ B, IL6, CD40, and antigen presentation, the enhancers normally repressed in human GC B cells, the various murine and human *CREBBP*-mutant signatures, and the BCL6-SMRT-repressed enhancers and genes induced upon BCL6 knockdown. Overall, this analysis yielded significant enrichment of almost all of these gene sets across the board in both the murine and human gene-expression profiles (Fig. 7H). Collectively, these data indicate the relevance of the murine *Tet2* knockout experiments to the human disease. These results also underline the link between *TET2* and *CREBBP* in coordinating the GC exit transcriptional program through activation of gene enhancers and suggest that *TET2* and *CREBBP* mutations induce malignant transformation through at least partially overlapping mechanisms of action. Finally, our findings suggest that lymphomas with *TET2* loss of function might become addicted to HDAC3, as we previously demonstrated for *CREBBP*-deficient lymphoma

cells (11). To further investigate this hypothesis, we transduced *TET2* WT DLBCL cell lines (OCI-Ly18 and MD901) with two different lentiviral shRNA for *TET2* or a control shRNA and confirmed *TET2* knockdown by qPCR (Supplementary Fig. S8A and S8B). We then treated the cells with our HDAC3-selective inhibitor (HDAC3i; ref. 44) or vehicle for 4 days and quantified the number of live cells (Annexin V DAPI) in each condition. We observed that *TET2* loss of function made the cells more sensitive to HDAC3 inhibition, as reflected by the increase in growth inhibition after HDAC3i treatment in sh*TET2* compared with shScr (Supplementary Fig. S8C and S8D). Whether patients with *TET2* mutations could benefit from treatment with HDAC3 inhibitors requires further investigation.

DISCUSSION

Herein, we report that *TET2* is a tumor suppressor in DLBCL and not merely a passenger mutation that might be inherited from aberrant HSCs. This work expands on the known role of *TET2* in myeloid malignancies and T-cell malignancies (23, 31, 45). Hence, the transforming effect of *TET2* deficiency goes beyond previous results showing B1 transformation in non-immunologically challenged mice (46) and oncogenic potential in fully mature B2 cells participating in the GC reaction. *Tet2* deficiency resulted in impairment of B cells to undergo terminal differentiation in a B cell-autonomous manner as manifested by the deficient formation of GC-derived antibody secreting cells and reduced titers of high-affinity antibodies. This phenotype is indicative of a defect in the ability of B cells to exit the GC reaction, which is a canonical feature of DLBCL lymphomagenesis (47). Indeed, *TET2* loss of function in GC B cells results in accelerated formation of DLBCLs in cooperation with constitutive expression of *BCL6*, a major oncogene in DLBCL, highlighting the significance of *TET2* mutations in patients with primary DLBCL. The GC reaction is a cyclical process, whereby proliferative centroblasts from the dark zone migrate to the light zone as centrocytes, which interact with TFH cells and follicular dendritic cells (3). Centrocytes that encode for higher-affinity BCRs are selected to return to the dark zone as centroblasts for further proliferation and mutagenesis. Repeated rounds of somatic hypermutation lead to enrichment and selection of B cells with the highest-affinity BCR, which then undergo class switch recombination and differentiate to PC that confer long-term immunity by secreting high-affinity antibodies (3). *TET2* loss of function impairs this process as manifested by failure of GC B cells to undergo these final steps, which indicates that *TET2* is not only a tumor suppressor but is also critically required for B cells to successfully undergo the high-affinity humoral immune response.

Exit from the GC reaction requires the engagement of epigenetic mechanisms that activate enhancers for genes that must be induced to resolve the GC phenotype and mediate the PC phenotype (9, 11, 34). Some of the key signaling pathways driving this process are CD40, NF κ B (downstream of BCR signaling), antigen presentation genes such as MHC class II, and cytokines such as IL6 and others (4, 48). Our data suggest that the defect in class switch recombination and PC

differentiation evident in *Tet2*-deficient GCs is linked to aberrant repression of enhancers that regulate these genes. These enhancers accordingly manifested a reduction in the 5hmC mark, suggesting that *TET2* is required to induce enhancer activation and concordant transcriptional activation of genes in these pathways.

Indeed, we find that *TET2* is part of a more complex enhancer-switching mechanism controlling egress of B cells from the GC reaction. Specifically, GC exit enhancers require histone acetyl-transferase CREBBP, which is disrupted by somatic loss-of-function mutations in DLBCL and is a *bona fide* DLBCL tumor suppressor (11, 12). During the GC reaction, CREBBP target enhancers are repressed and H3K27 deacetylated by the transcriptional repressor BCL6 through its recruitment of SMRT/HDAC3 complexes (11, 34). BCL6/SMRT complexes are then suppressed in the light zone downstream of BCR and CD40 signaling, enabling CREBBP to induce H3K27 acetylation and activation of GC exit enhancers and genes (11, 34, 49–52). *CREBBP* mutations disrupt this balance resulting in sustained repression of these genes and a corresponding defect in GC exit (11, 12). Along these lines, we show that there is striking similarity in the sets of genes aberrantly repressed in *Tet2*- and *Crebbp*-deficient mouse GC-derived B cells and, most importantly, in primary human patients with DLBCL. We find that loss of enhancer 5hmC is associated with loss of H3K27 acetylation, suggesting that CREBBP requires the presence of *TET2* and the 5hmC mark to mediate H3K27 acetylation. A functional link between *TET2* and CREBBP is supported by other studies, which show that CREBBP and p300 can acetylate *TET2*, increasing its enzymatic activity and protein stability (39), and that 5hmC colocalizes with active enhancers bound by p300 (53). The fact that patients with *TET2*-mutant DLBCL manifest similar gene-expression profiles as B cells in the *Tet2*-deficient mouse models supports the relevance of our experimental systems to the human disease. Our data suggest that *TET2* loss of function most likely mediates malignant transformation by suppressing/impairing CREBBP-mediated enhancer acetylation, disrupting GC exit transcriptional programming and causing accumulation of B cells in the light zone, which manifests as GC hyperplasia.

CREBBP mutations are generally described as early lesions in DLBCL and induce a similar phenotype as *Tet2* deficiency with defective PC formation (12, 54). Along these lines, it is notable that *TET2* somatic mutations have been demonstrated to be present in the HSCs of patients with *TET2*-mutant DLBCL and are present at high variant allele frequencies in DLBCL. Hence, *TET2* loss-of-function mutations are the only known lymphoma-driving mutations shown to occur at the HSC level. Even though *TET2* loss of function causes B cell-autonomous effects, our mouse models that disrupt *Tet2* expression in HSCs are likely the most physiologically relevant. *CREBBP* mutations likely occur at a later stage than *TET2* mutations during early B-cell development. Regardless, both *TET2* and *CREBBP* are early hits during lymphomagenesis, which represents an additional similarity between these lesions. The fact that *CREBBP* and *TET2* mutations tend to be mutually exclusive in patients with DLBCL lends further support to this model, consistent with these being pathway mutations with a shared mechanism of action. *TET2* loss of

function can also lead to gain of cytosine methylation (18), although this effect is less pronounced than loss of 5hmC (26, 55, 56). Although the loss of 5hmC seems to adequately explain the aberrant transcriptional profiles of *TET2*-deficient lymphomas, the potential contribution of 5mC should not be excluded and is worth exploring in future studies, as suggested by our analysis of the *Prdm1* locus.

In conclusion, we propose that *TET2* be considered a *bona fide* tumor suppressor in DLBCL and included in targeted resequencing panels used to characterize mature B-cell lymphomas. Perturbation of chromatin mechanisms is a hallmark of DLBCL, with >50% of the recently identified 150 recurrent somatic mutations affecting either TFs or epigenetic modifiers (6). *TET2* mutation thus joins the set of alleles in DLBCL that disrupt the GC reaction through perturbations in chromatin modifications, a finding with implications for design and interpretation of studies testing the impact of epigenetic modifier drugs in patients with DLBCL. In this regard, our results showing sensitivity to HDAC3 inhibition of *TET2*-deficient lymphomas suggest that this new therapeutic approach should be tested in patients with *TET2* mutations.

METHODS

Animal Models

Vav-Cre/*Tet2*^{fl/fl} and *Tet2*^{fl/fl} mice were a generous gift from R. Levine, Memorial Sloan Kettering Cancer Center (45). Cγ1-Cre mice [B6.129P2(Cg)-*Ighg1*^{tm1(Cre)Cgn}/J; stock number 010611] and CD19-Cre mice [B6.129P2(C)-*Cd19*^{tm1(Cre)Cgn}/J; stock number 006785] were obtained from The Jackson Laboratory. *IqBcl6* mice for the lymphoma studies were obtained from R. Dalla-Favera (41), Columbia University. All mice were followed until any one of several criteria for euthanizing were met, including severe lethargy and more than 10% body-weight loss in accordance with our Weill Cornell Medicine Institutional Animal Care and Use Committee-approved animal protocol (protocol #2011-0031). All animals were maintained according to the guidelines of the Research Animal Resource Center of Weill Cornell Medicine.

Mice carrying constitutively inactivated *Tet2* alleles (*Tet2*^{-/-}) and CD19-Cre/*Tet2*^{-/-} mice have been previously described (31, 46). These mice, including WT C57BL/6 mice, were maintained in the breeding facility of SCEA, Gustave Roussy (Villejuif, France) and were sacrificed at the indicated times. Experiments with *Tet2*^{-/-} mice were conducted according to the Gustave Roussy Institutional guidelines and authorized by the Direction Départementale des Services Vétérinaires du Val de Marne.

DLBCL Patient Samples

A cohort of 128 tumors with a pathologic diagnosis of DLBCL were interrogated by targeted sequencing, as previously described (42). Informed written consent was obtained from all the subjects. In brief, 500 ng to 1 μg of tumor DNA was sheared by sonication using a Covaris S2 instrument, and end repaired, A-tailed, and ligated with Illumina TruSeq adaptors (Bio-O Scientific) using Kapa HyperPrep Kits (Kapa Biosystems). Adapter ligated products were enriched using 6 cycles of PCR, and 12-plexed for hybrid capture using a custom hybrid capture reagent (Nimblegen) that included all coding exons for *CREBBP* and *TET2*. Each pool was sequenced with 100 bp paired-end reads on a single lane of a HiSeq 2500 Instrument in high output mode at the Hudson Alpha Institute for Biotechnology. Raw sequencing reads were aligned to the human genome (hg19) using a BWA-Mem (57), realigned around InDels using GATK, sorted and

deduplicated using Picard tools, and variants were called according to a consensus between VarScan2 (58) and GATK Unified Genotyper (59). This approach has been validated to have a specificity of 92.9% and a sensitivity of 86.7% (43). Average on-target rate for this data set was 88% and average depth of coverage 641× (min = 263×, max = 1396×). Affymetrix U133 plus 2 gene-expression microarrays were performed on 84 matched DLBCL tumors previously [ref. 60; Gene Expression Omnibus (GEO) Accession, GSE10846]. Research on all human subjects was approved by the respective institutional review boards.

Mouse B-cell Isolation

For the isolation of GC B cells, Vav-Cre/*Tet2*^{-/-} and Vav-Cre/*Tet2*^{+/-} mice were immunized intraperitoneally with SRBCs (1 × 10⁸ cells per mouse) or NP-CGG ratio 20–25 (Biosearch Technologies) in alum (1:1) to induce GC formation. Mice were sacrificed 10 days after immunization, spleens were dissected, and mononuclear cells were purified using Histopaque gradient centrifugation. Cell suspensions were enriched in B cells by positive selection with anti-B220 magnetic microbeads and GC B cells (B220⁺GL7⁺FAS⁺DAPI⁻) were sorted using a BD FACSAria II or a BD Influx sorter (BD Biosciences).

In Vitro Generation of iGCB and iPB

Splenic murine B cells were isolated from 2- to 3-month-old mice using the mouse untouched B-cell isolation kit [CD43 (Ly-48) MicroBeads, Miltenyi Biotec] according to the manufacturer's instructions (yield > 90% CD19⁺ B220⁻). 40LB cells expressing CD40L and producing BAFF were a kind gift of D. Kitamura (33) and cultured in DMEM media with 10% FBS and penicillin G/streptomycin.

For B-cell culture, RPMI-1640 medium (Thermo Fisher Scientific) was supplemented with 10% FCS, 0.055 mmol/L 2-ME, 10 mmol/L HEPES, 1 mmol/L sodium pyruvate, 100 units/mL penicillin, and 100 μg/mL streptomycin (GIBCO). Purified B cells (110 × 10³ cells per well) were cultured in a 6-well plate in the presence of 40LB cells (0.5 × 10⁶ cells per well) that had been irradiated with 80 Gy of γ-ray. rIL4 (1 ng/mL; PeproTech) was added to the primary culture for 4 days, and on day 4, the cells were replated onto a new irradiated feeder layer and cultured with rIL21 (10 ng/mL; PeproTech). Cells were harvested at days 4 (D4) or 8 (D8) and processed for flow-cytometric analysis and RNA or protein isolation.

For viral transduction, naive B cells were first stimulated by anti-CD40 antibody (3 μg/mL; COGER) and rIL4 (25 ng/mL; PeproTech), transduced after 48 hours, and plated on 40LB cells 48 hours after.

A *PRDM1* (human) cDNA fragment was obtained from Genescript and was subcloned in an MSCV-IRES-GFP retrovirus backbone. Viral particles were generated as described (61).

DLBCL Cell Cultures

Human DLBCL cell lines were grown in medium containing RPMI-1640, 10% FBS, 1% HEPES, 1% glutamine, and 1% penicillin/streptomycin. OCI-Ly18 was obtained from the Ontario Cancer Institute in June 2011, and MD901 was provided by Jose Angel Martinez-Climent (CIMA, Pamplona, Spain) in June 2011. Cell line authentication testing was performed at IDEXX BioResearch (<http://www.idexxbioresearch.com/cellcheck>), using methods recommended by the American National Standards Institute (ANSI ASN-0002-2011). The cell lines were confirmed to be of human origin and tested for evidence of cross-species contamination (mouse, rat, Chinese hamster, and African Green monkey). Short tandem repeat testing was performed, and the genetic profile obtained was compared with the established cell line profile to confirm the cell lines are consistent with the established profile. These cell lines have also been routinely tested for *Mycoplasma* contamination in the laboratory. OCI-Ly18 and MD901 cells were transduced with shRNAs against *TET2*, or Scr as

control, for 4 days and then treated with a selective HDAC3 inhibitor, BRD3308 (20 μ mol/L), or vehicle (DMSO) for 96 hours. The shRNA sequences to knockdown human TET2 were shTET2#1: CCTCAA GCATAACCCACCAAT; shTET2#2: CCTTATAGTCAGACCATGAAA.

ELISA and ELISPOT

For analysis of T cell-dependent antibody production, mice were immunized intraperitoneally with 100 μ g NP-CGG ratio 20–25 in alum (1:1). On day 35 after immunization, serum was collected, and titers of isotype-specific antibodies to NP were measured in plates coated with NP₂₆-BSA or NP₇-BSA with the SBA Clonotyping System, according to the manufacturer's protocol (5300-05; Southern Biotechnology). Results were assessed by spectrophotometric measurement of absorbance at 405 nm using a Biotek Synergy Neo Alpha Plate Reader (BioTek). Background readings of absorbance in negative control wells were <0.050. For enzyme-linked immunospot (ELISPOT) assay, we collected bone marrow cells on day 60 after immunization. Cells were incubated for 20 hours at 37°C on NP₂₆-BSA-coated or NP₇-BSA-coated 96-well MultiScreen-HA filter plates (Millipore). NP-specific spots were visualized with goat antibody to mouse IgG1 (1034-05) or IgM (1021-05) conjugated to horseradish peroxidase (HRP; Southern Biotechnology), and color was visualized by the addition of 3,3',5,5'-tetramethylbenzidine (Southern Biotechnology). Plates were evaluated using an automated Zeiss ELISPOT reader system (ZellNet Consulting, Inc.).

Flow Cytometry Analysis

Flow cytometry analysis of spleen cell suspensions was performed using the following fluorescent-labeled anti-mouse antibodies: APC-conjugated anti-B220 (BD Biosciences; #553092), anti-CD138 (BD Biosciences; #558626), and anti-IgM (eBioscience; #17-5790-82); PE-conjugated anti-IgD (BD Biosciences; #558597), anti-CD23 (eBioscience; #5010271), and anti-CD95 (BD Biosciences; #554258); PE-Cy7-conjugated anti-CD21 (BioLegend; #123420), anti-CD95 (BD Biosciences; #557653), anti-GL7 (BD Biosciences; #561530), and anti-CD86 (BD Biosciences; #560582); PE-Vio770-conjugated anti-B220 (Miltenyi Biotec; #130-102-308); FITC-conjugated anti-CXCR4 (BD Biosciences; #551967) and anti-IgG1 (eBioscience; #11-4011-85); Brilliant Violet 421-conjugated anti-CD138 (BioLegend; #142507); PerCP-Cy5.5-conjugated anti-GL7 (BioLegend; #144610), anti-CD138 (BioLegend; #142510), and anti-IgM (BD Biosciences; #550881); APC-Cy7-conjugated anti-CD19 (eBioscience; #47-0193-82), and AlexaFluor647-conjugated anti-BLIMP1 (BD Biosciences; #563643). Sytox blue (Thermo Fisher Scientific) or DAPI was used for the exclusion of dead cells. Data were acquired on a BD FACSCanto II (BD Biosciences) and analyzed using FlowJo v10.1 software (TreeStar).

Histology and IHC

Mouse organs were fixed in 4% formaldehyde and embedded in paraffin. Sections (4 μ m) were deparaffinized and heat antigen-retrieved in citrate buffer pH 6.4, and endogenous peroxidase (HRP) activity was blocked by treating the sections with 3% hydrogen peroxide in methanol. Indirect IHC was performed with antispecies-specific biotinylated secondary antibodies, followed by avidin-horse-radish peroxidase or avidin-AP and developed by Vector Blue or DAB color substrates (Vector Laboratories). Sections were counterstained with hematoxylin if necessary. The following primary antibodies were used: biotin-conjugated anti-B220 (BD Biosciences; 550286), anti-CD3 (Abcam; ab16669), anti-PNA (Vector Laboratories; B1075), and anti-Ki67 (Abcam; ab16667). Photomicrographs were scanned using Aperio eSlide Manager (Leica Biosystems).

Genomic DNA and RNA Extraction

Murine gDNA was extracted using the Puregene Genra cell kit (Qiagen) and eluted in TE. The quality of purified DNA was checked

using an Agilent 2100 Bioanalyzer (Agilent Technologies). Total RNA was extracted from murine GC B cells or tumors using TRIzol (Life Technologies) and RNeasy Mini Kit (Qiagen) with DNase treatment. RNA concentration was determined using Qubit (Life Technologies), and integrity was verified using Agilent 2100 Bioanalyzer.

Targeted Bisulfite DNA Sequencing and Quantification of Methylation

Genomic DNA (200 ng) was modified by bisulfite treatment according to the manufacturer's instructions (MethylDetector, Active Motif). DNA fragments were PCR amplified, and direct sequencing reaction was performed using standard conditions (Applied Biosystems). Primers were: 5'-GTTGATAGTAGAAAGTTGTTTGG and 5'-CTCTAACTCAAATCATTCTATTCT.

The CpG methylation-sensitive digestion of genomic DNA was carried out with the EpiJET DNA Methylation Analysis Kit based on MspI/HpaII digestion (Thermo Scientific) following the manufacturer's instructions using 100 ng purified genomic DNA. qPCR was performed using SYBR Green Select Master Mix (Thermo Scientific) following manufacturer's instructions. The primers for common PCR were as follows: forward: 5'-CTCTGACTCTGGTCTGAAGT-3'; reverse: 5'-GTCTCCTGCTTCGGTGTATC-3'. The methylation level was quantified using the 2^{- Δ CT} method and presented as the percent cytosine methylation at the target-gene site.

Chromatin Immunoprecipitation

Cells were cross-linked with 1% formaldehyde and neutralized with 0.125 mol/L glycine. Cell lysates were sonicated to a length of 300–500 base pairs with a Covaris E-220 Focused-ultrasonicator (Covaris), and proteins were immunoprecipitated with antibody to H3K27ac (Active Motif 39133) or IgG (ab-37415; abcam) as control. After complete washing, immunoprecipitated DNA was eluted in elution buffer and reverse-cross-linked overnight at 65°C. DNA was purified and quantified by real-time PCR. Primer sequences are shown in Supplementary Table S4. Enrichment was calculated relative to input.

Quantitative Real-Time PCR

cDNA synthesis from RNA was performed using the Verso cDNA Synthesis kit (Thermo Scientific). *Tet2* expression was detected using the Green FastMix kit (Applied Biosystems) on a QuantStudio6 Flex Real-Time PCR System (Applied Biosystems) or using TaqMan Universal PCR Master Mix (Applied Biosystems) on a ABI PRISM 7500 (Applied Biosystems). Gene expression was normalized to *Rpl13* or *Hprt* using the Δ C(t) method, and results were represented as mRNA expression. Primer sequences are shown in Supplementary Table S4. All mouse TaqMan probes were purchased from Applied Biosystems: *Pax5* (Mm00435501), *Irf4* (Mm00516431), *Irf8* (Mm00492567), *Xbp1* (Mm00457357), *SpiB* (Mm03048233), *Bcl6* (Mm00477633), and *Prdm1* (Mm00476128). *Polr2a* (Mm00839493) and *Gapdh* (Mm99999915) expression level was used for normalized expression values relative to control using the $\Delta\Delta$ CT method.

Hydroxymethylated DNA Immunoprecipitation Sequencing (hMeDIP-Seq)

A total of 1.5 μ g of genomic DNA was sonicated with a Covaris S220. End repair, A-Tailing, and ligation of the preannealed adapters were performed before immunoprecipitation with the 5hmC antibody (1 μ g/mL; Active Motif 39791). A 10% of the volume was set aside before immunoprecipitation as input. Spike-ins (Diagenode AF-107-0040) were used in order to control the efficiency of the 5hmC pulldown. Amplification of the libraries was performed using Kapa HiFi Hot Start Ready Mix (Kapa Biosystems, KK2601) for 12 cycles. Concentrations and quality of the final DNA libraries was measured by Qubit (Life Technologies) and Agilent 2100 Bioanalyzer, respectively.

Sequencing of the libraries was performed in an Illumina HiSeq2500 multiplexing a single pulldown and a single input on each lane. Reads were trimmed using cutadapt 1.9.1 (75) and aligned by bowtie 2.0.5. Unique reads were extracted using an in-house pipeline. MACS2 (62) peak caller “callpeak” was used to detect broad peaks in 5hmC Vav-Cre/*Tet2*^{-/-} and Vav-Cre/*Tet2*^{+/+} data sets after applying “-nomodel -extsize 200 -B -SPMR -broad -bw 200 -broad-cutoff 0.1” parameters. Subsequently, “bdgdiff” was used from MACS2 to detect differential regions between Vav-Cre/*Tet2*^{-/-} and Vav-Cre/*Tet2*^{+/+} samples. For UCSC visualization purposes, -SPMR normalization was complementarily invoked. For genomic annotation of DHMR regions, we used mouse GRC38 genome from gencode as a reference genome sequence. DHMR regions were annotated for introns, exons, promoters, and distal enhancers. Promoter regions were defined as 2 kb upstream or downstream of transcription start site of the gene. For distal enhancers, we first added a 2 kb upstream promoter region to the gene boundary and then defined the distal enhancer as 100 kb upstream or downstream region around it. For genome annotation of 5hmC peaks, they were overlapped with distal enhancers, introns, and promoter regions. Fold enrichment was averaged for genes containing multiple 5hmC peaks. Separately, transcriptome data sets from Vav-Cre/*Tet2*^{-/-} and Vav-Cre/*Tet2*^{+/+} cells were compared against each other to obtain a list of differential genes at fold-change cutoff of 1.2 and *P* value less than 0.05.

Identification of Peaks Losing 5hmC and H3K27ac Signals

First, the BAM files with the mapped reads for hMeDIP-Seq (5hmC signal) and chromatin immunoprecipitation sequencing (ChIP-seq; H3K27ac signal; ref. 11) were partitioned into 50-bp-long bins and normalized using the approach of Ashoor and colleagues (63) to allow the cross-comparison between replicates. The mean read depth coverage from each knockout or knockdown condition was then compared with the mean signal strength in the corresponding bins in WT condition. Next, all bins fulfilling the following criteria were connected into temporary peaks: (i) minimal number of 30 reads in WT condition, that is, the minimal sum of the reads in all WT replicates; (ii) at least 35% signal loss in knockout or knockdown condition; (iii) the maximum distance between the peaks losing signal cannot exceed 51 bp, and the minimal temporary peak length needs to be composed of at least 4 bins. The above analysis was conducted separately for hMeDIP-Seq and ChIP-seq data using the “merge” and “intersect” functions from the bedtools package (64) in combination with in-house scripts. Next, only the temporary peaks losing 5hmC or H3K27ac signals, which were located not further than 300 bp from each other, were retained and overlapped with the H3K27ac and 5hmC peaks from the corresponding mouse models in WT condition. Finally, the collection of peaks losing both 5hmC and H3K27ac signals was assigned to genes, if located in the region up to 100 kb upstream or downstream around them.

RNA-seq

Libraries were generated using the TruSeq-stranded-mRNA sample kit (Illumina) following the Illumina single-end library preparation protocol. Single-end sequencing (SE50) was performed on Illumina HiSeq2500. RNA-seq results were aligned to mm10 or hg19, respectively, using STAR (65) and annotated to RefSeq using the Rsubread package (66). DEGs were detected using DESeq2 (67) with fold-change cutoff of 1.2 and *P* value less than 0.05.

Microarray Data Analysis

Raw microarray readings were processed using the limma (version 3.30.13) package (68), including extraction from CEL files and normalization with RMA function. For the purpose of the GSEA analysis, normalized expression levels were subsequently exported as GCT files and postprocessed as described in the “GSEA” section below. For the purpose of the differential gene-expression analysis,

microarray probes expression levels were transformed to log₂ values and condensed using averseps function. DEGs were called using lmfFit and eBayes functions. FDR values lower than or equal to 0.05 were considered as significant, considering recomputed FDR scores for DEGs with a minimum of 1.5 FC.

GSEA

All GSEA analyses were conducted using the same settings as in the case of the previously published hematologic cancer studies (69), with the exception of the permute parameter, the value of which was set to “gene_set” as a consequence of dealing with sample sizes smaller than 7. The input files are all the genes’ normalized expression levels from each of the data sets:

- (1) *TET2*^{mut} DLBCL: Seven *TET2*-mutated samples compared against 60 samples with none of these genes mutated.
- (2) *CREBBP*^{mut} DLBCL: Seven *CREBBP*-mutated samples compared against 60 samples with none of these genes mutated.
- (3) Vav-Cre/*Tet2*^{-/-}: 4 Vav-Cre/*Tet2*^{-/-} GC B cell samples were compared against three Vav-Cre/*Tet2*^{+/+} GC B cell samples.

All gene sets used in the GSEA study were based on the previous publications of the Melnick or Straudt (9, 11, 34, 70) laboratories. In addition, the enrichment of these gene sets among the 1,825 genes that enhancers overlap with DHMRs in mice, or 1,294 genes that were downregulated in Vav-Cre/*Tet2*^{-/-} mouse, was conducted using hypergeometric test, followed by Benjamini-Hochberg correction for multiple testing. Gene sets were also tested for the enrichment in the shuffled gene lists as an additional background control.

Statistical Analysis

Mutual exclusivity of *CREBBP* and *TET2* mutations in the human patients with DLBCL cohort was computed using CoMet (71).

Data Deposition

Raw data from RNA-seq and hMeDIP-Seq in murine GC B cells were deposited in GEO under accession number GSE111700.

Disclosure of Potential Conflicts of Interest

M.A. Rubin reports receiving a commercial research grant from Janssen, has received honoraria from the speakers bureau of Roche, and is a consultant/advisory board member for ViridBio and Astellas. R.L. Levine is on the supervisory board of Qiagen and scientific advisory board of Loxo, C4 Therapeutics, and Isoplexis; and is a consultant for and receives research support from Roche and Celgene. A.M. Melnick reports receiving research support from Janssen. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: P.M. Dominguez, H. Ghamlouch, W. Rosikiewicz, R.L. Levine, O.A. Bernard, S. Aoufouchi, S. Li, R. Shaknovich
Development of methodology: P.M. Dominguez, H. Ghamlouch, W. Rosikiewicz, W. Béguelin, M.A. Rivas, P. Pawlikowska, M. Torres-Martin, A.S. Doane, V. Della-Valle, M. Teater, L. Cimmino, S. Tadros, M.A. Rubin, M.E. Figueroa, S. Li, A.M. Melnick
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): P.M. Dominguez, H. Ghamlouch, L. Fontán, M. Armand, E. Mouly, M. Torres-Martin, M. Durant, N. Droin, S. Tadros, S. Motanagh, A.H. Shih, G. Inghirami, M.R. Green
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): P.M. Dominguez, H. Ghamlouch, W. Rosikiewicz, P. Kumar, L. Fontán, P. Pawlikowska, M. Armand, E. Mouly, M. Torres-Martin, A.S. Doane, M. Durant, M. Teater, N. Droin, S. Tadros, O. Elemento, G. Inghirami, M.E. Figueroa, S. Aoufouchi, S. Li, R. Shaknovich, A.M. Melnick

Writing, review, and/or revision of the manuscript: P.M. Dominguez, H. Ghamlouch, W. Rosikiewicz, P. Kumar, W. Tam, G. Inghirami, M.E. Figueroa, O.A. Bernard, S. Aoufouchi, S. Li, R. Shaknovich, A.M. Melnick

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): H. Ghamlouch, W. Rosikiewicz, M.T. Calvo Fernandez, V. Della-Valle, M.A. Rubin, W. Tam, I. Aifantis, G. Inghirami, S. Aoufouchi, S. Li, A.M. Melnick

Study supervision: P.M. Dominguez, O.A. Bernard, S. Aoufouchi, S. Li, R. Shaknovich

Acknowledgments

We thank people from the Epigenomics Core Weill Cornell Medicine (WCM), Genomics Core WCM, Flow Cytometry Core WCM, Laboratory of Comparative Pathology WCM and MSKCC; Edward Holson and Kdac Therapeutics Inc. for the HDAC3 inhibitor; members of the Gustave Roussy Platforms, including Patrick Gonin, Yann Lecluse, Philippe Rameau, and Olivia Bawa; and Daisuke Kitamura (Tokyo University of Science, Japan) for the kind gift of the 40LB cell line. Research reported in this publication was supported by the Office of the Director of the NIH under Award Number S10OD019986 to Hospital for Special Surgery. P.M. Dominguez is supported by a Lymphoma Research Foundation Postdoctoral Fellowship. H. Ghamlouch is supported by a grant from Fondation de France (No 00067113). O. Elemento is supported by NSF CAREER, LLS SCOR, Hirschl Trust Award, Starr Cancer Consortium I6-A618, NIH1R01CA19454. MR. Green is supported by R01 CA201380. O.A. Bernard is supported by the Institut National du Cancer (2013-PLBIO-09, 2016-PLBIO-068 and INCa-DGOS-INSERM 12551), Association Laurette Fugain, and équipe labellisée Ligue Nationale Contre le Cancer. S. Aoufouchi is supported by Comité Val-d'Oise de la Ligue contre le cancer. S. Li is supported by Leukemia Research Foundation Research Grant, The Jackson Laboratory Cancer Center New Investigator Award, and The Jackson Laboratory Director's Innovation Fund. Research reported in this publication was partially supported by the NCI of the NIH under Award Number P30CA034196. A.M. Melnick is supported by the Chemotherapy Foundation, The Follicular Lymphoma Research Consortium, Leukemia and Lymphoma Society SCOR Grant #7012-16, and the Starr Cancer Consortium.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

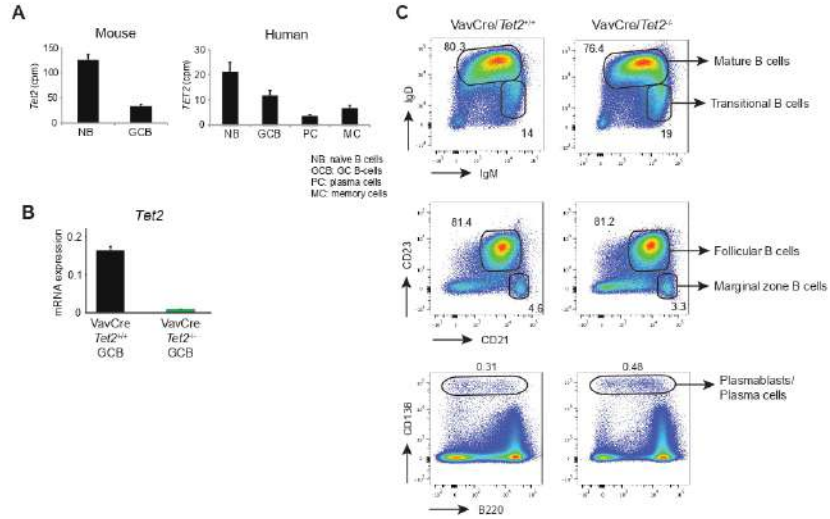
Received June 13, 2018; revised August 26, 2018; accepted September 26, 2018; published first October 1, 2018.

REFERENCES

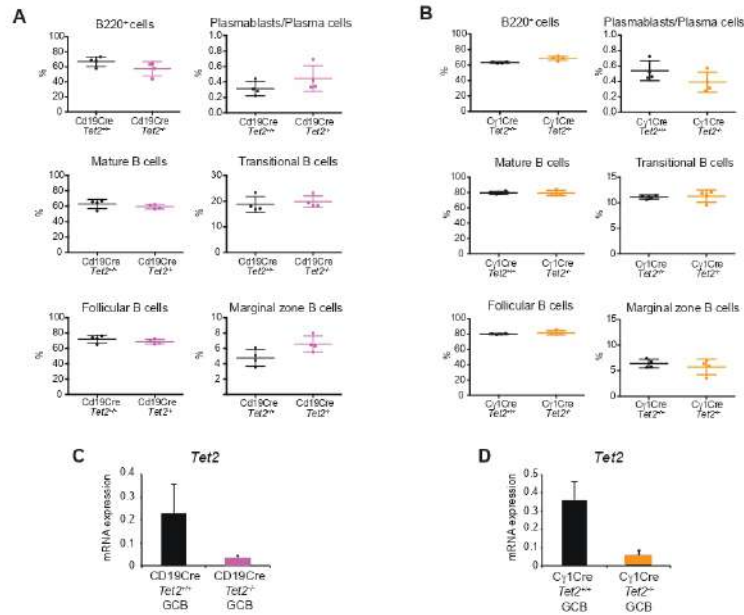
- Coiffier B, Thieblemont C, Van Den Neste E, Lepage G, Plantier I, Castaigne S, et al. Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. *Blood* 2010;116:2040-5.
- De Silva NS, Klein U. Dynamics of B cells in germinal centres. *Nat Rev Immunol* 2015;15:137-48.
- Mesin L, Ersching J, Victora GD. Germinal center B cell dynamics. *Immunity* 2016;45:471-82.
- Klein U, Dalla-Favera R. Germinal centres: role in B-cell physiology and malignancy. *Nat Rev Immunol* 2008;8:22-33.
- Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* 2006;31:89-97.
- Reddy A, Zhang J, Davis NS, Moffitt AB, Love CL, Waldrop A, et al. Genetic and functional drivers of diffuse large B cell lymphoma. *Cell* 2017;171:481-94 e15.
- Beguelin W, Popovic R, Teater M, Jiang Y, Bunting KL, Rosen M, et al. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell* 2013;23:677-92.
- Caganova M, Carrisi C, Varano G, Mainoldi F, Zanardi F, Germain PL, et al. Germinal center dysregulation by histone methyltransferase EZH2 promotes lymphomagenesis. *J Clin Invest* 2013;123:5009-22.
- Ortega-Molina A, Boss IW, Canela A, Pan H, Jiang Y, Zhao C, et al. The histone lysine methyltransferase KMT2D sustains a gene expression program that represses B cell lymphoma development. *Nat Med* 2015;21:199-208.
- Zhang J, Dominguez-Sola D, Hussein S, Lee JE, Holmes AB, Bansal M, et al. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat Med* 2015;21:1190-8.
- Jiang Y, Ortega-Molina A, Geng H, Ying HY, Hatzi K, Parsa S, et al. CREBBP inactivation promotes the development of HDAC3-dependent lymphomas. *Cancer Discov* 2017;7:38-53.
- Zhang J, Vlasevska S, Wells VA, Nataraj S, Holmes AB, Duval R, et al. The CREBBP acetyltransferase is a haploinsufficient tumor suppressor in B-cell lymphoma. *Cancer Discov* 2017;7:322-37.
- Dominguez PM, Teater M, Chambwe N, Kormaksson M, Redmond D, Ishii J, et al. DNA methylation dynamics of germinal center B cells are mediated by AID. *Cell Rep* 2015;12:2086-98.
- Shaknovich R, Cerchiatti L, Tsikitas L, Kormaksson M, De S, Figueroa ME, et al. DNA methyltransferase 1 and DNA methylation patterning contribute to germinal center B-cell differentiation. *Blood* 2011;118:3559-69.
- De S, Shaknovich R, Riester M, Elemento O, Geng H, Kormaksson M, et al. Aberration in DNA methylation in B-cell lymphomas has a complex origin and increases with disease severity. *PLoS Genet* 2013;9:e1003137.
- Pan H, Jiang Y, Boi M, Tabbo F, Redmond D, Nie K, et al. Epigenomic evolution in diffuse large B-cell lymphomas. *Nat Commun* 2015;6:6921.
- Teater M, Dominguez PM, Redmond D, Chen Z, Ennishi D, Scott DW, et al. AICDA drives epigenetic heterogeneity and accelerates germinal center-derived lymphomagenesis. *Nat Commun* 2018;9:222.
- Asmar F, Punj V, Christensen J, Pedersen MT, Pedersen A, Nielsen AB, et al. Genome-wide profiling identifies a DNA methylation signature that associates with TET2 mutations in diffuse large B-cell lymphoma. *Haematologica* 2013;98:1912-20.
- Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* 2011;333:1300-3.
- Abdel-Wahab O, Mullally A, Hedvat C, Garcia-Manero G, Patel J, Wadleigh M, et al. Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood* 2009;114:144-7.
- Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A, et al. Mutation in TET2 in myeloid cancers. *N Engl J Med* 2009;360:2289-301.
- Smith AE, Mohamedali AM, Kulasekararaj A, Lim Z, Gaken J, Lea NC, et al. Next-generation sequencing of the TET2 gene in 355 MDS and CMML patients reveals low-abundance mutant clones with early origins, but indicates no definite prognostic value. *Blood* 2010;116:3923-32.
- Lemonnier F, Couronne L, Parrens M, Jais JP, Travert M, Lamant L, et al. Recurrent TET2 mutations in peripheral T-cell lymphomas correlate with TFH-like features and adverse clinical parameters. *Blood* 2012;120:1466-9.
- Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, et al. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev* 2011;25:679-84.
- Cimmino L, Abdel-Wahab O, Levine RL, Aifantis I. TET family proteins and their role in stem cell differentiation and transformation. *Cell Stem Cell* 2011;9:193-204.
- Rampal R, Alkalini A, Madzo J, Vasanthakumar A, Promier E, Patel J, et al. DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. *Cell Rep* 2014;9:1841-55.

27. Stroud H, Feng S, Morey Kinney S, Pradhan S, Jacobsen SE. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol* 2011;12:R54.
28. Hon GC, Song CX, Du T, Jin F, Selvaraj S, Lee AY, et al. 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol Cell* 2014;56:286–97.
29. Rasmussen KD, Jia G, Johansen JV, Pedersen MT, Rapin N, Bagger FO, et al. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. *Genes Dev* 2015;29:910–22.
30. Cimmino L, Dolgalev I, Wang Y, Yoshimi A, Martin GH, Wang J, et al. Restoration of TET2 function blocks aberrant self-renewal and leukemia progression. *Cell* 2017;170:1079–95 e20.
31. Quivoron C, Couronne L, Della Valle V, Lopez CK, Plo I, Wagner-Ballon O, et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell* 2011;20:25–38.
32. Bunting KL, Soong TD, Singh R, Jiang Y, Beguelin W, Poloway DW, et al. Multi-tiered reorganization of the genome during B cell affinity maturation anchored by a germinal center-specific locus control region. *Immunity* 2016;45:497–512.
33. Nojima T, Haniuda K, Moutai T, Matsudaira M, Mizokawa S, Shiratori I, et al. In-vitro derived germinal centre B cells differentially generate memory B or plasma cells in vivo. *Nat Commun* 2011;2:465.
34. Hatzi K, Jiang Y, Huang C, Garrett-Bakelman F, Gearhart MD, Giannopoulos EG, et al. A hybrid mechanism of action for BCL6 in B cells defined by formation of functionally distinct complexes at enhancers and promoters. *Cell Rep* 2013;4:578–88.
35. Figueroa ME, Lugthart S, Li Y, Erpelinck-Verschueren C, Deng X, Christos PJ, et al. DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer Cell* 2010;17:13–27.
36. Bogdanovic O, Smits AH, de la Calle Mustienes E, Tena JJ, Ford E, Williams R, et al. Active DNA demethylation at enhancers during the vertebrate phylogenetic period. *Nat Genet* 2016;48:417–26.
37. Mahe EA, Madgou T, Serandour AA, Bizot M, Avner S, Chalmel F, et al. Cytosine modifications modulate the chromatin architecture of transcriptional enhancers. *Genome Res* 2017;27:947–58.
38. Yamazaki J, Jelinek J, Lu Y, Cesaroni M, Madzo J, Neumann F, et al. TET2 mutations affect non-CpG island DNA methylation at enhancers and transcription factor-binding sites in chronic myelomonocytic leukemia. *Cancer Res* 2015;75:2833–43.
39. Zhang YW, Wang Z, Xie W, Cai Y, Xia L, Easwaran H, et al. Acetylation enhances TET2 function in protecting against abnormal DNA methylation during oxidative stress. *Mol Cell* 2017;65:323–35.
40. Barwick BG, Scharer CD, Bally APR, Boss JM. Plasma cell differentiation is coupled to division-dependent DNA hypomethylation and gene regulation. *Nat Immunol* 2016;17:1216–25.
41. Cattoretti G, Pasqualucci L, Ballon G, Tam W, Nandula SV, Shen Q, et al. Deregulated BCL6 expression recapitulates the pathogenesis of human diffuse large B cell lymphomas in mice. *Cancer Cell* 2005;7:445–55.
42. Garcia-Ramirez I, Tadros S, Gonzalez-Herrero I, Martin-Lorenzo A, Rodriguez-Hernandez G, Moore D, et al. Crebbp loss cooperates with Bcl2 overexpression to promote lymphoma in mice. *Blood* 2017;129:2645–56.
43. Green MR, Kihira S, Liu CL, Nair RV, Salari R, Gentles AJ, et al. Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. *Proc Natl Acad Sci U S A* 2015;112:E1116–25.
44. Wagner FF, Lundh M, Kaya T, McCarren P, Zhang YL, Chattopadhyay S, et al. An isochemogenic set of inhibitors to define the therapeutic potential of histone deacetylases in beta-cell protection. *ACS Chem Biol* 2016;11:363–74.
45. Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* 2011;20:11–24.
46. Mouly E, Ghamlouch H, Della-Valle V, Scourciz L, Quivoron C, Roos-Weil D, et al. B-cell tumor development in Tet2-deficient mice. *Blood Adv* 2018;2:703–14.
47. Hatzi K, Melnick A. Breaking bad in the germinal center: how deregulation of BCL6 contributes to lymphomagenesis. *Trends Mol Med* 2014;20:343–52.
48. Victora GD, Nussenzweig MC. Germinal centers. *Annu Rev Immunol* 2012;30:429–57.
49. Niu H, Ye BH, Dalla-Favera R. Antigen receptor signaling induces MAP kinase-mediated phosphorylation and degradation of the BCL-6 transcription factor. *Genes Dev* 1998;12:1953–61.
50. Polo JM, Juszczynski P, Monti S, Cerchiotti L, Ye K, Grealley JM, et al. Transcriptional signature with differential expression of BCL6 target genes accurately identifies BCL6-dependent diffuse large B cell lymphomas. *Proc Natl Acad Sci U S A* 2007;104:3207–12.
51. Ranuncolo SM, Polo JM, Dierov J, Singer M, Kuo T, Grealley J, et al. Bcl-6 mediates the germinal center B cell phenotype and lymphomagenesis through transcriptional repression of the DNA-damage sensor ATR. *Nat Immunol* 2007;8:705–14.
52. Saito M, Gao J, Basso K, Kitagawa Y, Smith PM, Bhagar G, et al. A signaling pathway mediating downregulation of BCL6 in germinal center B cells is blocked by BCL6 gene alterations in B cell lymphoma. *Cancer Cell* 2007;12:280–92.
53. Ichiyama K, Chen T, Wang X, Yan X, Kim BS, Tanaka S, et al. The methylcytosine dioxygenase Tet2 promotes DNA demethylation and activation of cytokine gene expression in T cells. *Immunity* 2015;42:613–26.
54. Horton SJ, Girotopoulos G, Yun H, Vohra S, Sheppard O, Bashford-Rogers R, et al. Early loss of Crebbp confers malignant stem cell properties on lymphoid progenitors. *Nat Cell Biol* 2017;19:1093–104.
55. Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* 2010;468:839–43.
56. Shih AH, Jiang Y, Meydan C, Shank K, Pandey S, BarreYRO L, et al. Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. *Cancer Cell* 2015;27:502–15.
57. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013; 1303. <https://arxiv.org/abs/1303.3997>.
58. Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;22:568–76.
59. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–303.
60. Lenz G, Wright G, Dave SS, Xiao W, Powell J, Zhao H, et al. Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med* 2008;359:2313–23.
61. Damm F, Mylonas E, Cosson A, Yoshida K, Della Valle V, Mouly E, et al. Acquired initiating mutations in early hematopoietic cells of CLL patients. *Cancer Discov* 2014;4:1088–101.
62. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 2008;9:R137.
63. Ashoor H, Louis-Brennetot C, Janoueix-Lerosey I, Bajic VB, Boeva V. HMCAn-diff: a method to detect changes in histone modifications in cells with different genetic characteristics. *Nucleic Acids Res* 2017;45:e58.
64. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010;26:841–2.
65. 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.
66. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res* 2013;41:e108.
67. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.

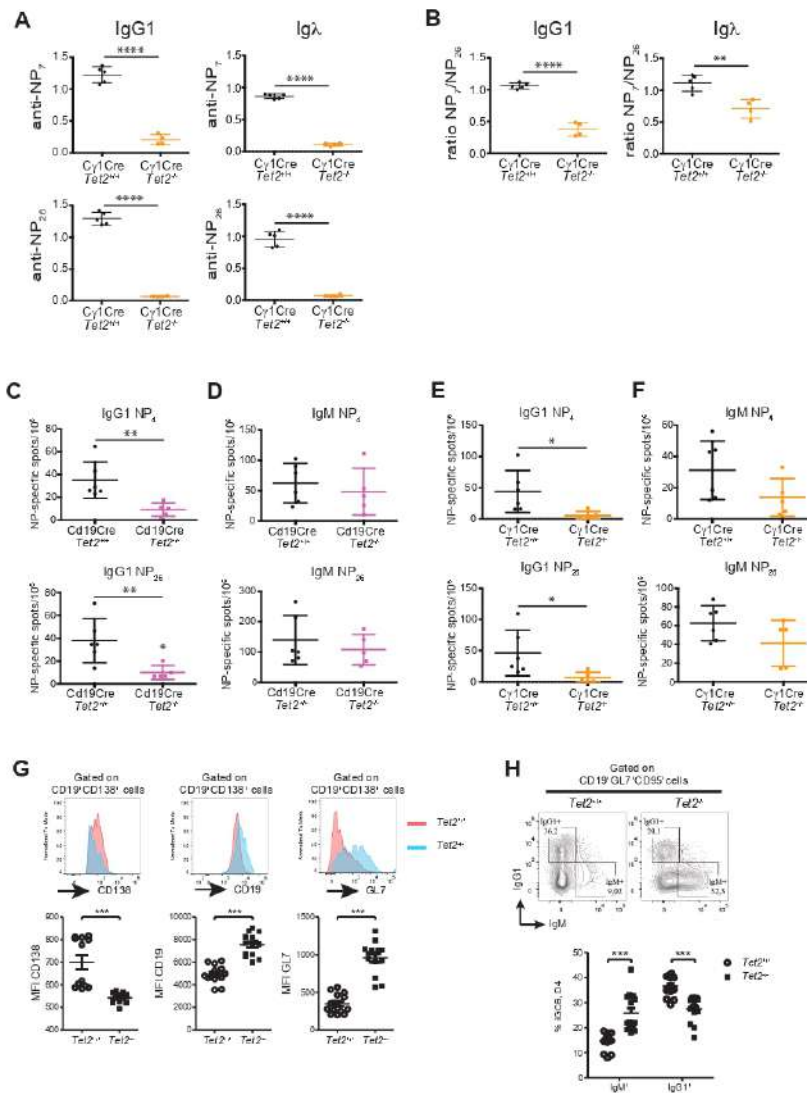
68. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015;43:e47.
69. Armstrong SA, Staunton JE, Silverman LB, Pieters R, den Boer ML, Minden MD, et al. MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. *Nat Genet* 2002;30:41–7.
70. Shaker S, Bernstein M, Momparler RL. Antineoplastic action of 5-aza-2'-deoxycytidine (Dacogen) and depsipeptide on Raji lymphoma cells. *Oncol Rep* 2004;11:1253–6.
71. Leiserson MD, Wu HT, Vandin F, Raphael BJ. CoMEt: a statistical approach to identify combinations of mutually exclusive alterations in cancer. *Genome Biol* 2015;16:160.
72. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 2016;44(W1):W160–5.
73. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013;6:p11.
74. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012;2:401–4.
75. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 2011;17:10–2.



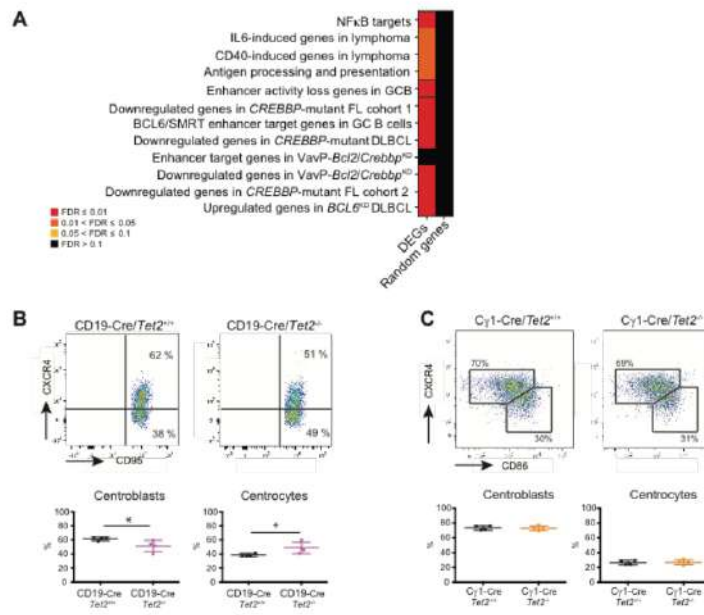
Supplementary Figure 1. **A**, Gene expression levels by RNAseq of *Tet2/TET2* in mouse and human B cells. **B**, Gene expression levels by RT-PCR of *Tet2* in GC B-cells from *Vav-Cre/Tet2^{+/+}* and *Vav-Cre/Tet2^{-/-}* mice. **C**, Representative flow cytometry plots corresponding to total B cells ($B220^+$), mature B cells ($B220^+IgD^+IgM^+$), transitional B cells ($B220^+IgD^{int}IgM^+$), follicular B cells ($B220^+CD23^+CD21^+$), marginal zone B cells ($B220^+CD23^{low}CD21^+$) and plasmablasts/PC ($CD138^+$) in the spleens of *Vav-Cre/Tet2^{+/+}* and *Vav-Cre/Tet2^{-/-}* mice.



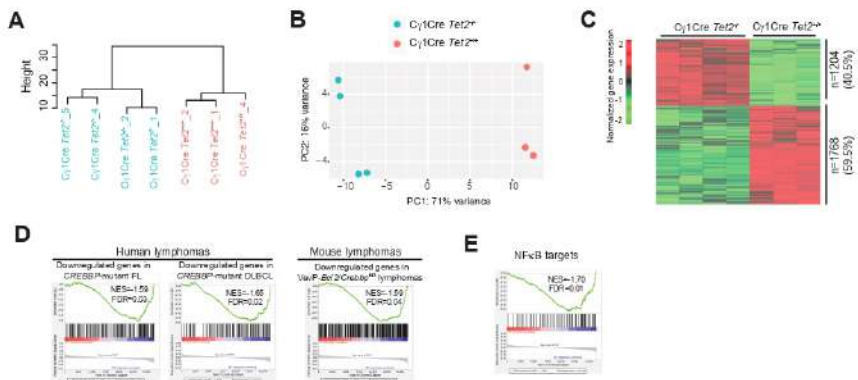
Supplementary Figure 2. **A**, Quantification of flow cytometry data corresponding to total B cells ($B220^+$), mature B cells ($B220^+IgD^+IgM^+$), transitional B cells ($B220^+IgD^{int}IgM^+$), follicular B cells ($B220^+CD23^+CD21^+$), marginal zone B cells ($B220^+CD23^{low}CD21^+$) and plasmablasts/PC ($CD138^+$) in the spleens of *Cd19-Cre/Tet2^{+/+}* and *Cd19-Cre/Tet2^{-/-}* mice. **B**, Quantification of flow cytometry data corresponding to total B cells ($B220^+$), mature B cells ($B220^+IgD^+IgM^+$), transitional B cells ($B220^+IgD^{int}IgM^+$), follicular B cells ($B220^+CD23^+CD21^+$), marginal zone B cells ($B220^+CD23^{low}CD21^+$) and plasmablasts/PC ($CD138^+$) in the spleens of *Cγ1-Cre/Tet2^{+/+}* and *Cγ1-Cre/Tet2^{-/-}* mice. **C**, Gene expression levels by qPCR of *Tet2* in GC B-cells from *Cd19-Cre/Tet2^{+/+}* and *Cd19-Cre/Tet2^{-/-}* mice. **D**, Gene expression levels by qPCR of *Tet2* in GC B-cells from *Cγ1-Cre/Tet2^{+/+}* and *Cγ1-Cre/Tet2^{-/-}* mice.



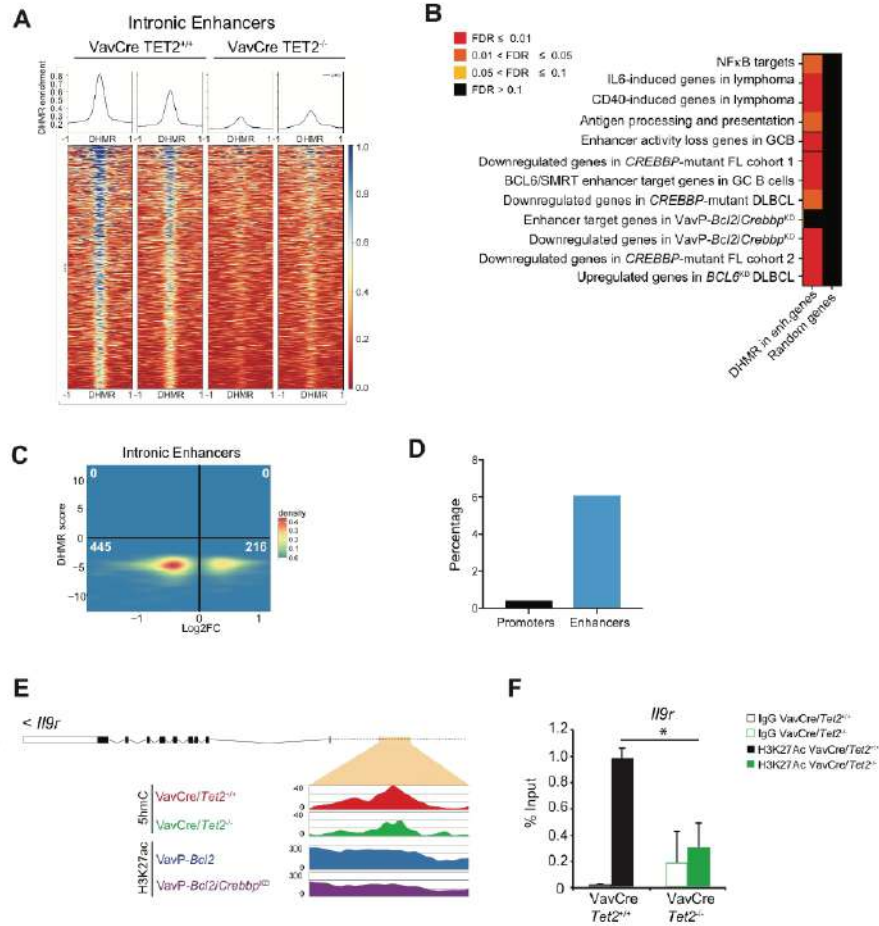
Supplementary Figure 3. A, Twenty one days after immunization, NP-specific antibodies (IgG1 and Igλ) were measured in the sera of $C\gamma 1\text{-Cre}/Tet2^{+/+}$ and $C\gamma 1\text{-Cre}/Tet2^{-/-}$ mice by ELISA. **B**, NP7/NP20 ratio of NP-specific antibodies detected in A. **C, D**, Sixty days after immunization (forty days after boost), NP-specific IgG1 (**C**) and IgM (**D**) secreting cells from the bone marrow of $Cd19\text{-Cre}/Tet2^{+/+}$ and $Cd19\text{-Cre}/Tet2^{-/-}$ mice were quantified by ELISPOT. **E, F**, Sixty days after immunization (forty days after boost), NP-specific IgG1 (**E**) and IgM (**F**) secreting cells from the bone marrow of $C\gamma 1\text{-Cre}/Tet2^{+/+}$ and $C\gamma 1\text{-Cre}/Tet2^{-/-}$ mice were quantified by ELISPOT. two-tailed t test, * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$ **** $p < 0.0001$. **G**, Representative flow cytometry plots analyzing the expression of CD138, CD19 and GL7 on $Tet2^{+/+}$ and $Tet2^{-/-}$ $CD19^+CD138^+$ cells at D8. Histograms represent the median fluorescence intensity (MFI) on D8; symbols represent triplicate of 4 independent experiments. **H**, Flow cytometry plots showing IgM and IgG1 expression by iGCB cells ($CD19^+GL7^+FAS^+$) on D4. Histograms represent the percentage of IgM^+ iGCB cells and $IgG1^+$ iGCB cells on D4; symbols represent triplicate of 5 independent experiments.



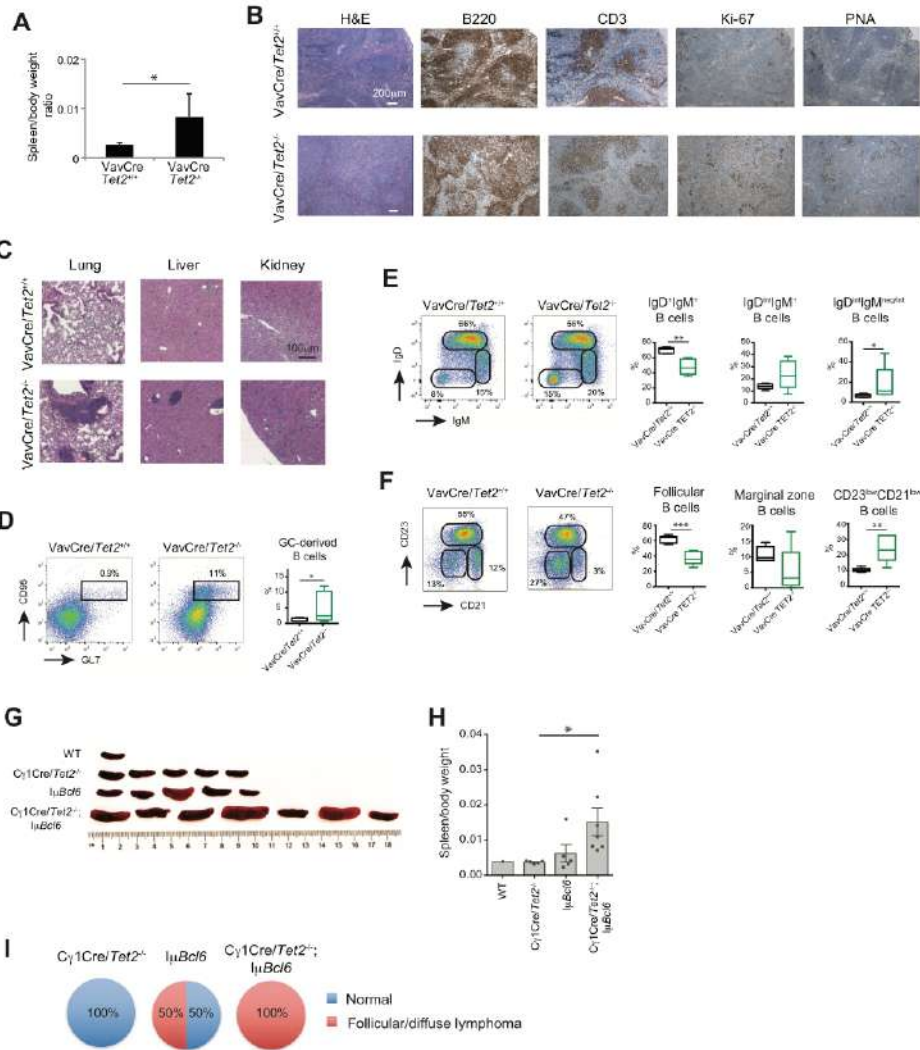
Supplementary Figure 4. A, Heatmap of the FDR scores for the enrichment of 12 gene sets involved in the exit of GC B-cells from the GC. FDR scores were obtained using hypergeometric test for a total of 1294 downregulated genes in *Vav-Cre/Tet2*^{-/-} GC B-cells, described as "DEGs". The entire procedure was subsequently repeated for a set of 1294 randomly selected genes, described as "Random genes". **B**, Representative flow cytometry plot and quantification of centroblasts (CXCR4^{high}CD86^{high}) and centrocytes (CXCR4^{low}CD86^{low}) in the GC of *Cd19-Cre/Tet2*^{+/+} and *Cd19-Cre/Tet2*^{-/-} mice at day 10 after immunization with SRBC. **C**, Representative flow cytometry plot and quantification of centroblasts (CXCR4^{high}CD86^{high}) and centrocytes (CXCR4^{low}CD86^{low}) in the GC of *Cγ1-Cre/Tet2*^{+/+} and *Cγ1-Cre/Tet2*^{-/-} mice at day 10 after immunization with SRBC.



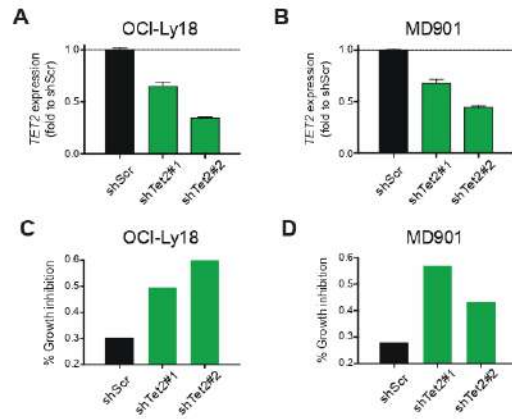
Supplementary Figure 5. Hierarchical clustering (A) and principal component analysis (B) of RNAseq data of *Cγ1Cre/Tet2*^{+/+} (n=3) and *Cγ1Cre/Tet2*^{-/-} (n=4) GC B-cells. C, Normalized gene expression levels of the differentially expressed genes between *Cγ1Cre/Tet2*^{-/-} and *Cγ1Cre/Tet2*^{+/+} GC B-cells. Expression levels were scaled by row. **D, E** GSEA enrichment plots showing correlation of different genesets with ranked expression change between *Cγ1Cre/Tet2*^{+/+} and *Cγ1Cre/Tet2*^{-/-} GC B-cells. NES, normalized enrichment score; FDR, false discovery rate.



Supplementary Figure 6. **A**, Heatmap of the 5hmC enrichment in all DHMR overlapping with intronic enhancers in Vav-Cre/*Tet2*^{+/+} and Vav-Cre/*Tet2*^{-/-} GC B-cells. Region spans 1 kbp up and down from the DHMR peak center. Heatmap was generated using deepTools (67). **B**, Heatmap of the FDR scores for the enrichment of 12 gene sets involved in the exit of GC B-cells from the GC. FDR scores were obtained using hypergeometric test for a total of 1825 genes with enhancers overlapping with DHMRs, described as "DHMR in enh. genes". The entire procedure was subsequently repeated for a set of 1825 randomly selected genes, described as "Random genes". **C**, Correlation of differential hydroxymethylation in intronic enhancers and differential gene expression between Vav-Cre/*Tet2*^{+/+} and Vav-Cre/*Tet2*^{-/-} GC B-cells. **D**, Percentage of promoters and enhancers with DHMR and H3K27Ac peaks that lose >35% signal in the absence of TET2 and CREBBP, respectively. **E**, Read-density tracks of H3K27ac histone modification as well as 5hmC levels in *IIR* locus in Vav-Cre/*Tet2*^{-/-}, Vav-Cre/*Tet2*^{+/+}, VavP-Bcl2 and VavP-Bcl2/Crebbp^{KD} B-cells. Area marked in orange indicates region losing 5hmC and H3K27ac signals. **F**, H3K27Ac enrichment in region shown in A.



Supplementary Figure 7. A, Spleen to body weight ratio of *Vav-Cre/Tet2^{+/+}* (n=5) and *Vav-Cre/Tet2^{-/-}* (n=5) mice **B**, Representative histologic sections of formalin-fixed, paraffin-embedded spleens from *Vav-Cre/Tet2^{+/+}* and *Vav-Cre/Tet2^{-/-}* mice. Sections were stained with H&E and antibodies specific for B220, CD3, Ki-67 and PNA. **C**, Representative histologic sections of formalin-fixed, paraffin-embedded liver, lung and kidney from *Vav-Cre/Tet2^{+/+}* and *Vav-Cre/Tet2^{-/-}* mice. Sections were stained with H&E. **D**, Representative flow cytometry plot and quantification of CD95⁺GL7⁺ GC-derived B-cells. **D, E** Representative flow cytometry plot and quantification of IgD^{hi}IgM⁺, IgD^{int}IgM⁺ and IgD^{low}IgM^{low} B cells (**D**) and CD23⁺CD21⁺ (follicular B cells), CD23⁺CD21⁺ (marginal zone B cells) and CD23^{hi}CD21^{hi} B cells in the spleens of *Vav-Cre/Tet2^{+/+}* and *Vav-Cre/Tet2^{-/-}* mice; two-tailed t test, *p < 0.05 **p < 0.01. **G**, Picture of the spleens of *Cγ1Cre/Tet2^{+/+}; IμBcl6*, *Cγ1Cre/Tet2^{-/-}; IμBcl6* and WT mice. **H**, Spleen to body weight ratio of *Cγ1Cre/Tet2^{-/-}; IμBcl6*, *Cγ1Cre/Tet2^{+/+}; IμBcl6* and WT mice *p < 0.05. **I**, Diagnosis of *Cγ1Cre/Tet2^{-/-}* (n=7), *IμBcl6* (n=6) and *Cγ1Cre/Tet2^{-/-}; IμBcl6* (n=12) mice.



Supplementary Figure 8. A, B, *TET2* knockdown efficiency in OCI-Ly18 (**A**) and MD901 (**B**) cells after 4 days of puromycin selection (1 μ g/ml). Graphs show *TET2* expression in sh*TET2* relative to shScr. **C, D,** Growth inhibition of OCI-Ly18 (**C**) and MD901 (**D**) cells upon *TET2* depletion and treatment with either a selective HDAC3 inhibitor (HDAC3i; 20 μ M) or vehicle for 4 days. Graphs show percentage of growth inhibition calculated as (live cells in vehicle – live cells in HDAC3i)/live cells in vehicle.

2. Résultats supplémentaires

2.1 Contexte et objectifs

L'étude des conséquences de la perte de TET2 dans la réaction du CG était en cours lors de mon arrivée dans l'équipe. Ma contribution sur ce projet a consisté à développer plusieurs stratégies *in vitro* pour documenter et reverser le phénotype observé (illustré dans les Fig 3F-L ; Fig4 H-J ; Fig 6 E-F de l'article n°1). J'ai en particulier mis au point une expérience permettant d'apporter l'expression ectopique de BLIMP1 (que nous avons identifié comme régulé à la baisse dans les BGC et PC TET2KO) par transduction rétrovirale puis analysé ses conséquences sur la réaction du CG. Par ailleurs, il a été démontré que la vitamine C stimule l'activité enzymatique des protéines TET, *in vivo* et *in vitro* (voir paragraphe 3.3.3 de l'introduction). J'ai donc cherché à savoir si la présence de vitamine C dans le milieu de culture de notre modèle *in vitro* pouvait modifier le phénotype observé en l'absence de l'activité TET2 et restaurer la différenciation plasmocytaire. Le premier point a été publié dans l'article n°1, le second est présenté ici en résultat supplémentaire.

Dans un deuxième temps, je me suis intéressée au rôle de TET3 dans la différenciation B terminale. Au cours du développement des cellules B, l'expression de TET2 et de TET3 augmente progressivement, avec un pic dans les cellules B spléniques transitionnelles. Leur expression est maintenue dans les cellules du centre germinatif, avec néanmoins quelques différences : l'ARNm de TET3 n'est pas exprimé de manière différentielle entre les centroblastes et les centrocytes alors que l'ARNm de TET2 atteint son niveau le plus bas dans les centrocytes avant de ré augmenter au stade plasmablaste/plasmocyte. TET1 n'est que très faiblement exprimé.

Nous nous sommes donc demandé si les protéines TET2 et TET3 jouaient un rôle spécifique dans la différenciation B terminale ou bien si la cellule B avait seulement besoin d'une activité catalytique suffisante de protéines TET, qu'elle soit apportée par TET2 ou TET3. Dans le premier cas, le dépôt de 5hmC par TET2 ou TET3 se ferait sur des sites différents du génome, ou éventuellement sur les mêmes sites mais à des stades de différenciation différents. Dans le second cas, il suffirait à la cellule B d'atteindre un certain niveau de 5hmC aux loci nécessaires, indépendamment de la protéine TET responsable de ce dépôt. Les études publiées à ce jour s'intéressent soit à la perte de TET2 soit au double KO TET2/TET3, mais aucune ne rapporte les conséquences d'une perte de TET3 seule sur la réaction du CG. Nous souhaitons comparer les fonctions des deux protéines, pour explorer leur redondance ou leur complémentarité. Cette partie de mon travail a donc consisté à caractériser le phénotype de la perte de TET3 au cours de la réaction immunitaire, *in vivo* et *in vitro*.

2.2. Matériel et Méthodes supplémentaires

2.2.1 Culture cellulaire

- Traitement à la vitamine C

La vitamine C (ou acide ascorbique) a été ajoutée directement au milieu de culture liquide à une concentration finale de 250µM, pendant toute la durée de la culture, selon le protocole décrit précédemment (cf article n°1).

- Traitement à la TAT-Cre.

La TAT-Cre (également connue sous le nom de His-TAT-NLS-Cre) a été achetée chez Ozyme (EG-1001-100, Ozyme). Les cellules B naïves spléniques ont été resuspendues dans du milieu RPMI, préchauffé à 37°C et traitées avec 50 µg/mL de TAT-Cre pendant 45 minutes à 37°C. Les cellules ont ensuite été lavées dans du milieu complet et cultivées selon le protocole décrit précédemment (cf article n°1).

2.2.2 Modèle murins et greffe des progéniteurs hématopoïétiques

Des souris portant l'allèle *Tet3* avec les séquences codantes de l'exon 10 et 11 flanquées de deux sites loxP, ont été générées par la plateforme de recombinaison homologue (Institut Cochin, Paris, France) et ont été croisées avec des souris exprimant le transgène Cre inducible au tamoxifène (Cre-ERT) sous le contrôle du promoteur/enhancer Scl (*Scl-Cre/Tet3^{fl/fl}*). Pour supprimer les allèles *Tet3*-floxés, le tamoxifène a été solubilisé à 20 mg/ml dans de l'huile de tournesol (Sigma-Aldrich, Saint-Quentin Fallavier, France) et 8 mg de tamoxifène ont été administrés aux animaux, deux fois à 5 jours d'intervalle par gavage oral.

Les expériences de greffe ont été réalisées chez des souris femelles C57/BL6 CD45.1 âgées de 8 semaines (laboratoires Charles River), irradiées à 9 Gray. Deux millions de cellules Lin⁻ provenant de souris *Scl-Cre/Tet3^{+/+}* ou *Scl-Cre/Tet3^{-/-}* âgées de 5 mois et induites depuis 3 mois ont été injectées en rétro-orbital.

2.2.3 PCR quantitative en temps réel

Toutes les sondes TaqMan murines ont été achetées chez Applied Biosystems, TET1 : Mm01169089 (exon 11-12) ; TET2 : Mm01312907 (exon 10-11) ; TET3 : Mm01184936 (exon 10-11)

2.3 Résultats

La vitamine C compense la perte de TET2 et favorise la différenciation terminale en plasmocytes in vitro.

Nous avons utilisé le modèle de culture *in vitro* utilisé dans l'article n°1, reposant sur l'utilisation d'une lignée de fibroblastes exprimant le CD40 et sécrétant du BAFF, couplée à l'ajout d'IL4 les quatre premiers jours et d'IL21 les quatre suivants (**Figure 1**).

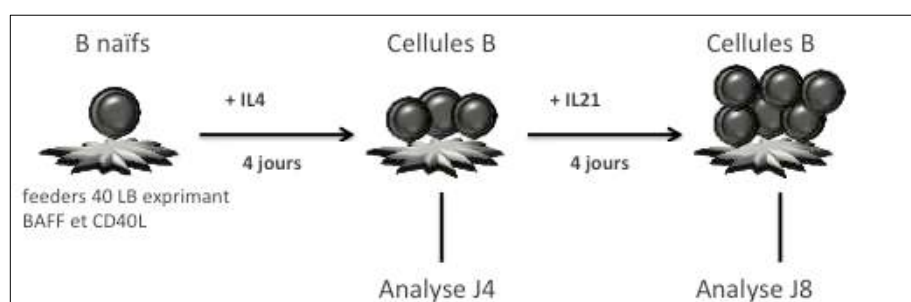


Figure 1 : Représentation schématique du modèle de culture *in vitro*

A J4, après 4 jours de stimulation CD40/IL4, la majorité des cellules analysées sont des BGC, exprimant les marqueurs Fas et GL7. En contexte TET2KO, les BGC représentent 90% et 85% des cellules totales en moyenne, avec ou sans vitamine C respectivement, ce qui est significativement supérieur aux cellules WT. En revanche, les cellules TET2KO traitées par la vitamine C présentent la même proportion de plasmocytes CD138+ que les cellules WT avec ou sans vitamine C (entre 3 et 3,5% des cellules totales en moyenne), supérieure à celles des cellules TET2 KO non traitées (moins de 2% des cellules totales) (**Figure 2A**).

A J8, après 4 jours de stimulation supplémentaire par CD40/IL21, une partie des cellules se différencie en plasmocytes exprimant le CD138+. En contexte TET2KO, 30% des cellules se sont différenciées avec le traitement à la vitamine C, ce qui est supérieur à ce que l'on observe pour les cellules WT non traitées (entre 20 et 25% de CD138+) (**Figure 2B**).

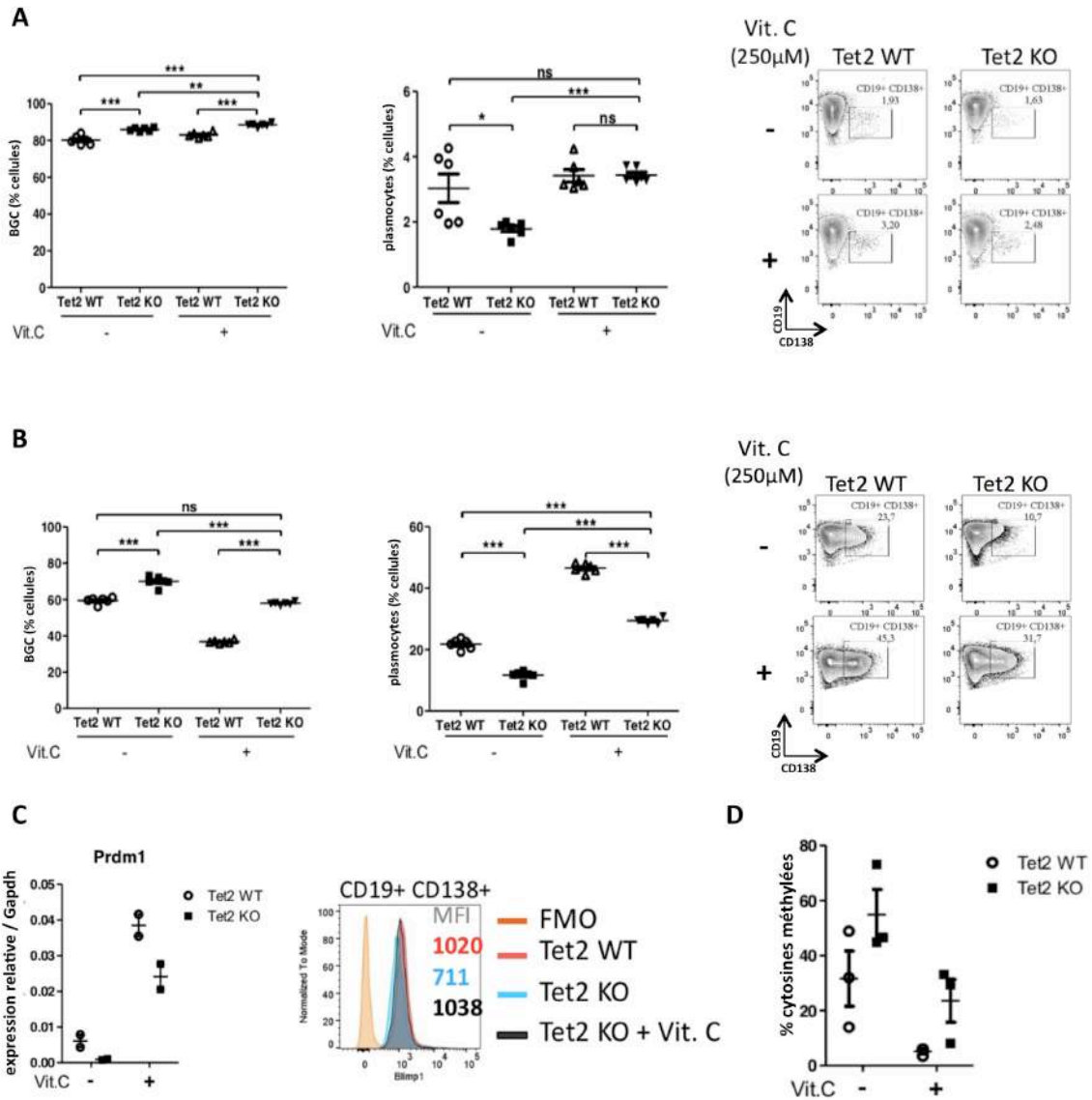


Figure 2. La vitamine C compense la perte de *Tet2* et favorise la différenciation terminale en plasmocytes.

A. Analyse au jour 4 par cytométrie en flux des BGC (CD19+GL7+Fas+) et des plasmocytes (CD19^{low} CD138+), cultivés en présence ou en l'absence de vitamine C. Les graphiques (à gauche) représentent le pourcentage de cellules vivantes (2 expériences indépendantes, en triplicat). Les profils FACS (à droite) sont un exemple de l'analyse de la population CD19+BCD138+.

B. Analyse au jour 8 par cytométrie en flux des BGC (CD19+GL7+Fas+) et des plasmocytes (CD19^{low} CD138+), cultivés en présence ou en l'absence de vitamine C. Les graphiques (à gauche) représentent le pourcentage de cellules vivantes (2 expériences indépendantes, en triplicat). Les profils FACS (à droite) sont un exemple de l'analyse de la population CD19+BCD138+.

C. Expression de *Prdm1* (à gauche) dans les cellules induites in vitro au jour 8 en présence et en l'absence de vitamine C, mesurée par RT-PCR quantitative (n=2). Profil FACS représentatif du marquage intracellulaire BLIMP1 (à droite) des cellules CD19+ CD138+ au jour 8. Les chiffres indiquent l'intensité moyenne de fluorescence (MFI) de BLIMP1.

D. Quantification de la méthylation de l'ADN par PCR quantitative couplée à l'utilisation d'une enzyme de restriction sensible à la méthylation. Les régions différemment méthylées sont représentées par le pourcentage de cytosines méthylées dans le gène *Prdm1*, dans les cellules induites in vitro au jour 8 en présence ou en l'absence de vitamine C

Nous avons lié ces observations à l'expression de *Prdm1* par PCR quantitative en temps réel et par marquage intracellulaire par cytométrie en flux. En effet, lorsque nous ajoutons de la vitamine C, nous constatons une régulation à la hausse de *Prdm1*, au niveau de l'ARNm et au niveau protéique (**Figure 2C**). Ceci est corrélé au % de méthylation du gène *Prdm1* (région dans l'intron 2, cf article n°1), diminué avec le traitement à la vitamine C (**Figure 2D**).

Nos résultats indiquent que l'ajout de vitamine C dans le milieu restaure la différenciation plasmocytaire dans les cellules TET2KO, illustrée par l'acquisition du marqueur CD138+ et par l'augmentation de l'expression de *Prdm1*.

Obtention d'un modèle murin Tet3 KO

Huit semaines après induction de l'expression de la recombinaise Cre par le tamoxifène dans le modèle murin inducible Scl-Cre/*Tet3^{fl/fl}*, le génotypage des cellules triées du sang des souris montrait une efficacité d'excision insuffisante dans les cellules B CD19+ (<50%, données non montrées). Nous avons donc dû mettre en place des stratégies *in vivo* et *in vitro* nous permettant d'obtenir des lymphocytes B naïfs avec une meilleure excision de *Tet3*.

En comptant sur l'avantage de prolifération des progéniteurs *Tet3*-déficients rapporté dans la littérature (243), nous avons réalisé une greffe de cellules Lin- triées à partir de souris Scl-Cre/*Tet3^{-/-}* dans des souris receveuses irradiées de façon sublétales. L'expression de *Tet3* est très faible dans les progéniteurs des souris Scl-Cre/*Tet3^{-/-}* par rapport aux progéniteurs des animaux sauvages (**Figure 3A**). Six semaines après la greffe, le chimérisme sanguin a été évalué. Il indique que plus de 90% des cellules sanguines et plus de 99% des lymphocytes B des souris receveuses ont le phénotype du donneur (**Figure 3B et 3C**). A 8 semaines post-greffe, les animaux ont été immunisés à une ou deux reprises par SRBC. L'absence d'expression de *Tet3* dans les cellules B spléniques des animaux Scl-Cre/*Tet3^{-/-}* après immunisation a été confirmée par PCR quantitative en temps réel (**Figure 3D**).

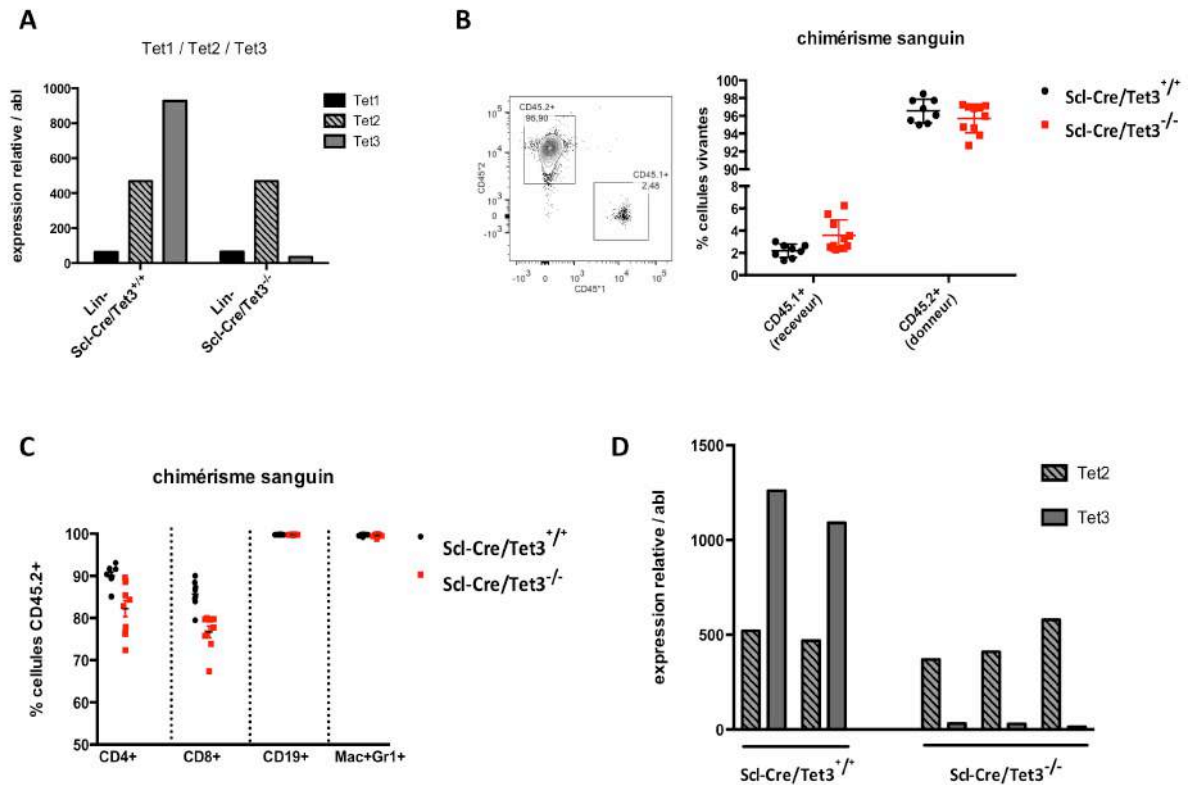


Figure 3. Obtention d'un modèle murin TET3KO

- Expression de *Tet1*, *Tet2* et *Tet3* dans les progéniteurs Lin⁻ utilisés pour la greffe. Les cellules Lin⁻ ont été isolées à partir de la moelle de souris Scl-Cre/*Tet3*^{-/-} et Scl-Cre/*Tet3*^{+/+} et une PCR quantitative en temps réel a été réalisée. Les résultats présentés sont normalisés par rapport à l'expression d'*Abi*.
- Chimérisme global dans le sang des souris receveuses six semaines après la greffe. Le pourcentage de cellules totales vivantes issues du receveur (CD45.1) et du donneur (CD45.2) est représenté sous forme de graphique (à droite) ou de profil FACS (à gauche).
- Chimérisme par sous-type cellulaire dans le sang des souris receveuses six semaines après la greffe, exprimé en pourcentage de cellules du donneur (CD45.2)
- Expression de *Tet2* et *Tet3* dans les lymphocytes B de la rate après immunisation. Les résultats pour 2 animaux Scl-Cre/*Tet3*^{+/+} et 3 animaux Scl-Cre/*Tet3*^{-/-} sont représentés.

Analyse de la perte de *TET3* in vivo

Après une immunisation, nous ne voyons pas de différences significatives entre le groupe contrôle et le groupe Scl-Cre/*Tet3*^{-/-}, à l'exception de la répartition centroblaste/centrocyte au sein des BGC (**Figure 4A** et **4B**). En effet, les animaux Scl-Cre/*Tet3*^{-/-} ont en proportion plus de centrocytes et moins de centroblastes que les animaux sauvages (**Figure 4B**). Ce phénotype est également observé chez les souris TET2KO (cf article n°1). Cependant, après deux immunisations, nous notons une splénomégalie chez les Scl-Cre/*Tet3*^{-/-} et une augmentation du pourcentage de BGC dans la rate (12% en moyenne des cellules totales pour les animaux Scl-Cre/*Tet3*^{-/-} contre 7,5% en moyenne chez les contrôles) (**Figure 4C** et **4D**).

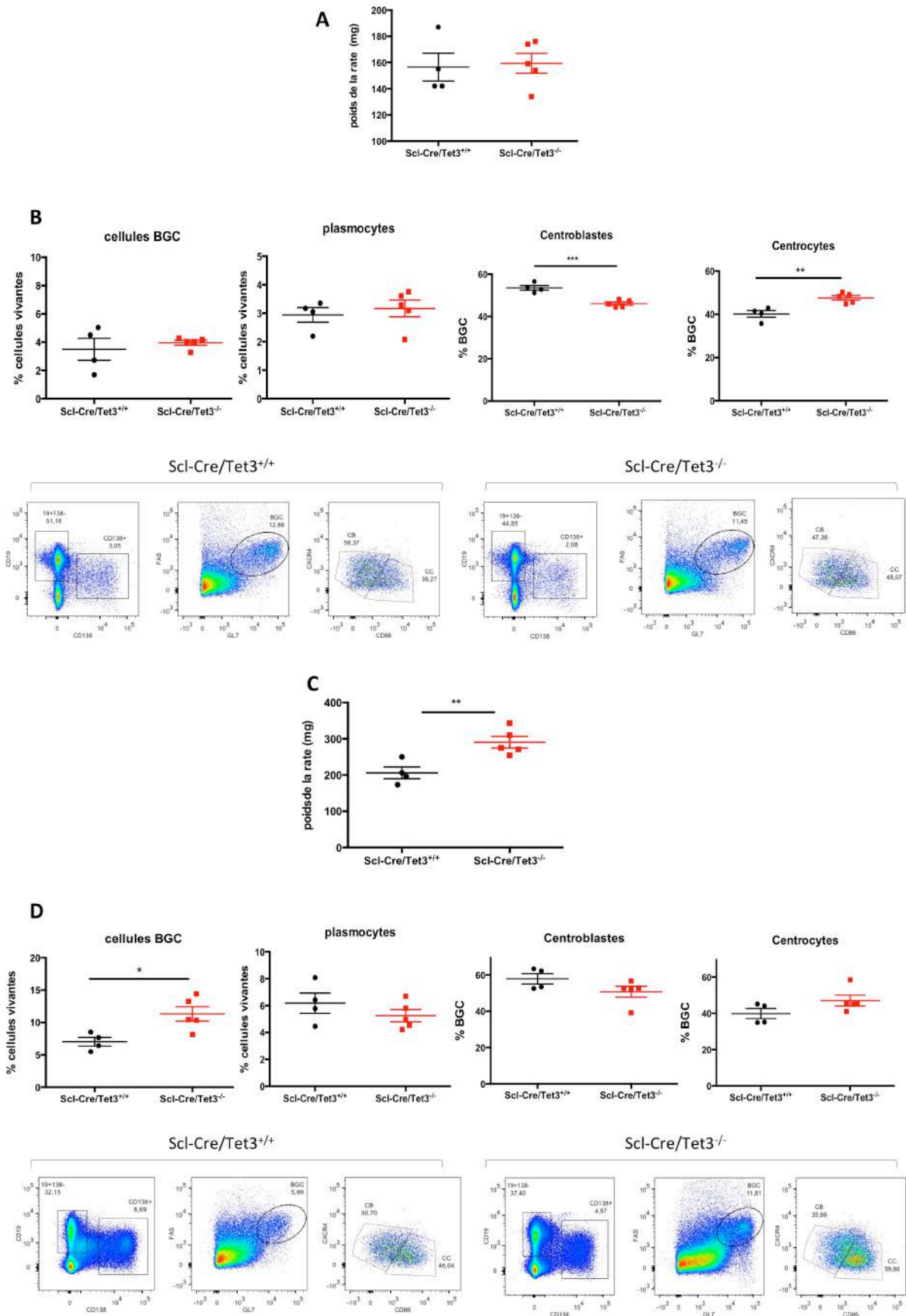


Figure 4. Impact de la perte de Tet3 dans la réaction du centre germinatif *in vivo*

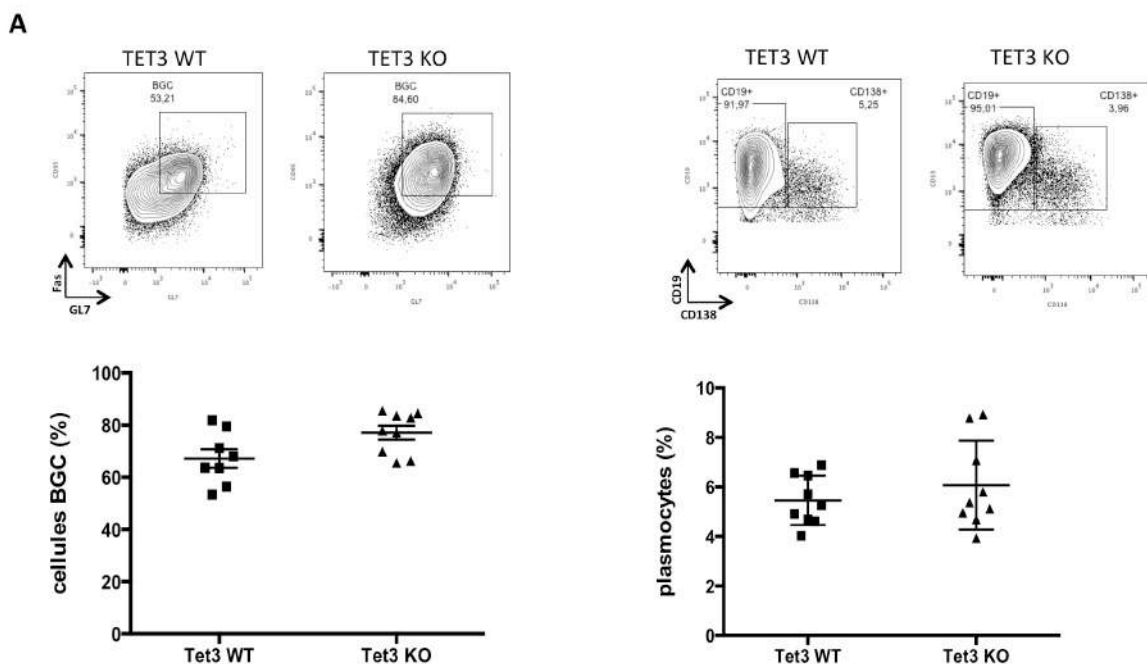
A. Poids de la rate des animaux 10 jours après une immunisation par globules rouges de sang de mouton

- B. Pourcentage des cellules BGC (CD19+GL7+FAS+), plasmocytes (CD19+CD138+), centroblastes (CXCR4^{high}CD86^{high}) et centrocytes (CXCR4^{low}CD86^{low}) dans la rate après une immunisation SRBC; présentées sous forme de profils FACS (en bas) et de quantification (en haut) au jour 10 ; chaque symbole représente une souris (n=4 pour les souris Scl-Cre/*Tet3*^{+/+} et n=5 pour les souris Scl-Cre/*Tet3*^{-/-}).
- C. Poids de la rate des animaux au jour 32 (après une immunisation au jour 0 et une seconde au jour 28) par globules rouges de sang de mouton
- D. Pourcentage des cellules BGC (CD19+GL7+FAS+), plasmocytes (CD19+CD138+), centroblastes (CXCR4^{high}CD86^{high}) et centrocytes (CXCR4^{low}CD86^{low}) dans la rate après deux immunisations SRBC; présentées sous forme de profils FACS (en haut) et de quantification (en bas) au jour 32 ; chaque symbole représente une souris (n=4 pour les souris *Tet3*-WT et n=5 pour les souris *Tet3*-KO).

Toutes les p values p ont été calculées à l'aide du t-test de Student non apparié, *p < 0,05, **p < 0,005, ***p < 0,0005 et ns : non significatif, dans toutes les expériences.

Analyse de la perte de *TET3* *in vitro*

Après traitement à la TAT-Cre, les lymphocytes B naïfs sont mis en culture selon le protocole habituel (cf article n°1). Une première expérience montre une tendance à avoir plus de BGC à J4 (**Figure 5A**) et moins de différenciation plasmocytaire à J8 (**Figure 5B**) en contexte *TET3*KO, sans atteindre la significativité. Les pourcentages d'excision de *Tet3* obtenus après traitement à la TAT-Cre variaient de 70 à 80% (**Figure 5C**). Il reste donc vraisemblablement une partie de l'activité catalytique de *TET3* dans ces cellules, qui expliquerait le phénotype modéré observé et la variabilité des résultats selon les réplicats biologiques. Cette difficulté, qui n'a pas pu être contournée dans les expériences *in vitro* suivantes, à laquelle s'ajoutent la complexité d'obtention du modèle *in vivo* et son absence de phénotype marquant, nous a conduit à privilégier un projet mené en parallèle (voir partie 2 de ce manuscrit).



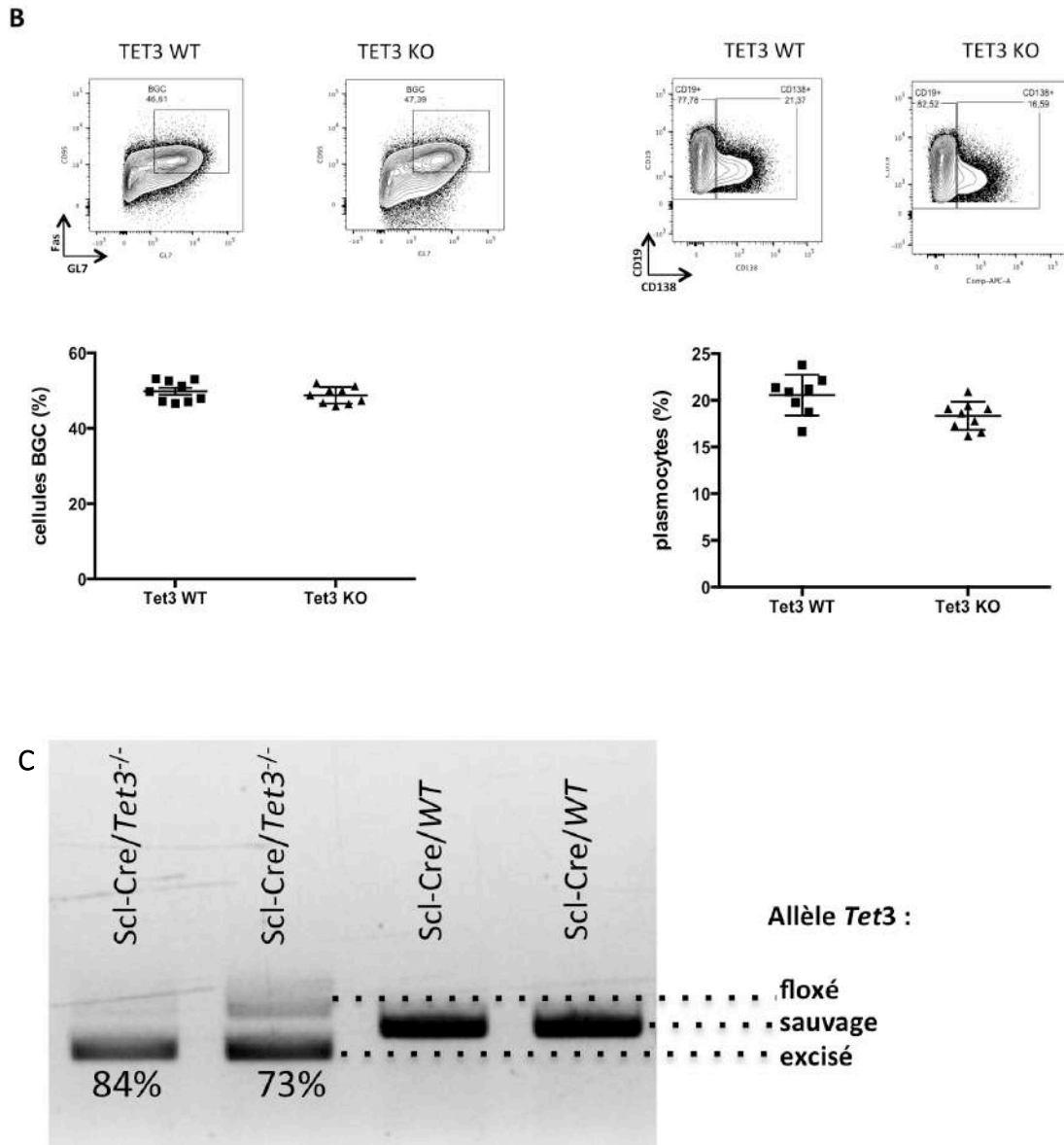


Figure 5. Impact de la perte de *Tet3* dans la réaction du centre germinatif *in vitro*.

- A.** Pourcentage au jour 4 des cellules BGC (CD19+GL7+Fas+), à gauche et plasmocytes (CD19+CD138+), à droite ; présentés sous forme de profils FACS (en haut) et de quantification (en bas); les symboles représentent des triplicats de 3 souris de chaque génotype.
- B.** Pourcentage au jour 8 des cellules BGC (CD19+GL7+Fas+), à gauche et plasmocytes (CD19+CD138+), à droite ; présentés sous forme de profils FACS (en haut) et de quantification (en bas); les symboles représentent des triplicats de 3 souris de chaque génotype.
- C.** Evaluation par PCR de l'efficacité d'excision par la TAT-Cre (50 µg/mL). A J4, l'ADN des cellules en culture a été extrait et une PCR avec des amorces spécifiques *Tet3* a été réalisée. Les résultats de la migration des produits de PCR sur gel d'agarose 2% pour deux souris *Scl-Cre/Tet3^{-/-}* et deux souris contrôles sont présentés.

2.4 Discussion

Rôle de TET2 dans les cellules souches ou dans les cellules engagées dans la différenciation

Les mutations de *TET2* ont d'abord été identifiées dans les hémopathies myéloïdes (SMD et LAM) comme des évènements précoces affectant les cellules souches, augmentant leur nombre et leur capacité d'auto-renouvellement (275,291,468). Quelques études ont depuis montré que ces mutations de *TET2* n'étaient pas toujours les premiers évènements acquis dans la cellule tumorale, comme en témoigne le travail de Ortmann *et al.* (273) dans les syndromes myéloprolifératifs. Ils y rapportent l'existence de 2 groupes de patients, en fonction de l'ordre d'acquisition des mutations de *JAK2* et *TET2*. Les patients « *JAK2*-first » ont plus souvent des PV que des TE, un plus grand risque de thrombose et une meilleure réponse *in vitro* des cellules au ruxolutinib. Les patients « *TET2*-first » présentent une expansion des CSH mais l'excès de différenciation érythroïde et mégacaryocytaire n'est observé qu'après l'acquisition de la mutation *JAK2*. A contrario, les patients « *JAK2*-first » ont un excès de cellules progénitrices érythroïdes et mégacaryocytaires, et l'avantage de prolifération est apporté par *TET2* ultérieurement. Cette étude illustre l'effet coopératif des mutations affectant la différenciation et celles affectant la prolifération. Elle montre également que le rôle de *TET2* n'est pas limité aux CSH, mais que la survenue d'une mutation dans une cellule déjà engagée dans un processus tumoral peut participer à la transformation. En d'autres termes, soit *TET2* fait le lit de la transformation en augmentant la prolifération et l'auto-renouvellement des progéniteurs, soit il vient renforcer l'avantage compétitif d'une cellule tumorale déjà engagée dans la différenciation et ayant acquis d'autres évènements somatiques. Nous nous sommes donc intéressés à cette question dans la lignée lymphoïde B, où des mutations de *TET2* sont également rapportées, notamment dans les DLBCL. Comme pour les SMP, *TET2* pourrait y avoir un rôle de 'pré-amplificateur', sélectionnant des progéniteurs qui seront alors prédisposés à la transformation ou bien avoir un rôle spécifique dans une cellule engagée dans la différenciation lymphoïde.

Dans l'article n°1, nous avons cherché à préciser les liens entre la perte de fonction de *TET2* et la transformation maligne des cellules BGC, en réalisant des expériences fonctionnelles sur des modèles de souris et en reliant les observations aux données des patients humains ayant des DLBCL mutés pour *TET2*. Nous avons ainsi montré que la suppression de *TET2* dans les cellules hématopoïétiques ou les cellules B perturbe la capacité des cellules BGC à engager le CSR et la différenciation terminale.

Concernant le phénotype des animaux à long terme, les souris Vav-Cre *TET2*KO développent des syndromes lymphoprolifératifs avec l'expansion d'une population aberrante de cellules B CD23⁻CD21⁻ tandis que les souris Cg1-Cre *TET2*KO ne montrent

pas d'anomalies particulières. La différence entre ces deux modèles est le stade du développement auquel TET2 est supprimé, dans tous les progéniteurs pour le modèle Vav-Cre et beaucoup plus tardif, dans les B matures du centre germinatif pour le modèle Cg1-Cre.

Étant donné le lien étroit entre prolifération et différenciation, il était envisageable que le phénotype observé (diminution de la différenciation plasmocytaire) s'explique par une modification de la prolifération en contexte TET-déficient. Nos analyses *in vivo* (expression du marqueur de prolifération Ki67 dans la rate des animaux Vav-Cre TET2KO) et *in vitro* (expansion du nombre de cellules à J4 et J8) montrent une prolifération plus importante lors de la perte de TET2. Par ailleurs, une étude de Caron *et al.* (114) rapporte que l'engagement dans la différenciation plasmocytaire *in vitro* est associée à une déméthylation progressive de l'ADN. Ils montrent la participation d'une déméthylation active, associée au dépôt de 5hmC au niveau des gènes de l'identité plasmocytaire.

L'hypothèse que nous pouvons formuler est que ne pouvant pas se différencier, faute d'expression des gènes clés de la différenciation terminale (*PRDM1* par exemple), les cellules TET2KO continuent à proliférer, s'accumulant au stade centrocyte.

Pour la suite de ce travail, étant donné que les cellules BGC sont un ensemble hétérogène de cellules dotées d'importantes capacités de prolifération (centroblastes) ou de différenciation (centrocytes), il serait intéressant d'isoler ces deux sous populations afin d'analyser séparément les conséquences de la perte de TET2 sur leurs profils d'expression génique.

Coopération de TET2 avec d'autres évènements oncogéniques

Plusieurs éléments plaident en faveur de la nécessité de mutations coopérantes pour observer une transformation tumorale dans les contextes *TET2* mutés. De façon générale, les modèles murins TET2KO développent des hémopathies avec une longue latence, ce qui laisse penser que des mutations additionnelles surviennent avant d'aboutir à une transformation avérée.

En pathologie humaine, les mutations de *TET2* peuvent être retrouvées chez des individus sains dans le contexte des CHIP, sans qu'une hémopathie soit systématiquement développée dans les années suivantes (bien que l'incidence de ces CHIP augmentent particulièrement après 60 ans, on peut donc supposer que notre espérance de vie limite la possibilité de pouvoir diagnostiquer d'éventuels cancers subséquents).

Dans l'article n°1, le croisement des souris Cg1-Cre TET2KO (perte de TET2 dans les BGC) avec les souris porteuses de l'oncogène *Bcl6* (expression de BCL6 sous le contrôle du promoteur I μ de la chaîne lourde d'Ig) a accéléré le développement de lymphome. De façon générale, dans les hémopathies, les mutations de *TET2* sont très souvent associées

à d'autres mutations. Chez l'Homme, dans le sous-groupe ST2 de la nouvelle classification des DLBCL proposée par Wright *et al.* (407), les mutations de *TET2* sont associées de manière significative à celles de *SGK1* et *P2RY8*, codant pour des protéines impliquées dans la signalisation. La majorité des mutations de *SGK1* sont tronquantes et pourrait impliquer la signalisation PI3K, SGK1 étant un proche parent des kinases de la famille AKT (469). *P2RY8* est un récepteur couplé à la Gα13 (codé par la guanine nucleotide-binding protein α13 (GNA13), servant de médiateur pour l'inhibition de la migration et régulant la croissance des cellules B dans les tissus lymphoïdes. L'inactivation de *P2RY8* et son médiateur de signalisation GNA13 dans les BGC empêche leur confinement spatial dans les structures folliculaires du CG et l'inhibition de leur activité AKT (470).

Interactions entre Tet2 et CBP

Un autre point important soulevé par notre travail est le chevauchement important entre les ensembles de gènes réprimés de manière aberrante dans les cellules BGC de souris déficientes en *Tet2* ou en *Crebbp*, ainsi que chez les patients atteints de DLBCL avec des mutations de l'une ou l'autre protéine. Nous constatons que la perte de 5hmC est associée à la perte de l'acétylation de H3K27, ce qui suggère que CREBBP nécessite la présence de TET2 et de la marque 5hmC pour assurer l'acétylation de H3K27. Un lien fonctionnel entre TET2 et CREBBP est soutenu par d'autres études, qui montrent que CREBBP et son paralogue p300 peuvent acétyler TET2, augmentant son activité enzymatique et sa stabilité (128), et que le 5hmC co-localise avec les enhancers liés par p300 (262).

Cependant, la perte de CREBBP ou de TET2 dans la réaction du CG n'entraînent pas exactement le même phénotype. Zhang *et al.* (128) rapportent que les animaux *Crebbp^{fl/fl}/CD19-Cre* ont une augmentation du rapport centroblastes/centrocytes, par rapport à leurs congénères de type sauvage, à l'inverse de nos observations chez les TET2KO. De surcroît, Le nombre de BGC *in vivo* après immunisation n'est pas augmenté chez les CREBBPKO. Une explication possible est que bien qu'ils semblent agir sur les mêmes régions régulatrices, les effets de TET2 sont limités à la chromatine alors que CREBBP n'acétyle pas seulement les histones, mais également d'autres protéines comme P53 et BCL6.

Globalement, ces résultats soutiennent l'idée que la régulation précise de l'expression des gènes dans la réponse immunitaire humorale nécessite un dialogue entre les modifications de l'ADN et des histones, qui sont toutes deux gravement perturbées dans les BGC déficientes en TET2.

Protéines TET et Vitamine C

Nos essais avec la vitamine C ont montré un rétablissement de la différenciation plasmocytaire en contexte TET2KO, et une diminution de la méthylation au niveau des régions régulatrices de *Prdm1*. La vitamine C augmente l'activité catalytique de toutes les dioxygénases dépendantes du fer, ses effets ne sont donc pas attribuables de façon spécifique aux protéines TET.

La concentration utilisée dans le milieu de culture (250 μ M) est 3 à 5 fois supérieure aux concentrations normales sanguines, les adultes sains ayant des taux plasmatiques de vitamine C entre 50 et 80 μ M (252). Des essais cliniques ont montré que lorsque la vitamine C était administrée par voie orale, les concentrations plasmatiques étaient étroitement contrôlées (en raison d'une régulation gastro-intestinale de son absorption) : la dose orale maximale testée donnait des niveaux plasmatiques submillimolaire (~200 μ M), tandis que l'administration par voie intraveineuse donnait des concentrations plasmatiques 30 à 70 fois plus élevées, avec une bonne tolérance (252).

Une équipe s'est intéressée à l'effet de la vitamine C sur des lignées de lymphomes B et rapportent une meilleure sensibilité aux drogues (doxorubicine et cisplatine) médiée par la réactivation du gène *SMAD1*, hyperméthylé chez les patients résistant à la chimiothérapie et codant pour un médiateur intracellulaire de la voie du TGF β (471). Néanmoins, aucune étude clinique à ce jour ne démontre les bénéfices de cette molécule dans les hémopathies lymphoïdes B. Il serait intéressant de tester les effets préventifs de la vitamine C sur la survenue de lymphome, par exemple dans le modèle murin de DLBCL *Tet2/Bcl6*. En postulant que la vitamine C pourrait favoriser la réexpression des gènes régulés à la baisse de façon aberrante en l'absence de TET2, et donc rétablir la différenciation terminale de cellules bloquées dans la réaction du CG, nous avons testé le traitement à la vitamine C sur plusieurs lignées de DLBCL humain, mais les résultats préliminaires ne montrent pas d'effets sur la prolifération des cellules ou sur l'expression des gènes clés de la réaction du CG comme *BCL6*, *IRF4* ou *PRDM1*.

Plus de données sont disponibles dans les hémopathies myéloïdes, notamment grâce au travail de Cimmino *et al.* (257) sur les lignées de LAM. Les souris *Tet2*^{-/-} traitées par la vitamine C développent une pathologie LMMC-like avec une latence plus longue et un chiffre de globules blancs moindre. Pour ce qui est du mécanisme d'action, ils ne mettent pas en évidence d'effet de la vitamine C sur l'expression des protéines TET1,2,3 mais une augmentation de leur activité, mesurée en dotblot par la quantité de 5hmC dans les cellules. En cohérence avec ceci, dans leur étude, les souris *Tet2Tet3DKO* présentent une capacité diminuée de générer de la 5hmC en réponse à la vitamine C.

Rôle de TET3 dans la différenciation lymphoïde B

Schoeler *et al.* (295) ont montré que dans les cellules B sauvages murines, les ARNm de *Tet2* et *Tet3* sont régulés à la baisse en fonction de la division cellulaire, mais avec une cinétique différente. Alors que TET2 est initialement régulé à la baisse puis modérément régulé à la hausse à partir des cycles de division 5-6, la régulation à la baisse de TET3 n'est apparente qu'après au moins 4 divisions cellulaires. Ces résultats suggèrent que les cellules BGC régulent à la baisse les protéines TET pour empêcher une différenciation terminale prématurée, mais qu'une régulation à la hausse de TET2 est nécessaire pour que cette différenciation terminale ait lieu.

Puisque TET3 est exprimé dans les cellules B activées, au moins au début de leur différenciation, nous avons cherché à établir le rôle de TET3 dans la réaction du centre germinatif. Nous avons été confrontés à des difficultés pour obtenir un modèle approprié TET3KO, aussi bien *in vivo* qu'*in vitro*. Nous travaillons avec un modèle murin inducible Scl-Cre/Tet3^{fl/fl} (le modèle Tet3KO constitutif étant létal en période périnatale). Ce modèle ne nous a pas permis d'obtenir des lymphocytes B totalement excisés pour *Tet3* 8 semaines après induction, bien que les lymphocytes T soient totalement excisés. Le turnover des lymphocytes B du sang circulant n'est pourtant pas connu pour être moins rapide que celui des lymphocytes T (plusieurs mois chez la souris). L'hypothèse d'une contre-sélection des lymphocytes B TET3KO est difficilement soutenable car les souris induites il y a plus d'un an ont bien des lymphocytes B 100% excisés. Il reste à envisager la possibilité que les lymphocytes T TET3KO soient, eux, sélectionnés positivement.

Le phénotype de la perte de TET3 dans la différenciation B terminale et la réaction du CG semble modéré, aussi bien *in vivo* qu'*in vitro*. Il est à noter que le modèle de culture *in vitro* n'était pas optimal au moment des expériences, et que l'excision par la TAT-Cre n'est pas complète, ce qui peut expliquer ces résultats. Nous pourrions croiser des souris CD19-Cre avec les souris Tet3^{fl/fl} afin d'obtenir un modèle conditionnel spécifique à la lignée B. Concernant le phénotype de la perte de TET3 à long terme, une dizaine d'animaux que nous avons laissé vieillir jusqu'à deux ans ne développent pas de pathologie B (une souris présentait des anomalies de la lignée lymphoïde T). Globalement, il semble donc qu'en l'absence de TET3, l'activité de TET2 soit suffisante pour assurer une différenciation B et une réponse humorale apparemment normale. En accord avec ceci, aucune étude à ce jour ne rapporte les conséquences d'une perte de TET3 seule sur la différenciation B, tandis que de nombreuses études se sont intéressées au double KO *Tet2/Tet3* qui présente les effets les plus spectaculaires, car il abolit de façon très marquée les capacités d'hydroxyméthylation de la cellule.

Interaction entre protéines TET et facteurs de transcription

Une conséquence notable de la perte de fonction de TET2 dans les cellules BGC est la perte focale d'hydroxyméthylation au niveau des enhancers et la répression transcriptionnelle des gènes qui interviennent dans la sortie des GC, comme *PRDM1*. D'autres gènes majeurs de la réaction du CG voient leur expression modifiée en contexte TET2KO, notamment *BCL6* (régulé à la hausse) et *IRF4* (régulé à la baisse).

L'hydroxylation des cytosines méthylées, étape intermédiaire dans la déméthylation, joue un rôle majeur dans le recrutement des FT à la chromatine, en augmentant l'accessibilité de leurs sites de fixation sur l'ADN. Un rapport récent a montré que le facteur de transcription BATF recrute les protéines TET sur des éléments régulateurs spécifiques dans le locus du gène AID, favorisant le dépôt de 5hmC et la déméthylation, assurant l'accessibilité de la chromatine et renforçant l'expression d'AID (244). Conformément à ce rapport et au défaut de CSR observé dans les cellules BGC Cg1-Cre/Tet2^{f/f}Tet3^{f/f}, les niveaux d'ARNm d'AID ont été réduits de 50% dans les cellules BGC (295). Cependant, il a été montré que chez les souris avec le seul knockout *Tet2*, il n'y a pas de réduction de l'expression d'AID (295,297), ce qui suggère que TET3 suffit à maintenir l'expression d'AID en l'absence de TET2. Par conséquent, le maintien de la méthylation des cytosines chez les souris *Tet2*^{-/-} est plus probablement due à une altération de la désamination médiée par AID, ce qui concorde avec les études qui suggèrent une interdépendance entre TET2 et AID dans la déméthylation de l'ADN (472,473).

Dans les cellules Tet2KO, un enrichissement des sites de liaison pour SPI1, E2A et BATF dans les régions hyperméthylées a été rapporté (474), ce qui est remarquable car ces FT ont été précédemment liés à l'activation des gènes par le recrutement de protéines TET dans les cellules B (240,244). Il a été démontré que SPI1 et E2A interagissent physiquement avec les protéines TET2 et TET3, les recrutant dans les enhancers, où elles contribuent, via leur activité d'hydroxyméthylation, à augmenter l'accessibilité de la chromatine (240).

Dans l'article n°1, nous montrons que le déficit en TET2 dans les cellules BGC entraîne une perte de 5hmC, nous pouvons donc supposer qu'il pourrait en parallèle être associé à un gain de 5mC. En effet, notre étude du locus de *Prdm1* montre que les régions affectées par la différence de 5mC et 5hmC ne sont pas les mêmes. Une publication récente à laquelle notre équipe est associée a montré que dans les cellules BGC *Tet2*^{-/-}, les gènes affectés par le gain de 5mC au niveau des promoteurs sont impliqués dans des voies similaires à celles réprimées par la perte de 5hmC au niveau des enhancers (474). La perte de 5hmC et le gain de 5mC dans les cellules BGC déficientes en TET pourrait donc entraîner une diminution relative de l'accessibilité de la chromatine et l'incapacité des autres FT en aval à réguler l'expression des gènes.

Au total, il semblerait que TET2 ait 1) un rôle sur des sites spécifiques (enhancers gènes sortie du GC), pour les démétyler et rendre la chromatine accessible à d'autres FT ; 2) un rôle de sentinelle pour assurer la maintenance de déméthylation de gènes qui doivent rester dans un état déméthylé pour assurer les fonctions cellulaires essentielles (gènes du cycle cellulaire par exemple).

PARTIE 2 : Facteur de transcription SPI1

ARTICLE n°2: A Recurrent Activating Missense Mutation in Waldenström macroglobulinemia affects the DNA binding of the ETS transcription factor SPI1 and enhances proliferation.

1. Résumé de l'article

La macroglobulinémie de Waldenstrom (MW) est une prolifération lymphocytaire B représentant 1 à 2% des hémopathies malignes. Elle fait partie des lymphomes lymphoplasmocytaires et se caractérise par une infiltration médullaire par des lymphoplasmocytes, associée à la sécrétion d'une IgM monoclonale. Sa physiopathologie demeure mal comprise, même si des anomalies génétiques acquises ont déjà été décrites, comme la mutation L265P du gène *MYD88*, présente dans plus de 95% des cas.

Chez 5/85 patients, nous avons identifié une nouvelle mutation somatique dans le domaine de liaison à l'ADN de SPI1/PU.1, un facteur de transcription de la famille ETS. Les protéines ETS sont regroupées en 4 classes (I, IIa/IIb, III, IV). Elles peuvent remplir des fonctions spécifiques qui sont liées en partie à des spécificités d'expression, de liaison à l'ADN et à leur interaction avec des partenaires différents. Toutes les protéines ETS partagent un domaine de liaison à l'ADN hautement conservé qui reconnaît une séquence centrée autour d'un tétranucléotide "GGAA". Le gène *SPI1* est essentiel à l'hématopoïèse et intervient dans le contrôle de différentes étapes de différenciation, *via* son niveau d'expression et son interaction avec d'autres facteurs de transcription ou avec des facteurs épigénétiques. Tous les patients atteints de MW mutés pour *SPI1* présentaient la même substitution nucléotidique hétérozygote (C→G), entraînant le changement d'une Gln (Q) pour un Glu (E) en position 226 (Q226E, QE) dans le domaine ETS de la protéine. Q226 est présent dans les protéines de classe III et responsable de la liaison d'un G en position -2 du motif ETS GGAA, tandis que le E, dans les protéines de classes I et IIa, se lie à un C en position -2.

Nous avons d'abord comparé les propriétés des protéines sauvages et mutées, par des approches d'EMSA (Electromobility Shift Assay) et de transactivation de vecteur indicateur. Ces expériences ont montré que la forme mutée conserve la capacité de liaison à l'ADN et est capable d'activer le promoteur de cibles connues de SPI1 (c-Fes, lambda B1 et GPAA1), et généralement de façon plus importante que la protéine sauvage.

Nous avons ensuite comparé les séquences d'ADN reconnues par les deux formes, ainsi que leurs gènes cibles, en les sur-exprimant dans des lignées cellulaires. Nous avons ainsi montré que la mutation modifie la spécificité de liaison de l'ADN de la protéine en favorisant sa liaison aux sites d'autres classes (I et IIa) des protéines ETS. Par conséquent, SPI1

mutant passe d'un état sauvage, dans lequel les sites de liaison à l'ADN sont situés majoritairement dans les enhancers, à un profil de type classe I/IIa, pour lequel les sites de liaison sont plus fréquemment situés dans les promoteurs. Nos données montrent également un enrichissement transcriptionnel significatif des gènes cibles des protéines ETS de classes I/IIa, à la fois dans une lignée cellulaire et dans des cellules primaires de patients exprimant SPI1QE.

Une différence importante entre SPI1 et les protéines ETS de classes I/IIa est que les activités de liaison et de transcription de l'ADN de ces dernières sont étroitement régulées par des voies de signalisation intracellulaires, ce qui suggère que les promoteurs liés par SPI1QE pourraient échapper à la régulation classique. En effet, nous avons constaté que la surexpression de SPI1QE entraîne une augmentation de la croissance des lignées cellulaires testées, par rapport à SPI1 WT. Dans ces lignées, l'expression de SPI1 mutant était associée à l'enrichissement des signatures transcriptionnelles liées au cycle cellulaire, à la prolifération, aux voies de signalisation intracellulaire et aux programmes transcriptionnels de MYC et IRF4.

Des études ont montré que SPI1 et son proche parent SPIB agissent physiologiquement comme régulateurs négatifs de la différenciation terminale des lymphocytes B. La MW se caractérisant par un degré variable de différenciation des lymphocytes en plasmocytes, nous avons voulu confirmer l'impact de la mutation de SPI1 sur la différenciation des cellules B. Nous avons effectivement constaté, dans un modèle *in vitro*, que l'expression de SPI1 (WT et QE) par des lymphocytes B entraînait une diminution de leur différenciation en plasmocytes.

Les analyses bio-informatiques des profils de transcription montrent un enrichissement transcriptionnel des voies MYC et IRF4 chez les patients mutés pour *SPI1*, en comparaison aux patients non mutés. Nous avons montré *in vitro* que les cellules exprimant SPI1QE sont plus sensibles que les cellules SPI1 WT au Lénalidomide, molécule connue pour cibler de façon indirecte des facteurs de transcription importants pour les lymphocytes B comme IRF4, IKZF1 et IKZF3. Ceci valide les prédictions bio-informatiques et pourrait permettre de traiter spécifiquement les patients mutés pour SPI1.

Nous rapportons dans ce travail le premier exemple de mutation activatrice de SPI1 chez l'Homme. La mutation semble faire adopter à cette protéine ETS de classe III un comportement qui ressemble à celui d'une classe I/IIa. La mutation affecte le programme transcriptionnel de SPI1, et entraîne un enrichissement transcriptionnel des cibles d'autres membres de la famille ETS comme ETS1.

A Recurrent Activating Missense Mutation in Waldenström Macroglobulinemia Affects the DNA Binding of the ETS Transcription Factor SPI1 and Enhances Proliferation



Damien Roos-Weil^{1,2,3,4,5}, Camille Decaudin^{1,2,3,4}, Marine Armand^{1,2,3,4}, Véronique Della-Valle^{1,2,3,4}, M'boyba K. Diop^{1,2,4,6}, Hussein Ghamlouch^{1,2,3,4}, Virginie Ropars⁷, Cécile Hérate^{1,2,3,4}, Diane Lara^{1,2,3,4,8}, Eric Durot^{1,2,3,4}, Rima Haddad⁹, Elena Mylonas^{1,2,3,4,10}, Frederik Damm¹⁰, Françoise Pflumio⁹, Bilyana Stoilova¹¹, Marlen Metzner¹¹, Olivier Elemento¹², Philippe Dessen^{1,2,3,4,5}, Valérie Camara-Clayette^{1,2,6}, François-Loïc Cosset^{1,3}, Els Verhoeyen^{1,3,14}, Véronique Leblond⁵, Vincent Ribrag^{1,15}, Pascale Cornillet-Lefebvre^{1,6}, Philippe Rameau⁶, Nabih Azar⁵, Frédéric Charlotte⁵, Pierre Morel^{1,7}, Jean-Baptiste Charbonnier⁷, Paresh Vyas¹¹, Thomas Mercher^{1,2,3,4}, Said Aoufouchi^{2,3,4,18}, Nathalie Droin^{1,2,3,4,6}, Christel Guillouf^{1,2,3,4}, Florence Nguyen-Khac^{5,8}, and Olivier A. Bernard^{1,2,3,4}

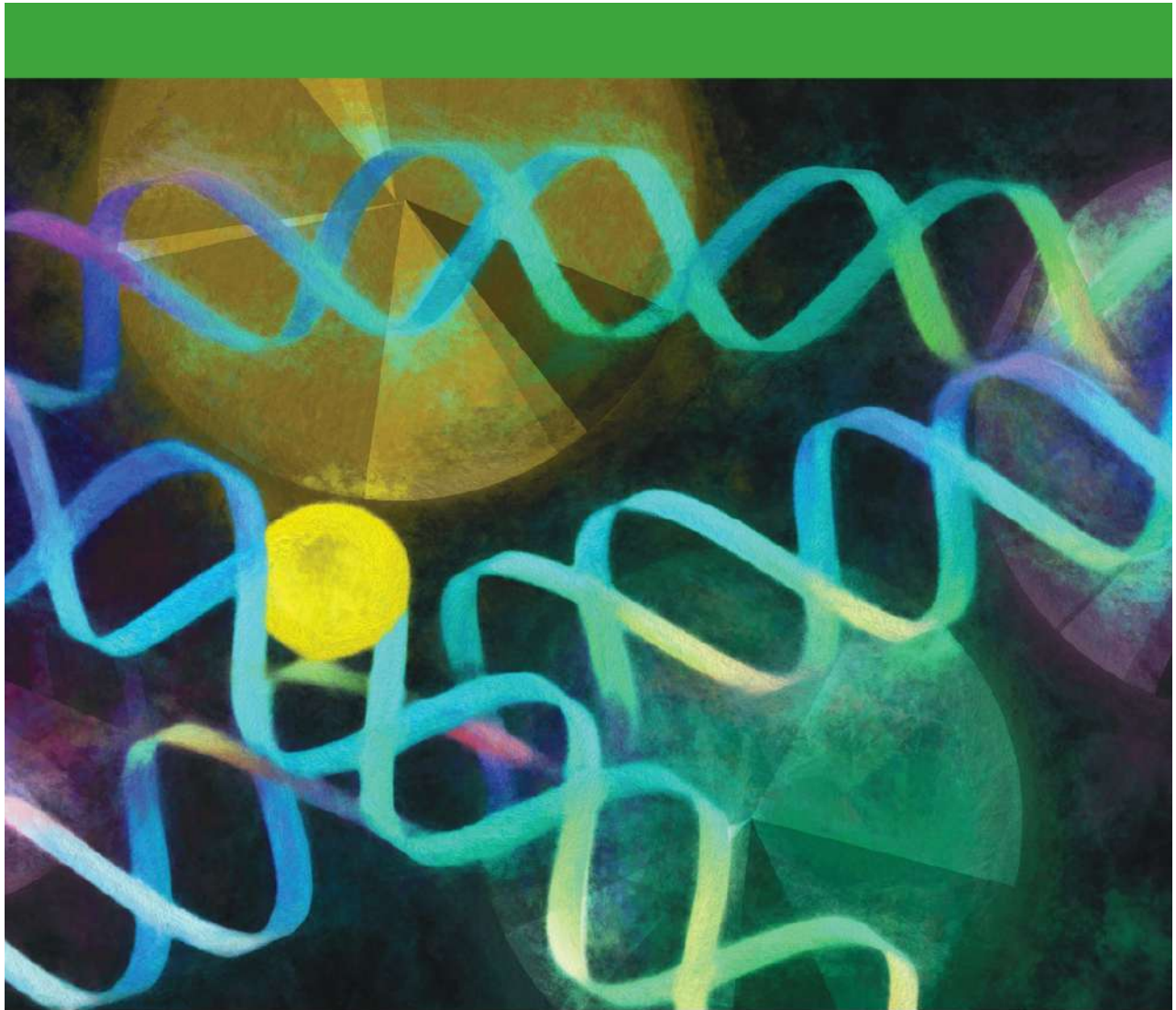
ABSTRACT

The ETS-domain transcription factors divide into subfamilies based on protein similarities, DNA-binding sequences, and interaction with cofactors. They are regulated by extracellular clues and contribute to cellular processes, including proliferation and transformation. ETS genes are targeted through genomic rearrangements in oncogenesis. The *PU.1/SPI1* gene is inactivated by point mutations in human myeloid malignancies. We identified a recurrent somatic mutation (Q226E) in *PU.1/SPI1* in Waldenström macroglobulinemia, a B-cell lymphoproliferative disorder. It affects the DNA-binding affinity of the protein and allows the mutant protein to more frequently bind and activate promoter regions with respect to wild-type protein. Mutant SPI1 binding at promoters activates gene sets typically promoted by other ETS factors, resulting in enhanced proliferation and decreased terminal B-cell differentiation in model cell lines and primary samples. In summary, we describe oncogenic subversion of transcription factor function through subtle alteration of DNA binding leading to cellular proliferation and differentiation arrest.

SIGNIFICANCE: The demonstration that a somatic point mutation tips the balance of genome-binding pattern provides a mechanistic paradigm for how missense mutations in transcription factor genes may be oncogenic in human tumors.

¹INSERM U1170, Gustave Roussy, Villejuif, France. ²Gustave Roussy, Villejuif, France. ³Université Paris-Saclay, Villejuif, France. ⁴Equipe labellisée Ligue Nationale Contre le Cancer, Paris, France. ⁵Sorbonne Université, Hôpital Pitié-Salpêtrière, APHP, Paris, France. ⁶AMMICA, INSERM US23/CNRS UMS3655, Gustave Roussy, Villejuif, France. ⁷Institute for Integrative Biology of the Cell (I2BC), Institute Joliot, CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, Gif-sur-Yvette cedex, France. ⁸Sorbonne Université, INSERM UMR51138, Centre de Recherche des Cordeliers, Paris, France. ⁹Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA) DSV-IRCM-SCSR-LSHL, Université Paris Diderot Sorbonne Paris Cité, Fontenay-aux-Roses, France. ¹⁰Department of Hematology, Oncology and Tumor Immunology, Charité-Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany. ¹¹MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine,

NHR Oxford Biomedical Research Centre Haematology Theme, Radcliffe Department of Medicine and Department of Haematology, Oxford University and Oxford University Hospitals NHS Foundation Trust, United Kingdom. ¹²Department of Physiology and Biophysics, Institute for Computational Biomedicine, Caryl and Israel Englander Institute for Precision Medicine, Weill Cornell Medicine, New York, New York. ¹³CIRI-International Center for Infectiology Research, Team EVIR, Université de Lyon; INSERM, U1111; Ecole Normale Supérieure de Lyon; Université Lyon 1; CNRS, UMR5308, Lyon, France. ¹⁴Université Côte d'Azur, INSERM, C3M, Nice, France. ¹⁵DITEP Gustave Roussy, Villejuif, Paris, France. ¹⁶Laboratoire d'hématologie, Pôle de biologie, CHU de Reims-Hôpital Robert Debré, Avenue du Général Koenig, Reims, France. ¹⁷Centre Hospitalier Dr. Schaffner, Lens; Service d'Hématologie Clinique et Thérapie Cellulaire, CHU Amiens Picardie, Amiens cedex, France. ¹⁸CNRS UMR8200, Gustave Roussy, Villejuif, France.



INTRODUCTION

B-cell differentiation is a multistep process that starts in the bone marrow and continues in periphery. B cells activated by antigen contact undergo affinity maturation in the germinal center (GC) allowing generation of mature plasma cells excreting high-affinity immunoglobulins and memory B cells, keeping the souvenir of antigen encounter (1). B-cell differentiation and maturation are tightly controlled by precise tuning of epigenetic, intracellular signaling pathways [especially B-cell receptor (BCR) and Toll-like receptor (TLR)], and activity of transcription factors, including the ETS family

(1, 2). The ETS family of transcription factors shares a conserved winged helix-turn-helix DNA-binding domain (DBD) and binds to a consensus DNA sequence containing a core 5'-GGA(A/T)-3' motif (3). ETS factors regulate gene expression through binding to ETS motifs or to composite sites through interaction with other transcription factors, including IRF4, IRF8, C/EBP α and C/EBP β , and c-Jun (4). SPI1 is an ETS family member that plays a crucial role in hematopoiesis (5–7), including at different steps of B-cell development, coordinating the proliferation and differentiation of B-lymphoid progenitors, but also acting as a repressor of terminal B-cell differentiation (8, 9). In humans, inactivating *SPI1* mutations

Note Supplementary data for this article are available at Cancer Discovery Online (<http://cancerdiscovery.aacrjournals.org/>).

D. Roos-Weil, C. Decaudin, and M. Armand contributed equally to this article. F. Nguyen-Khac and O.A Bernard are co-senior authors of this article.

Corresponding Authors: Olivier A. Bernard, INSERM U1170, Gustave Roussy, 39 rue Camille Desmoulins, Villejuif Cedex 94805, France. Phone: 331-4211-4233; Fax: 331-4211-5101; E-mail: olivier.bernard@insERM.fr;

and Florence Nguyen-Khac, Unité de Cytogénétique Hématologique, Service d'Hématologie Biologique, GH Pitié-Salpêtrière/Charles Foix, 83 Bd de l'Hôpital, Paris 75013, France. E-mail: florence.nguyen-khac@aphp.fr
Cancer Discov 2019;9:796–811

doi: 10.1158/2159-8290.CD-18-0873

©2019 American Association for Cancer Research.

have been described in acute myeloid leukemia (AML; refs. 10, 11) and gene fusions involving *SPI1* have been identified in pediatric T-cell acute lymphoblastic leukemia (12).

Biological processes driving GC reaction or generation of antibody-secreting plasma cells can be somatically altered leading to B-cell lymphoma and multiple myeloma (1). Activating mutations targeting genes encoding proteins that convey proliferative signals downstream of antigen exposure and BCR signaling include the adapter protein MYD88 (13). Activating mutations of *MYD88* [the most frequent is L265P (LP)] are observed in a wide range of mature B-cell tumors and lead to constitutive activation of the NF- κ B and JAK signaling pathways (13).

Waldenström macroglobulinemia (WM) is defined as an IgM-secreting and bone marrow-infiltrating lymphoplasmacytic lymphoma, related to memory B cells and shows a *MYD88* mutation in more than 95% of cases (14). Despite its indolent nature and treatment advances, WM remains incurable, with most patients dying from disease progression, underscoring the need of more specific targeted therapies. This relies on better understanding of the genetic landscape and mechanisms of transformation. Here, we describe the somatic mutational landscape of WM and identify a novel recurrent activating somatic mutation of the hematopoietic transcription factor *SPI1*. The mutation subtly alters DNA-binding sequence preference, changing the patterns of genes activated from ones normally activated by SPI1 to those more typically activated by other ETS family members. Cumulatively, the altered gene expression programs result in oncogenic proliferative signaling and a block in B-cell differentiation.

RESULTS

Identification of a Recurrent *SPI1* Mutation in WM

To gain insights into WM pathophysiology, we compared tumor and CD3⁺ germline coding sequences of a discovery cohort of 16 patients with WM and identified 265 somatic mutations affecting 230 genes, corresponding to a median of 16.5 mutations per patient (Methods; Supplementary Tables S1–S3; Supplementary Fig. S1A–S1C). In addition, we analyzed an extended cohort of 69 patients with WM, by targeted sequencing of 20 recurrently mutated genes (Supplementary Tables S4 and S5, for coordinates and mutations, and Supplementary Fig. S1D and S1E for summary). The distribution and frequency of mutations in the 85 patients with WM is shown in Fig. 1A. *MYD88* (81/85, 95%), *CXCR4* (24/85, 28%), and *ARID1A* (7/85, 8%) were confirmed to be the most frequently mutated genes in WM. Among the 13 genes mutated in at least 2 patients, 12 (*MYD88*, *JGLL5*, *CXCR4*, *ARID1A*, *CD79B*, *TP53*, *LTB*, *HIST1H1E*, *TRAF2*, *ROBO2*, *TRRAP*, *EZH2*) had been previously identified in the context of WM or other B-lymphoid malignancies (14, 15).

A novel mutation in the transcription factor gene *SPI1* was observed in 5 of 85 patients with WM (5.9%) of both discovery (2/16) and validation (3/69) cohorts (Fig. 1A). All patients carried the same single-nucleotide change (C>G) leading to a Gln (Q) for Glu (E) amino acid change at position 226 (Q226E, QE) in the ETS domain of the protein (Fig. 1B). This mutation was not detected in the matched T-cell

fraction. The Q226 residue lies within the alpha 3 helix of the DBD and is a conserved amino acid specific for the class III subgroup of the ETS proteins, substituted by E in classes I and IIa proteins (Fig. 1C). The mutations were observed on both DNA and RNA sequences from flow-sorted tumor cells (Fig. 1D, left and middle), at clonal or subclonal levels (Fig. 1B). *SPI1* transcription was similar between *SPI1* wild-type (WT) and QE samples (Fig. 1D, right). This mutation was not reported in the Catalogue of Somatic Mutations in Cancer (COSMIC; v87) database and not observed in a large number of other B-cell lymphoproliferative disorders samples (Supplementary Fig. S1F), indicating it might be specific to WM. Its stereotyped and heterozygous nature suggests it represents an example of *SPI1*-activating mutations in humans.

Biological and Clinical Features of *SPI1*-Mutated Patients

The main clinical and biological characteristics of the patients with WM are presented in Supplementary Table S1. Most characteristics were similar between the *SPI1* WT and QE subgroups with the exception of a significantly higher proportion of pejorative WM International Prognostic Scoring System prognostic features (16) in the *SPI1*-mutated subgroup. Analyses of bone marrow biopsy samples for CD38 and CD138 B-cell differentiation markers showed fewer CD138⁺ cells in mutant *SPI1* samples, suggesting less differentiation, with respect to *SPI1* WT patients (Fig. 1E; Supplementary Fig. S2A and S2B).

When compared with *SPI1* WT cases, mutant *SPI1* cases exhibited significantly shorter overall survival (OS), with median survival times that were, respectively, not reached and 20.7 months [HR = 7.39; 95% confidence interval (95% CI), 1.5–89; log-rank P = 0.0004; Fig. 1F]. Although it was not significant, *SPI1*-mutated patients showed a trend to shorter time to relapse and to next treatment compared with *SPI1* WT cases (log-rank P = 0.13 and 0.36). OS of *SPI1*-mutated patients was also shorter than that of *CXCR4*-mutated patients, a previously described aggressive subgroup of patients with WM (Supplementary Fig. S2C–S2K; log-rank P = 0.0003).

Mutant *SPI1* Maintains DNA-Binding and Transactivation Capacities and Increases Cellular Proliferation

SPI1 Q226 lies at a position that specifies the -2 position upstream of the GGAA ETS-core sequence, which represents one of the differences between SPI1 and other ETS-binding sites (17). We first used electrophoretic mobility shift assays (EMSA) to confirm the binding of SPI1 QE protein to the ETS-DNA motif derived from a known SPI1 target (18). A specific shifted complex was observed with the *Fes* probe (Fig. 2A; Supplementary Fig. S3A). Using isothermal calorimetry, we then compared binding affinity of GST fusion proteins containing the WT or mutant DNA-binding domains of SPI1 to lambda B1, a classic SPI1-binding site (19), and mutants thereof. A higher affinity to the WT lambda B1 sequence was observed for the SPI1 WT protein compared with QE protein (the affinity constant of the WT protein was 2.7 times lower). The opposite was observed for the SPI1 consensus sequence mutated into an ETS1-like sequence by

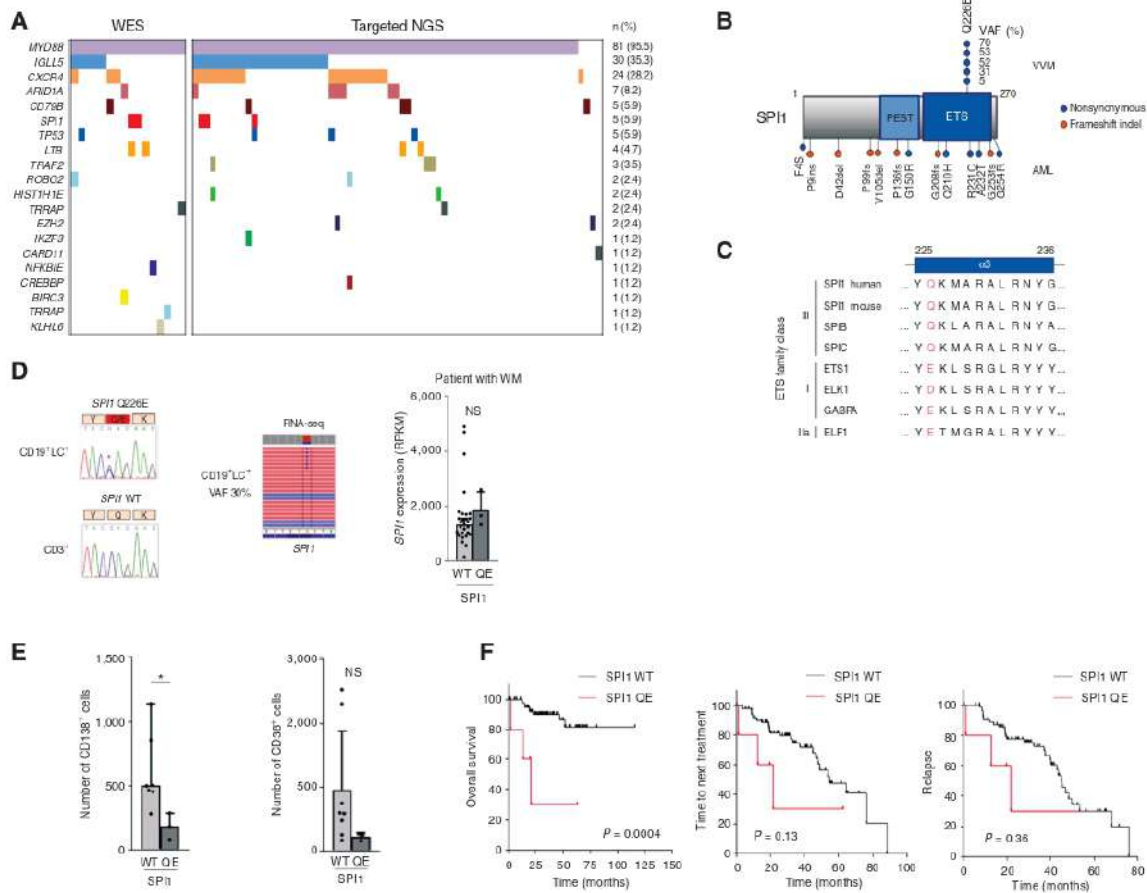


Figure 1. Mutational analyses of WM identify an acquired recurrent *SPI1* missense mutation. **A**, Mutational landscape of 85 patients with WM analyzed by WES ($n = 16$) and targeted ($n = 69$) sequencing. Each column represents a patient sample and each row a mutated gene. The number and percentage of each mutated gene in the whole cohort are indicated on the right of the grid. NGS, next-generation sequencing. **B**, Locations of *SPI1* mutations identified in WM ($n = 85$) and AML ($n = 541$) samples (10, 11). *SPI1* mutations are indicated at their amino acid position, color-coded by type of mutation (orange, frameshift; blue, nonsynonymous). Inactivating mutations observed in AML are shown beneath the protein, whereas WM QE mutations appear above. The estimated VAF in the 5 patients with WM and identified functional domains are indicated. PEST, proline-, glutamic acid-, serine-, and threonine-rich domain; ETS, ETS DNA-binding domain. Q, glutamine; E, glutamic acid; VAF, variant allele frequencies. **C**, Sequence alignment of members of ETS family proteins, in the vicinity of glutamine (Q) amino acid (red) lying at position 226 in alpha helix H3 of the conserved ETS domain. Numbers above indicate positions of amino acids of the human protein. **D**, Left, illustration of the *SPI1* Q226E (QE; c.676C>G, p.Gln226Glu) missense mutation detected in DNA from flow-sorted CD19 and light chain (LC)-positive tumor cells (top, left), as compared with matched CD3⁺ T cells of the same patient (bottom, left). A red star indicates the mutant base. Middle, expression of the mutant copy in RNA-seq data from the same patient is shown. Right, normalized read counts mapping to *SPI1* exons 1–5 are plotted to compare expression levels of *SPI1* transcripts in *SPI1* WT ($n = 29$) and QE ($n = 3$) WM samples. Histograms, median and bars 95% CI. Statistics were performed using Student *t* test. NS, not significant. **E–F**, Biological and clinical features of *SPI1* Q226E mutations in WM. **E**, Expression of differentiation markers analyzed by IHC on bone marrow biopsy samples of *SPI1* WT ($n = 8$) and QE ($n = 3$) patients with WM. The numbers of CD38⁺ and CD138⁺ plasma cells were evaluated on randomly placed squares representing a median whole area of 0.9 mm². Error bars, SD. Statistics were performed using Student *t* test. NS, not significant. **F**, Kaplan–Meier survival curves of overall survival (left), time to next treatment (middle) and relapse-free survival (right) for patients with ($n = 5$) and without ($n = 80$) *SPI1* QE mutation. *P* values are indicated and based on the log-rank test.

a double CC to AA mutation (at position -2 and -1 upstream of the GGAA core sequence). None of the proteins showed significant binding toward a mutant ETS core site (Fig. 2B; Supplementary Fig. S3B–S3F).

To evaluate the transcriptional potential of the mutant protein, we performed luciferase assays with reporter constructs containing the *Fes* promoter region (Fig. 2C), the lambda B1 site (Fig. 2D), or the ETS1-binding sequence of

the human *GPA1* gene promoter (Fig. 2E), a gene bound by *SPI1* QE in OCI-Ly10 cells (Supplementary Fig. S4A). Dose-dependent transcriptional activity was observed for both WT and QE proteins on each of the studied constructs that was abolished when the ETS-core sequence was mutated (Fig. 2C–E; Supplementary Fig. S3G–S3H). We noticed differences in the transcriptional activation capacities of the two proteins depending on the tested target. Up to two-fold more activation

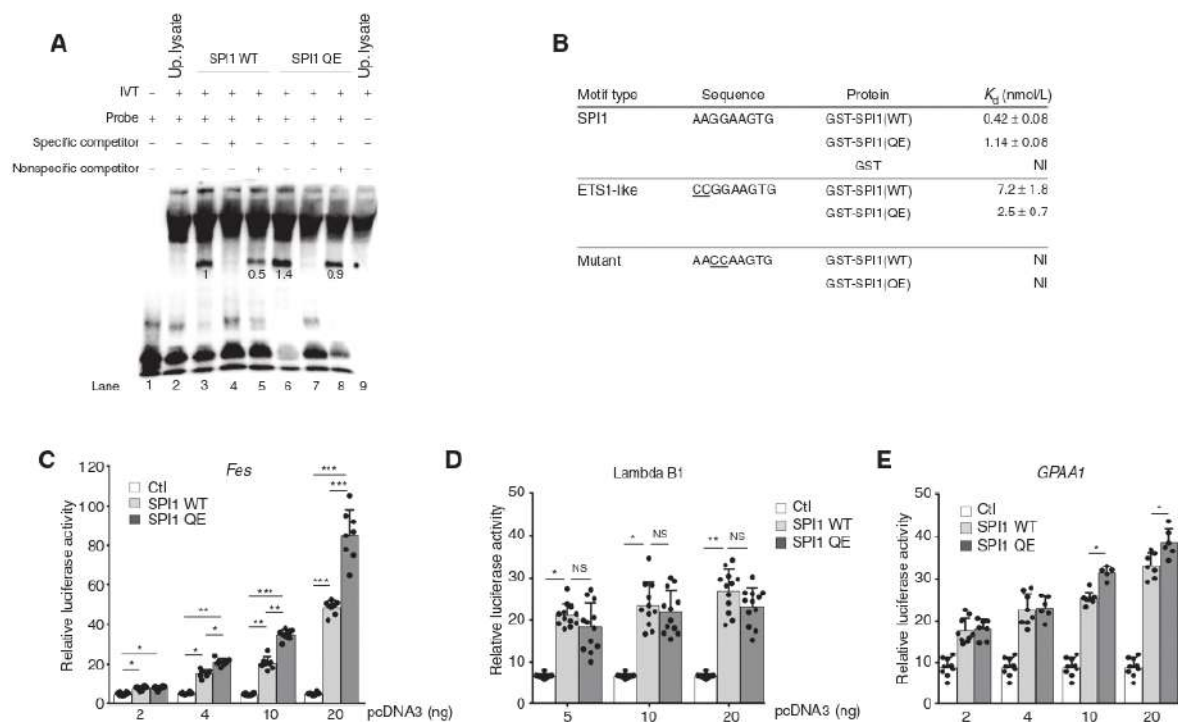


Figure 2. DNA-binding and transcriptional activation properties of mutant SPI1. **A**, Mutant SPI1 binds to ETS motif sequences. EMSA of control (up. lysate, unprogrammed lysate), SPI1 WT and QE IVT proteins in the presence of an ETS motif-containing probe with (+) or without (-) a 160x excess of a specific or nonspecific competitor, as indicated. The ETS motif-containing probe and competitors are from ref. 18. Sequences are provided in Methods. Estimation of the DNA-binding complex intensity in the different conditions is represented. Protein quantities produced by the indicated IVT lysates were similar (Supplementary Fig. S3A). IVT, *in vitro* translation; up., unprogrammed. **B**, *In vitro* comparison of DNA-binding affinities shows differences between SPI1 WT and QE proteins. Affinity was measured by isothermal calorimetry of the DNA-binding domain of SPI1 in fusion with GST with different DNA sequences. GST alone shows no high affinity site. The control DNA, with a GGAA > CCAA mutation in the core ETS-binding motif, demonstrated no high-affinity site. NI, no interaction. **C-E**, Mutant SPI1 transactivates a variety of ETS-reporter constructs. HEK293T cells were transfected with increasing amounts of the indicated [control (Ctl), SPI1 WT, or QE] expression plasmids and fixed amounts of a luciferase reporter gene. Reporter genes contain three copies of the following ETS motifs: *Fes* (**C**), *lambda B1* (**D**; see **B** for motif), and *GPAA1* promoter (CCGGAAGTG; **E**). Sequences are provided in Methods. Three independent experiments were performed ($n = 8-9$; mean ± SD). Error bars, SD. *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$.

was observed for SPI1 QE on the *Fes* promoter, whereas the two proteins showed similar activity on the SPI1-specific lambda B1 site. A stronger activation was observed with SPI1 QE at the highest doses on the ETS site of the *GPAA1* promoter. These *in vitro* experiments show that while maintaining the binding and transactivation capacity of the protein, the mutation differentially affects SPI1 activity depending on the tested region. More precisely, it appears that SPI1 QE retains binding and transactivation capacity on the classic SPI1-binding site but has an increased affinity and transactivation capacity on ETS1-like binding sites.

To gain insight into SPI1 QE binding to DNA on a genome-wide scale, we performed a chromatin immunoprecipitation sequencing (ChIP-seq) experiment with OCI-Ly10 cells expressing GFP fused to SPI1 WT or QE in an inducible fashion (Supplementary Fig. S4B and S4C). ChIP-seq performed with anti-GFP or anti-SPI1 antibodies showed similar results [Fig. 3A and B; Supplementary Tables S6-S9 (for anti-GFP); Supplementary Fig. S4D and S4E and Supplementary Tables S10-S13 (for anti-SPI1)].

DNA motif analyses of SPI1 WT- and QE-specific peaks identified different preferential motifs. The classic SPI1 motif was at the top of the motif list bound by SPI1 WT (Fig. 3A), whereas it ranked in 10th position for SPI1 QE (Fig. 3B). The top hit of motifs recognized by SPI1 QE is a common class I/IIa (ETS1/GABPA/ELF1/ELK1) motif (Fig. 3B). These two motifs differ in a single nucleotide (G or C) at position -2 of the core GGA(A/T), as clearly shown by the *de novo* analysis (Fig. 3C). Classes I/IIa motifs were observed in 77 and 95% ($\chi^2 P = 2.2 \times 10^{-9}$) and class III motifs in 85 and 57% ($\chi^2 P = 8.0 \times 10^{-163}$) of SPI1 WT and QE peaks, respectively (Supplementary Tables S6-S9). The composite SPI1-IRF site dropped from 9th in SPI1 WT to 19th rank in SPI1 QE. These data are in line with DNA-binding affinity, as evaluated by isothermal calorimetry (Fig. 2B) and the *in vitro* analyses of ETS family DNA-binding properties, which pointed at Q226 of SPI1 (class III) and E226 of ELF1 (class IIa) for recognizing G or C at position -2 (17). In addition, gene set enrichment analyses (GSEA) using gene sets specific for ETS1, ELK1, and GABPA indicate that SPI1 QE transcriptional activation in

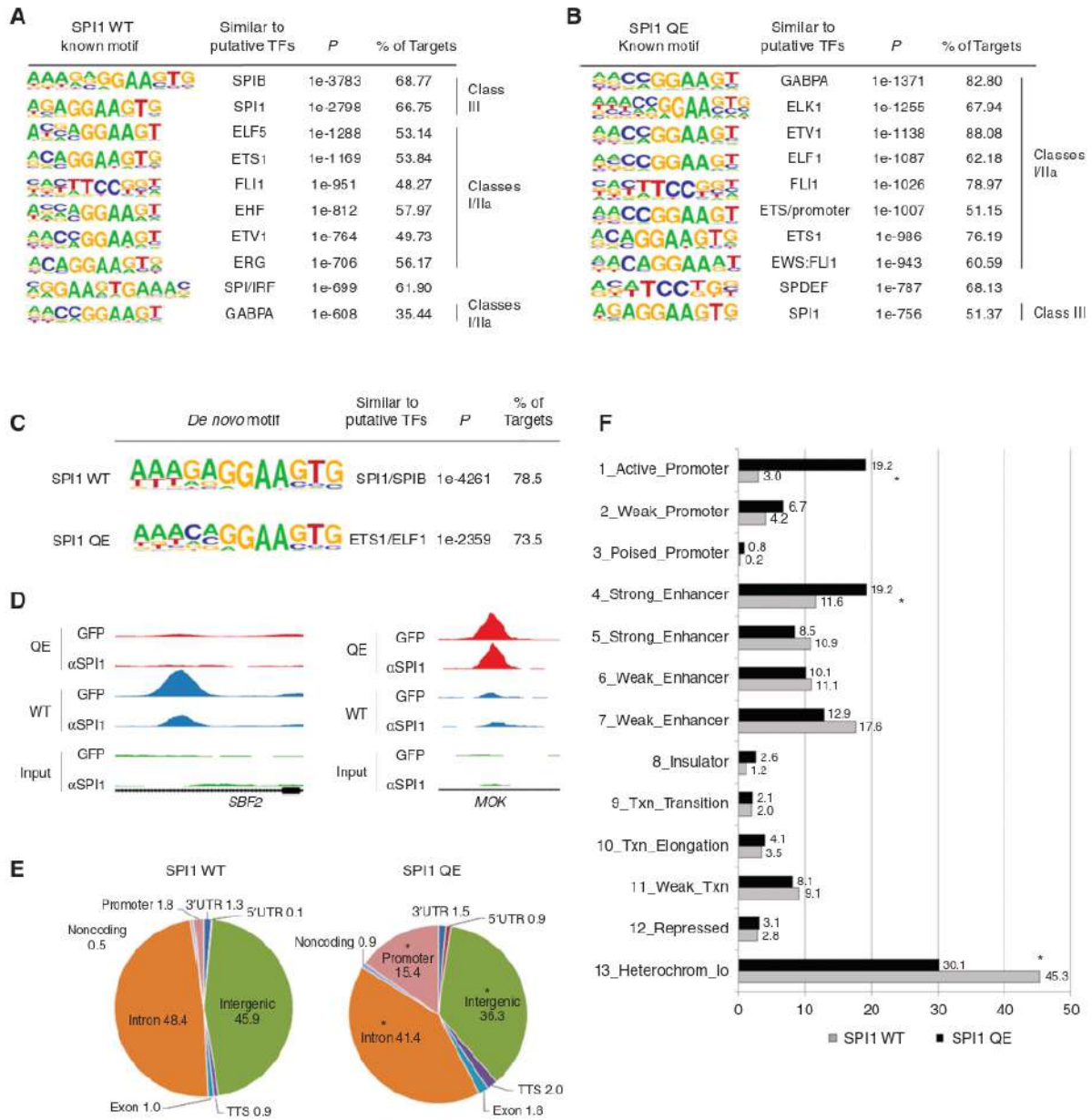


Figure 3. DNA-binding profile of mutant SPI1 in human B-cell lymphoma cell line. **A–F.** ChIP-seq analyses, performed with anti-GFP-Trap antibodies, in OCI-Ly10 cells transfected with doxycycline-inducible vectors expressing SPI1 WT or QE. Commonest known (**A** and **B**) and *de novo* (**C**) DNA motifs recognized by SPI1 WT (**A**) and QE (**B**) proteins. Homer motif known and *de novo* analyses were performed with a cut-off *P* value of 10^{-30} . The complete lists of *de novo* and known motifs are available in Supplementary Tables S6–S9 and S10–S13 for ChIP experiments performed with anti-GFP and anti-SPI1 antibodies, respectively. Motifs, associated transcription factors, *P* values, and percentages of peaks containing the motifs are indicated. **D.** ChIP-seq normalized signals with SPI1 QE (red) and WT (blue) proteins along with the input (green) are shown for two examples of differentially bound genes. The results of ChIP-seq data performed with GFP-Trap (top, GFP) and anti-SPI1 (bottom, α SPI1) antibodies are shown. *SBF2*: FDR = 0.00162; *MOK*: FDR = 0.00562. Genomic localization, defined by RefSeq (**E**) or ChromHMM (**F**, ref. 20) annotations, of DNA sequences (peaks) bound by SPI1 WT (left or gray) and QE (right or black; GFP) proteins. Different genomic regions and their respective percentage (with respect to the total number of peaks) are indicated. Statistics comparing the repartition of genomic regions were performed using χ^2 test. *, *P* \leq 0.05. TF, transcription factors; TTS, transcription termination sites; UTR, untranslated regions.

OCI-Ly10 cells more closely resembles ELK1 and ETS1-associated gene profiles (see Fig. 6B).

Examples of SPI1 QE- and WT-specific peaks are shown in Fig. 3D. Genomic regions bound by SPI1 QE were significantly enriched in promoters (defined by RefSeq; Fig. 3E), and in active promoters and strong enhancers (defined by ChromHMM annotations; Fig. 3F; ref. 20). Our data are consistent with frequent binding of classes I/IIa proteins at promoters, in contrast with what is observed for SPI1, in an Epstein-Barr virus-transformed B-cell line (Supplementary Fig. S4F and S4G and Supplementary Table S14).

Collectively, these data show that the *SPI1* QE mutation modifies the genome-wide DNA-binding pattern of the protein in a subtle way, in terms of both motif and localization, resulting in a specific transcriptional response.

We next investigated the effect of mutant SPI1 on cellular parameters. We first tested OCI-Ly10 cells that were stably transduced with doxycycline-inducible constructs encoding GFP [control (Ctl)] or GFP-SPI1 fusion proteins (WT or QE) and incubated or not with doxycycline for 48 hours. Both proteins were expressed at similar levels (Supplementary Fig. S4B). Both SPI1 WT and QE-expressing cells expanded in numbers significantly more than cells harboring the control construct (SPI1 WT cells expanded cell numbers twice as much and there were 20% more cells containing the SPI1 QE construct than the WT; Fig. 4A). To confirm this observation in another cellular context, we similarly overexpressed WT and mutant SPI1 proteins in BCWM.1 and MWCL-1 cell lines, established from *MYD88* LP WM samples. In both cell lines, overexpression of SPI1 QE resulted in more proliferation than overexpression of the WT form (see Fig. 6F and H). However, the OCI-Ly10 cell line represents a cleaner model than BCWM.1 and MWCL-1 cells for the functional investigations of SPI1 proteins, given the lack of detectable endogenous SPI1 expression in OCI-Ly10, unlike BCWM.1 and MWCL-1 (Supplementary Fig. S4B and SSA-SSD). Cell-cycle analyses in OCI-Ly10 cells showed that ectopic SPI1 expression decreased the percentage of cells in G₁ phase and proportionately increased cells in G₂-M phase (Fig. 4B-E). Similar variations of higher amplitude were observed in cells expressing SPI1 QE. These data indicate that mutant SPI1 likely promotes expansion via cellular proliferation in a manner that is similar to, but greater than SPI1 WT.

To characterize the consequences of mutant SPI1 binding, we used RNA sequencing (RNA-seq) to compare OCI-Ly10 cells expressing SPI1 WT or QE proteins. By GSEA, we observed that, as compared with SPI1 WT, expression profiles of mutant SPI1 samples were significantly enriched for functional annotations related to proliferation, cell cycle but also signal transduction programs that are central to B-cell physiology and pathology, including MYC, BCR, CD40 receptor (CD40), TLR, and PI3K signaling pathways (Fig. 4F; see Supplementary Table S15 for the list of enriched signatures). Differential expression was observed for target genes of key B-cell transcription factors, including XBP1, PRDM1, and IRF4 when signatures from the Lymphochip database (21) were used (Fig. 4F). These data confirm that SPI1 QE affects cellular proliferation to a higher extent than SPI1 WT.

SPI1 Mutation Cooperates with MYD88L265P and Interferes with Terminal B-cell Maturation in Murine Primary B Cells

We next investigated the effects of mutant SPI1 on hematopoietic and B-cell development in primary cells. To determine the effect of mutant SPI1 on hematopoietic development, mouse progenitor cells were transduced to express either SPI1 WT or QE and transplanted into lethally irradiated mice. SPI1 expression promoted accumulation of GFP⁺ cells in bone marrow and myeloid differentiation at the expense of B-lymphoid differentiation (Supplementary Fig. S6A-S6C). Mature B cells were significantly fewer and differentiation of B-lymphoid progenitors was abnormal with more cells accumulating at an earlier stage in SPI1 QE condition compared with SPI1 WT (Hardy's fraction A vs. B and C; Supplementary Fig. S6D). These data are consistent with previous reports describing SPI1 as a determinant of lymphoid and myeloid fate choices; high SPI1 levels favor the development of myeloid cells whereas B-cell development requires its downregulation (22) and indicates that SPI1 QE behaves as SPI1 WT in those settings. No transformation was observed during an 18-month follow-up.

Because *SPI1* QE mutations are detected in association with *MYD88* LP, we constructed retroviral vectors driving expression of *MYD88* and SPI1 (WT or mutant) either alone or in combination (Fig. 5A) to investigate potential oncogenic cooperation. All constructs expressed comparable levels of proteins (Supplementary Fig. S7A and S7B). Naive B-lymphocytes were purified from WT mice, stimulated, and transduced by retroviral constructs (Fig. 5B), and the proportion of fluorescent cells was monitored in the culture, as described previously (23). Ectopic expression of *MYD88* WT, SPI1 WT, or the empty backbone vector did not confer growth capacities to the transduced cells, whereas expression of mutant SPI1 expanded cell survival (Fig. 5C). As already reported (23), expression of mutant *MYD88* conferred a growth advantage over B cells expressing WT constructs. This advantage was statistically increased when SPI1 QE, but not WT, was coexpressed with *MYD88* LP, indicating cooperation between those two mutants. We also investigated the consequences on lymphocyte maturation toward plasmablastic/plasmacytic cells by examining the proportion of GFP⁺ B cells (CD19) expressing CD138, a marker of terminal B-cell differentiation. Expression of *MYD88* did not markedly affect the proportion of CD138⁺ B cells, whereas expression of SPI1, WT or QE, alone or in combination with *MYD88* LP, decreased the proportion of this cell population (Fig. 5D and E). The drop in the ratio of CD138⁺ B cells was more important when SPI1 was expressed alone. In contrast to cellular growth, SPI1 WT or QE did not differentially affect the differentiation ratio in these conditions.

To extend our investigation, we compared RNA-seq data from *MYD88* LP, *MYD88* LP-SPI1 WT (LP/WT), and *MYD88* LP-SPI1 QE (LP/QE)-expressing cells. In keeping with increased proliferation, MYC, E2F, and JAK-STAT signatures were enriched in SPI1 QE-expressing cells, with respect to SPI1 WT (Supplementary Fig. S7D-S7F).

To confirm our observations that SPI1 QE interferes with B-cell differentiation and their relevance in primary

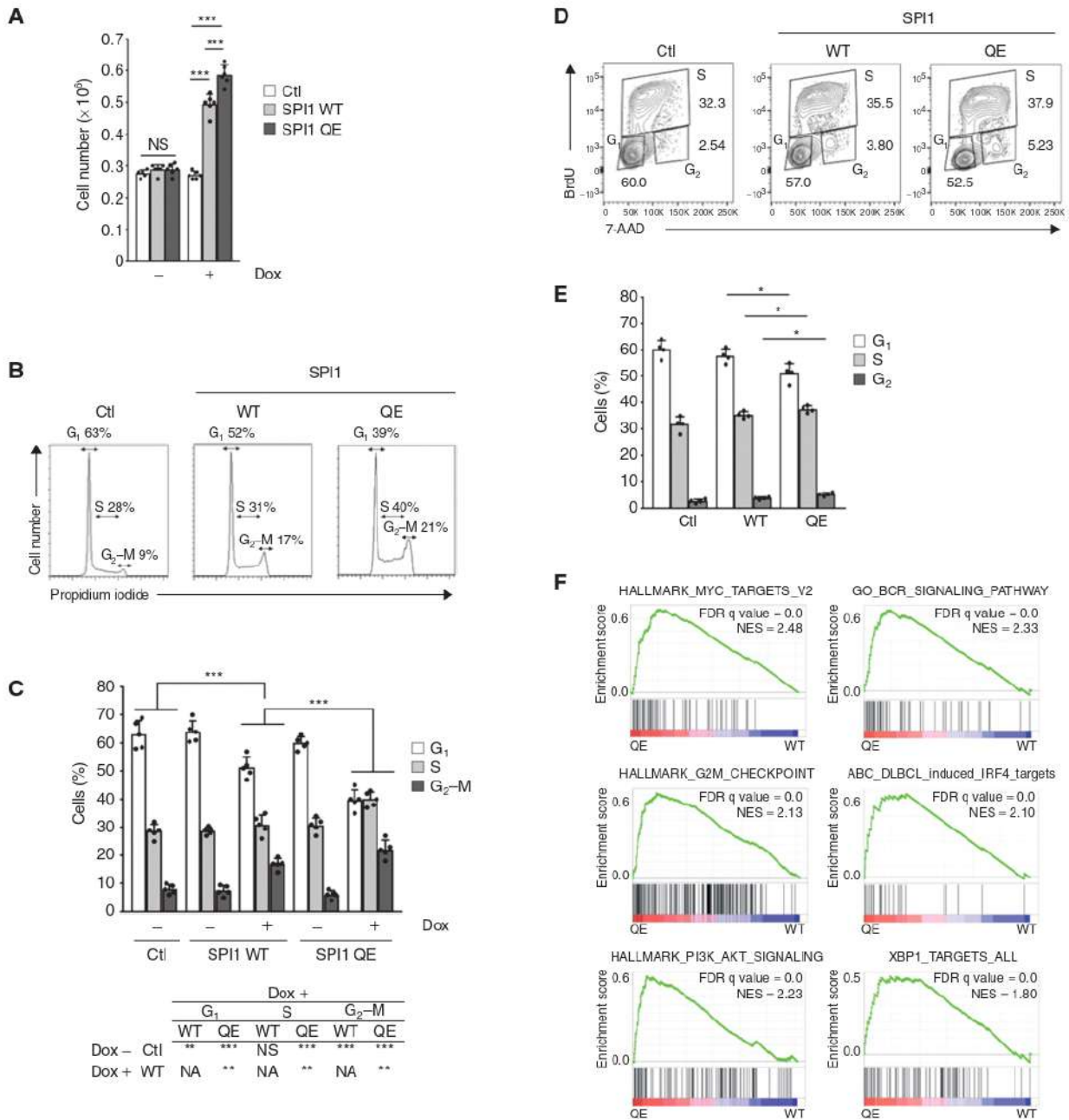


Figure 4. Cellular effects of mutant SPI1. **A**, Induction of SPI1 stimulates cellular proliferation. Quantitation of *in vitro* proliferation for OCI-Ly10 cells transduced with doxycycline-inducible constructs encoding GFP (Ctl), or SPI1-GFP fusion proteins (either WT or QE; $n = 6$, mean \pm SD), with (+) or without (-) the 48-hour addition of doxycycline. The levels of SPI1 proteins and their nuclear localization were checked in OCI-Ly10 cells bearing the different constructs (Supplementary Fig. S4B and S4C), and found to be equivalent for SPI1 WT and QE conditions. NS, not significant. ***, $P \leq 0.001$. Statistics were performed using two-tailed Student *t* test. Error bars, SD. **B-E**, Cell-cycle analyses in OCI-Ly10 cells transduced with doxycycline-inducible constructs encoding GFP (Ctl), or SPI1-GFP fusion proteins (either WT or QE). Representative flow cytometry analyses (**B**) of propidium iodide DNA staining of OCI-Ly10 cells bearing the indicated constructs and estimation (in %; **C**) of the fraction of cells in G₁-S-G₂-M cell-cycle phases. Cell cultures were treated with ($n = 5$) or without ($n = 5$) doxycycline during 24 hours before staining. The table beneath the graph details statistical comparisons between different conditions. Representative flow cytometry analyses (**D**) of BrdU incorporation of OCI-Ly10 cells bearing the indicated constructs and estimation (in %; **E**) of the fraction of cells in G₁-S-G₂ cell-cycle phases ($n = 4$). BrdU incorporation was monitored after 24 hours of doxycycline stimulation. G₁, S, G₂ cell-cycle phases are indicated in representative flow cytometry analyses. The results are shown as means \pm SD. Statistics were performed using two-tailed Student *t* test. * $P \leq 0.05$; ** $P \leq 0.01$; *** $P \leq 0.001$. Dox, doxycycline; NS, not significant; NA, not applicable. **F**, GSEA enrichment plots from RNA-seq analyses obtained upon SPI1 WT ($n = 3$) and QE ($n = 3$) expression in OCI-Ly10 cells showing enrichment in QE condition for selected published dataset regarding proliferation, cell cycle, B-cell signaling, and differentiation. The complete lists of significantly enriched gene sets are available in Supplementary Table S15. NES, normalized enrichment score.

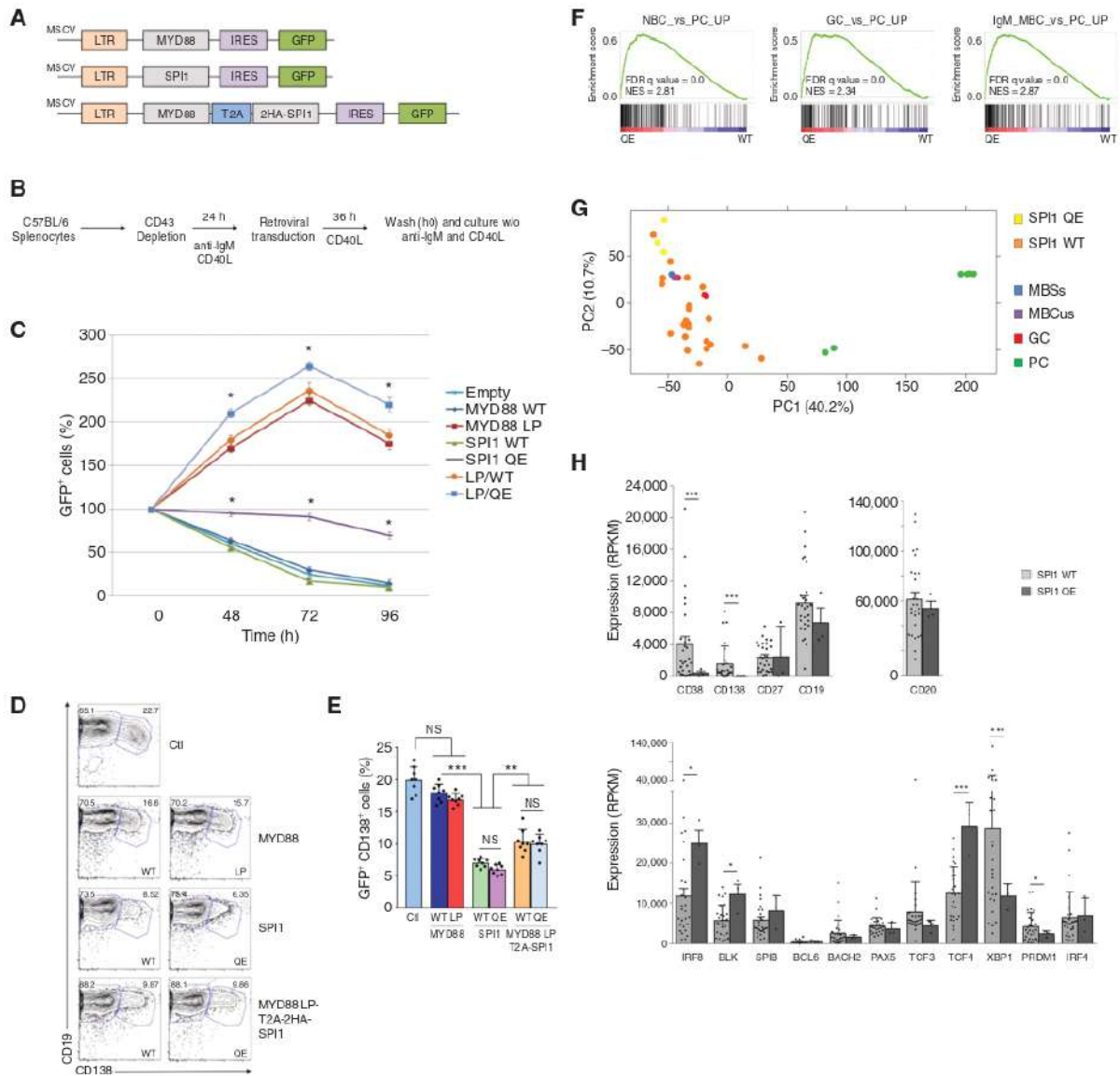


Figure 5. Effects of mutant SPI1 on maturation of murine B cells. **A**, Graphical illustration of MSCV constructs used for murine B-cell experiments. RNA and protein expressions of MYD88 and SPI1 from the different constructs were verified and are illustrated in Supplementary Fig. S7A and S7B. LTR, long terminal repeats sequence; IRES, internal ribosome entry site. **B**, Scheme of murine B-cell experiment. Procedure details are provided in the Methods section. CD40L, CD40 ligand; w/o, without. **C**, Time course of the abundance of GFP⁺ murine B cells transduced with empty (control) or the color-coded indicated MSCV plasmids [turquoise, empty; blue, MYD88 WT; red, MYD88LP; green, SPI1 WT; violet, SPI1 QE; orange, MYD88LP/SPI1 WT (LP/WT); light blue, MYD88LP/SPI1 QE (LP/QE)] and cultured for 96 hours without any mitogen agent. Results are represented relative to those of GFP⁺ cells at time 0, set as 100%. Data are representative of three independent experiments ($n = 8$, mean \pm SD). Error bars, SD. Statistics were performed using Student *t* test. The LP/QE condition was statistically different from the MYD88LP and LP/WT conditions, whereas SPI1 QE was different from MYD88 WT, SPI1 WT, and empty conditions. *, $P \leq 0.05$; LP, L265P. **D**, Representative flow cytometry analyses of CD19⁺CD138⁺ cells after 72 hours of culture. CD19 and CD138 expression analysis is gated on viable (Sytox⁻) transduced (GFP⁺) cells. **E**, Percentages of CD19⁺CD138⁺ GFP⁺ cells as means \pm SD ($n = 8$) are shown in histograms for each construct. **, $P \leq 0.01$; ***, $P \leq 0.001$. NS, not significant. **F**, GSEA enrichment plots for immunologic signatures using MSigDB, comparing SPI1 WT ($n = 29$) versus QE ($n = 3$) WM RNA-seq samples. The complete list is available in Supplementary Table S16. MBC, memory B cells; NBC, naïve B cells; PC, plasma cells; NES, normalized enrichment score. **G**, Principal component analysis of normal B cells and WM samples. This analysis included plasma cells (PC; green; $n = 6$), switched (MBCs; blue; $n = 1$) and unswitched (MBCus; violet; $n = 1$) memory, germinal center (GC; red; $n = 3$) B cells, and SPI1 WT (orange; $n = 29$) and QE (yellow; $n = 3$) WM samples. **H**, Expression of key genes controlling normal B-cell differentiation in SPI1 WT (light gray; $n = 29$) and QE (dark gray; $n = 3$) WM RNA-seq samples. Error bars, SD. Statistics were performed using Student *t* test. *, $P \leq 0.05$; ***, $P \leq 0.001$.

WM samples, we performed GSEA of RNA-seq data specifically for immunologic signatures comparing *SPI1* WT ($n = 29$) and QE ($n = 3$) WM samples with available RNA (Supplementary Table S16). The top hit was a gene set comparing naive B cells to plasma cells (Fig. 5F). The prevalence of this signature was further exemplified by 7 of the top 20 most enriched gene sets (out of more than 4,000 total gene sets) that represent a comparison of plasma cells with either naive B cells, GC or memory B cells. We then compared transcription profiles between these WM samples and normal B cells representing various steps of differentiation. Using principal component analyses (PCA), normal B-cell samples appeared to spread according to differentiation along the first principal component (PC1), with immature B cells on the left and mature, differentiated plasma cells on the right-hand side (Fig. 5G). Most of the WM samples were located between those two groups in keeping with the blocked plasmablast differentiation step of these tumor cells. Four WM samples and the three *SPI1* QE samples located left of memory B cells, suggesting an earlier arrest of differentiation. Specific examination of cell surface markers and key genes controlling germinal center reaction or plasma cell development confirmed that the *SPI1* QE samples were less differentiated than the other WM samples (Fig. 5H). We observed similar results when we compared *SPI1* QE samples along with the four WM samples, which segregated with them in PCA, with the rest of the cohort (Supplementary Fig. S8A). These observations are in keeping with a lower expression of CD138 in *SPI1* QE bone marrow biopsy WM samples (Fig. 1E; Supplementary Fig. S2A and S2B).

Proliferation of *SPI1* QE-Expressing Cells Is Sensitive to JQ1 and Lenalidomide

To support our findings, we extended GSEA using other MSigDB and Lymphochip signatures and compared model cell lines and primary samples data. We confirmed in both settings significant enrichment in classes I/IIa ETS motifs and signaling pathways related to proliferation and BCR and other intracellular signaling (MYC, JAK/STAT, PI3K/AKT, TLR, CD40; Fig. 6A). Most of these enrichments were also confirmed when we performed GSEA considering only genes bound by *SPI1* in OCI-Ly10 active promoters' regions (Supplementary Fig. S9A–S9C). Comparison of gene-expression signatures related to ETS1, ELK1, GABPA, and *SPI1* showed enrichment in *SPI1* QE-expressing samples and cell lines, with respect to WT samples. Lists of genes specifically bound by class I/IIa ETS factors were obtained from CHIP-seq data from GM12878 (Supplementary Table S17) and used in GSEA. Analyses confirmed greater ELK1 and ETS1 than GABPA signature enrichment in mutant samples (Fig. 6B). Note that a *SPI1* signature is also enriched in *SPI1* QE patients (Fig. 6B).

Signatures of MYC targets were among the most enriched signatures in *SPI1* QE-expressing OCI-Ly10 cells and *SPI1* QE WM samples (Fig. 6A). We thus interrogated whether *SPI1* QE-expressing OCI-Ly10 cells would be sensitive to JQ1, a BET inhibitor which targets MYC, among other activities. As described previously (24), JQ1 treatment induced significant cell death in OCI-Ly10 cells from 48 hours of exposition (data not shown). Although increasing concen-

trations of JQ1 did not affect viability of control cells at 24 hours, the proportion of viable *SPI1* WT- and QE-expressing OCI-Ly10 cells decreased with a significantly more pronounced effect in *SPI1* QE condition (Fig. 6C). We confirmed that JQ1 treatment lowered MYC expression in this context (data not shown).

Because of its promising clinical effects and its targeting of IRF4 (25) among other activities, we evaluated the impact of lenalidomide on cell growth. Lenalidomide treatment of cell lines revealed a higher sensitivity of *SPI1* QE, with respect to *SPI1* WT-expressing cells (Fig. 6D). Expression of *IRF4*-targeting shRNAs (sh*IRF4-1* and sh*IRF4-2*) in OCI-Ly10 cells resulted in 25% to 28% and 28% to 50% knockdown in, respectively, *SPI1* WT- and QE-expressing cells at the *IRF4* RNA level, and in an approximate 70% drop in IRF4 protein for both shRNAs in both *SPI1* WT- and QE-expressing cells (Supplementary Fig. S10A and S10B). IRF4 knockdown also blunted the proliferative stimulation induced by *SPI1* QE (Fig. 6E), confirming a role of IRF4 in *SPI1* QE consequences.

We also investigated JQ1 and lenalidomide treatments in BCWM.1 and MWCL-1 cells, which resulted in qualitatively similar differences between *SPI1* QE- and WT-overexpressing cells (Fig. 6F–I).

DISCUSSION

We defined the somatic mutation landscape of WM, a rare mature B-cell lymphoproliferative disorder. In addition to confirming recurrent mutations in *MYD88*, *CXCR4*, *ARID1A*, and *CD79B*, we identified a novel recurrent somatic mutation (c.676C>G, p.Q226E) in the DNA-binding domain of *SPI1*, a transcription factor of the class III subgroup of the ETS family. Q226 is present in class III proteins and responsible for the binding of a G at position -2 of the ETS-core GGA(A/T) motif, whereas the E, in classes I and IIa proteins, binds to C at position -2 (17). The mutation modifies the DNA-binding specificity of the protein in promoting binding to DNA recognition sequences of other classes (I and IIa) of ETS proteins. As a result, the mutant *SPI1* switches from a WT pattern of chromatin binding (sites preferentially located in enhancers and only 5%–15% in promoters) to a class I/IIa type profile, with binding sites more frequently located in promoters (40%–60%; refs. 26, 27). Our data also show a significant transcriptional enrichment of classes I/IIa ETS target genes in a model cell line and in patient samples expressing *SPI1* QE compared with WT.

An important feature of classes I/IIa ETS factors is that DNA binding and transcription activities are tightly regulated by intracellular signaling pathways, such as calcium and MAPK activity (28, 29), suggesting that promoters bound by the *SPI1* QE may escape normal regulation. Indeed, overexpression of *SPI1* QE was associated with growth stimulation with respect to the *SPI1* WT in all tested cell lines. Mutant *SPI1* expression was associated with enrichment in transcription signatures related to cell cycle, proliferation, intracellular signaling pathways, and also MYC- and IRF4-related transcriptional programs. As *SPI1* interaction with transcription factors (30) and epigenetic regulators, such as CBP (31) and TET2 (32), is important for their regulation,

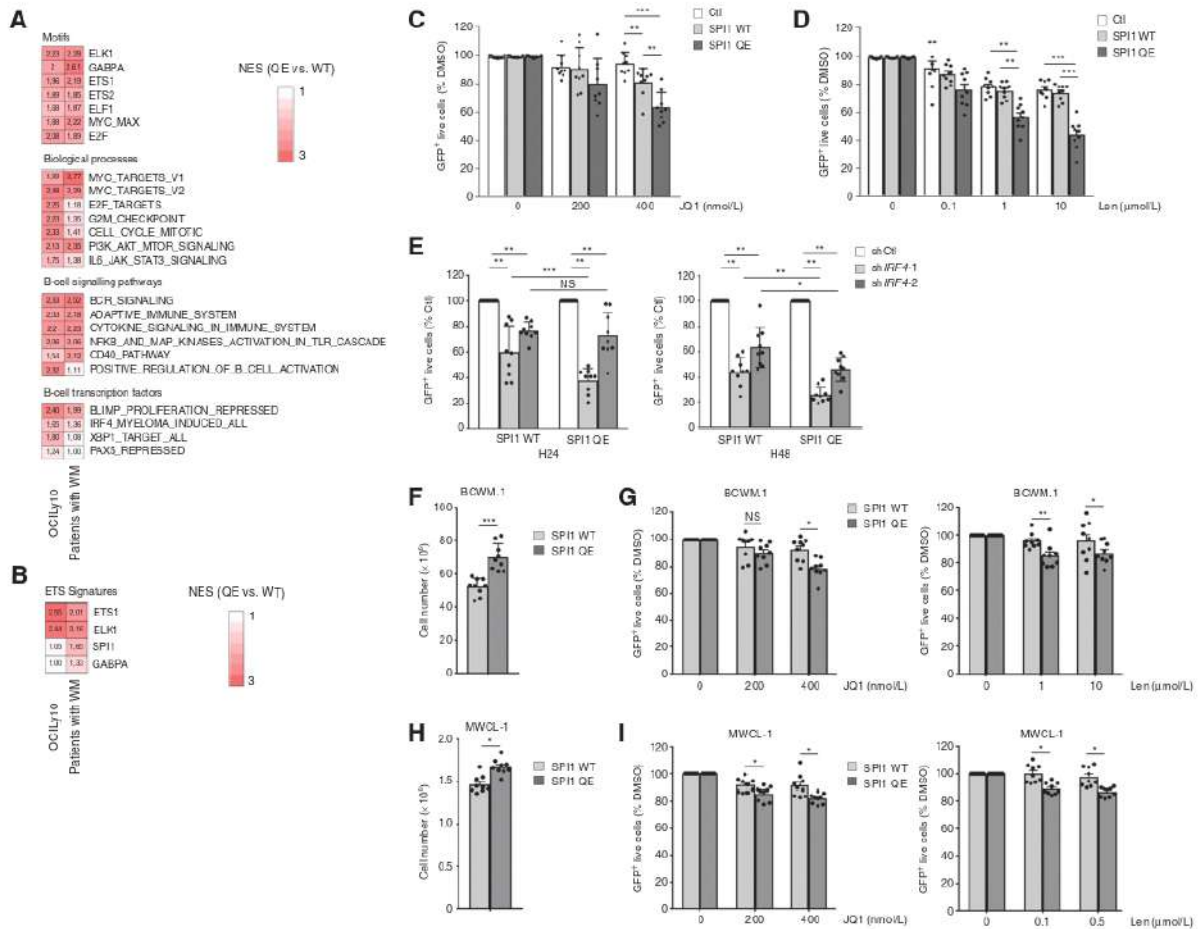


Figure 6. Genes and pathways activated by mutant SPI1 may be targeted. **A**, Integration of GSEA in OCI-Ly10 cells and WM cases, comparing SPI1 QE to WT RNA-seq samples, using MSigDB and the Lymphochip databases. Heat map represents the normalized enrichment score (NES) of each gene set signature tested. Only pathways relevant to transcription factor motifs, intracellular signaling pathways, and B-cell biology are shown in the figure. The complete lists are available in Supplementary Tables S15 and S16. NES values are indicated inside each box. **B**, GSEA of ETS1, ELK1, GABPA, and SPI1-specific signatures in SPI1 QE and WT OCI-Ly10 (left) and WM (right) samples. These specific signatures were generated from ENCODE ChIP-seq data of GM12878. Heat map represents the NES of each gene set signature tested. Represented signatures are enriched in SPI1 QE samples. NES values are indicated inside each box. **C**, The BET inhibitor JQ1 decreases cellular proliferation of SPI1 QE-transduced OCI-Ly10 cells. Viability of OCI-Ly10 cells bearing empty (Ct; white), SPI1 WT (light gray), or QE (dark gray) constructs ($n=9$, mean \pm SD). OCI-Ly10 cells were treated for 24 hours with doxycycline and the indicated concentrations of BET inhibitor (JQ1). Cell viabilities are represented relative to those without JQ1 treatment. Three independent experiments have been performed. NS, not significant. $^*P \leq 0.01$, $^{***}P \leq 0.001$. Statistics were performed using two-tailed Student *t* test. Error bars, SD. **D**, SPI1 QE-expressing OCI-Ly10 cells are more sensitive to lenalidomide (Len) than those expressing SPI1 WT. Viability of OCI-Ly10 cells bearing empty (Ct; white), SPI1 WT (light gray), or QE (dark gray) constructs ($n=9$, mean \pm SD). OCI-Ly10 cells were treated for 24 hours with doxycycline and the indicated concentrations of lenalidomide. Cell viabilities are represented relative to those without lenalidomide treatment. Three independent experiments have been performed. $^{**}P \leq 0.01$; $^{***}P \leq 0.001$. Statistics were performed using two-tailed Student *t* test. Error bars, SD. **E**, SPI1 QE-expressing OCI-Ly10 cells are more sensitive than those expressing SPI1 WT to shRNAs targeting IRF4. Viability of OCI-Ly10 cells transduced with puromycin-inducible shRNAs [one control (Ct; white) and two targeting IRF4 (shIRF4-1, light gray; shIRF4-2, dark gray; $n=9$, mean \pm SD)]. OCI-Ly10 cells were treated for 24 hours (left) or 48 hours (right) with puromycin. Cell viabilities are represented relative to those of control shRNA. Three independent experiments have been performed. NS, not significant. $^*P \leq 0.05$; $^{**}P \leq 0.01$; $^{***}P \leq 0.001$. Statistics were performed using two-tailed Student *t* test. Error bars, SD. **F-I**, Overexpression of SPI1 WT and QE proteins in WM-related cell lines (BCWM.1 and MWCL-1). SPI1 QE-expressing cells grow more and are more sensitive to JQ1 and lenalidomide (Len) treatment than those expressing SPI1 WT. **F**, Quantitation of *in vitro* proliferation of BCWM.1-transduced cells expressing either SPI1 WT (light gray) or QE (dark gray; $n=9$, mean \pm SD). **G**, Viability of BCWM.1 cells bearing SPI1 WT (light gray) or QE (dark gray) constructs ($n=9$, mean \pm SD) when treated with the indicated concentration of JQ1 (left) or lenalidomide (right). **H**, Quantitation of *in vitro* proliferation of MWCL-1-transduced cells expressing either SPI1 WT (light gray) or QE (dark gray; $n=9$, mean \pm SD). **I**, Viability of MWCL-1 cells bearing SPI1 WT (light gray) or QE (dark gray) constructs ($n=9$, mean \pm SD) when treated with the indicated concentration of JQ1 (left) or lenalidomide (right). The levels of SPI1 proteins were checked (Supplementary Fig. S5A–S5D) and found to be equivalent for SPI1 WT and QE conditions. Cell viabilities are represented relative to those without JQ1 or lenalidomide treatment. Three independent experiments have been performed in triplicate. Statistics were performed using two-tailed Student *t* test. Error bars, SD. NS, not significant. $^*P \leq 0.05$; $^{**}P \leq 0.01$; $^{***}P \leq 0.001$.

the impact of *SPI1* QE mutation on these activities will have to be investigated.

Missense mutations in transcription factor genes in hematopoietic malignancies, such as those of *ELF4*, *EGR2*, *MEF2C*, or *IRF4* in mature B-cell disorders (15, 33, 34), may affect the DNA binding of the protein. Together with the recent description of missense mutations modifying the DNA-binding properties of Krüppel-like factor 5 gene (*KLF5*; ref. 35), our work provides a framework for understanding the functional consequences of missense mutations.

Oncogenic properties of ETS proteins have been largely described in solid tumors. Rearrangements of ETS loci are a hallmark of prostate cancer with around half of tumors showing alterations at an ETS gene locus (36). However, abnormalities of ETS proteins have rarely been identified in mature B-cell malignancies, with the exception of *ELF4* mutations in chronic lymphocytic leukemia (33), translocations involving *SPIB*, and gain of chromosome 11q24, encompassing *ETS1* and *FLI1*, in diffuse large B-cell lymphoma (DLBCL; refs. 37, 38).

Various members of the ETS family also regulate B-cell differentiation and function. In murine models, the roles of six ETS proteins have been studied in detail in B-lymphocytes, including *ETS1* (2, 27), *FLI1* (39), *GABPA* (40), and the three class III factors *SPI1*, *SPIB*, and *SPIC* (41). Survey of their expression during human and murine B-cell maturation suggest the highest activity of the class I/IIa ETS factors in GC B cells, together with the highest level of *ETS1* transcription (Supplementary Fig. S8B–S8E).

Loss of *ETS1* leads to premature B-cell differentiation into antibody-secreting cells (42), and recent reports have emphasized that *SPI1* and *SPIB* physiologically act as negative regulators of plasma cell development and sensors to environmental stimuli (8, 9, 43). Considering that WM is characterized by a differentiation block at the transition from post-GC lymphocyte to plasma cell, we confirmed the impact of *SPI1* mutations on terminal B-cell differentiation. Indeed, *SPI1*-expressing B cells demonstrated less efficient plasma cell differentiation in an *in vitro* model. Moreover, our gene-expression analyses revealed that mutated *SPI1* WM cases were less differentiated than the other WM samples. Additional analyses will be required to comprehensively characterize the relationship between the different classes of ETS in the different steps of B-cell differentiation and in the differentiation blockade observed in WM. In this regard, it is worth noting that GSEA of RNA-seq from normal B-cell populations shows a significant enrichment of classes I/IIa motifs in naïve, GC, or memory B-cell populations versus plasma cells (Supplementary Fig. S9D–S9I). It also remains to be shown whether other missense mutations in genes of ETS family (*ETS1*, *ELK1*) or transcription factors (*EGRI*, *EGR2*, *MEF2C*; refs. 15, 34) identified in other B-cell lymphoproliferative disorders may have the same impact on DNA recognition and B-cell differentiation.

Finally, *SPI1* Q226E mutations are associated with poor prognosis, although impact on survival needs to be validated. The enrichment of MYC pathway associated with *SPI1* Q226E-expressing cases is of interest, as BET inhibitors target this pathway and may be of therapeutic benefit in these high-risk patients. Similarly, *in vitro* experiments show that cells expressing *SPI1* Q226E are sensitive to lenalido-

mid, which is also known to indirectly target the important B-lymphocyte transcription factors. Our results thus identify mutant *SPI1* as a new genetic biomarker for patients with poor-prognosis WM and point to treatment options for this patient group.

METHODS

Detailed materials and methods are available in the Supplementary Data.

Patients and Materials

A total of 85 patients meeting the diagnostic criteria for WM (44), were enrolled in this study. Written informed consent for bone marrow and biological analyses were obtained from all patients in accordance with the Declaration of Helsinki and with ethical approval from the local ethics committee (CPP Ile-De-France 05/21/2014). Characteristics of this cohort are summarized in Supplementary Table S1. Cytogenetic analyses were performed as described previously (45). Samples from patients with DLBCL ($n = 147$), follicular lymphoma ($n = 56$), marginal zone lymphoma ($n = 12$), and mantle-cell lymphoma ($n = 19$), were obtained with their informed consent and the approval of the local Research Ethics Committees (Gustave Roussy, Centre Henri Becquerel, Pitié-Salpêtrière, and Cochin hospitals). Patients with chronic lymphocytic leukemia ($n = 360$) were enrolled in two clinical trials (www.ClinicalTrials.gov; NCT00564512 and NCT00645606). Primer sequences are available upon request.

Statistical Analysis

Data are depicted as mean \pm SD. *P* values were calculated with the two-tailed unpaired Student *t* test unless otherwise specified. All statistical analyses were performed with Prism software version 6.0 (GraphPad Software, Inc.).

Clinical and laboratory variables were compared across patients with or without mutation using the Wilcoxon rank-sum test (for quantitative variables) or the Fisher exact test (for qualitative variables). Time-to-treatment (time between diagnosis and first treatment) was compared across groups using log-rank tests. For overall survival (survival since study enrollment) a Kaplan–Meier estimator was used and survival curves were compared using the log-rank test. All tests were two-sided, with *P* value less than 0.05 considered statistically significant. Softwares were SAS 9.3 (SAS, Inc.) and R 3.0.2.

Cell Sorting and Nucleic Acid Extraction

Samples were processed as described previously (34). Tumor cells were defined by CD19 and monotypic light chain positivity. Controls were matched CD3⁺ T cells. The cellular populations were sorted using an Influx Cell Sorter (BD Biosciences). Sorted cells were subjected to an additional round of sorting to reach purity over 98%. Nucleic acids were extracted using the AllPrep DNA/RNA Kit (Qiagen).

Sequencing

Whole-exome sequencing (WES) was conducted on paired samples from 16 patients as described previously (34). Library was built with SureSelect All Exon Kit (Agilent Technologies). Sequences were analyzed as described previously (34) with minor modifications (Supplementary Methods; Supplementary Tables S2 and S3; Supplementary Fig. S1A–S1C).

Twenty recurrently mutated genes from our WES cohort or in the literature were selected for targeted sequencing of an extended series of WM samples using Ion AmpliSeq Designer (Thermo Fisher

Scientific). See Supplementary Table S4 for genomic coordinates of targeted regions and Supplementary Fig. S1D for the list of genes. Accession numbers for sequence data are E-MTAB-7881, E-MTAB-7882, and EGAS00001003603.

RNA-seq

RNA-seq was performed and analyzed as described previously (refs. 34, 46; see also Supplementary Material). Molecular pathway analysis and GSEA were performed using GSEA (2.2.3), the Molecular Signatures Database (6.2, Broad Institute) and the Lymphoid Signature Database (21). Comparison with normal B-cell differentiation expression profiles was performed using Blueprint data (accession numbers: EGAD00001002347, EGAD00001002414, EGAD00001001164, EGAD00001002452, EGAD00001002323, EGAD00001002476).

qPCR Analyses

cDNA was generated using SuperScript IV Reverse Transcriptase (Invitrogen). TaqMan probes were purchased from Applied Biosystems: ABL1 (Hs1104728), GUSB (Hs00939627), IRF4 (Hs01056533), MYC (Hs00905030). The expression level of each gene was assessed by qRT-PCR with ABI Prism 7500 and calculated following the $\Delta\Delta C_t$ method. Each sample was analyzed in triplicate and normalized to ABL1 or GUSB expression as indicated.

Cell Culture and Drugs

All cell lines were maintained in standard conditions and verified for known molecular features (described in Supplementary Fig. S11A–S11C). OCI-Ly10 cells were a kind gift from K. Leroy (Institut Cochin, Paris, France). BCWM.1 and MWCL-1 cells were a kind gift from R. Guizé and R. Lema (CHU Clermont-Ferrand, France). OCI-Ly10, BCWM.1, and MWCL-1 cells were negative for *SPI1* Q226E mutation (Supplementary Fig. S11A). Ba/F3 cells were a kind gift from P. Dubreuil (CRCM, Marseille, France) and were supplemented with 1 ng/mL of mlL3. The BET inhibitor JQ1 was a kind gift from J. Bradner (Dana-Farber Cancer Institute, Boston, MA). Lenalidomide was purchased from Aldrich Chemistry (reference CDS022536).

Vectors, Retroviral and Lentiviral Production, Transduction, and Transplantation

SPI1 (NM_003120.2), *MYD88* (NM_002468.4), and derivative cDNAs were subcloned into murine stem cell virus (MSCV)-IRES-GFP, pV81, or pcDNA3. Mutations were introduced using the Quick Change site-directed Mutagenesis Kit (Agilent) or by PCR. Every PCR-amplified or mutagenized fragment was sequence-checked. Viral particles, transduction procedures, bone marrow transplantation assays, and hematopoietic differentiation analyses were performed as described previously (47, 48).

Cells were transduced with WT or mutated pV81/GFP-*SPI1* viruses and flow-sorted twice gating on viable GFP⁺ cells 24 hours after doxycycline addition (1,000 ng/mL; Sigma, reference D-9891) to reach more than 90% of GFP⁺ cells. These flow-sorted cells were used for subsequent analyses. Both mutant and WT proteins were expressed at similar levels (Supplementary Fig. S4B and SSA–5D). Cells transduced by lentivirus containing shRNA targeting human IRF4 [TRCN000014765 (shIRF4-1) and TRCN000014767 (shIRF4-2; Sigma)] were grown 3 days in puromycin (1,000 ng/mL) for selection.

EMSA

The *SPI1* core-binding sites GGAA boxes (in bold) and have been described previously (5'GAGGAAACCGCGGAATCAGGAAGAACTGGCCGCGC3'; ref. 18). EMSA was performed using nonradioactive reagents, as described in Supplementary Methods. Three independent experiments were performed.

Isothermal Calorimetric Assay

Interactions between the WT or mutant DNA-binding domains of *SPI1* fused to GST (49) and DNA containing "lambda B1," a classic *SPI1*-binding site (19), and mutants thereof, were analyzed using a VP-ITC calorimeter (Microcal). Prior to measurements, all solutions were degassed under vacuum. The reaction cell was loaded with GST-DBD WT or QE and the syringe was filled with 20mer DNA substrates containing the *SPI1*-binding site. The lambda B1-derived sequences used as probes were as follows: WT, TAAAAGGAA GTGAAACCAAG; mutant same as WT, with a mutant sequence, TAAACCAAGTGAACCAAG; ETS1-like same as WT with an ETS1-like sequence, TAACCGGAAGTGAACCAAG. The proteins were extensively dialyzed against the buffer (25 mmol/L Na phosphate pH 8.0, 150 mmol/L NaCl, 5 mmol/L β -mercaptoethanol). K_d were obtained by nonlinear least-squares fitting of the experimental data using the single set of independent binding sites model of the Origin software provided with the instrument. All binding experiments were performed in duplicate or triplicate at 25°C.

Luciferase Assay

HEK293T cells were transiently transfected with equimolar amounts of empty pcDNA3 vector (control, ctrl), pcDNA3 vectors encoding WT or mutant *SPI1*, along with modified firefly luciferase reporter constructs and a *Renilla* luciferase control reporter. Cells were harvested 24 hours after transfection and the Dual Luciferase Reporter Assay (Promega) was used for detecting luciferase activity. A minimum of three independent experiments in optimized conditions were performed. Checking for protein expression indicated that the mutant protein was consistently overexpressed with respect to the WT protein. Quantifying protein expression of over 5 Western blots (3 of them are shown in Supplementary Fig. S31–S3K) indicated an average 1.5 times stronger *SPI1* protein expression in *SPI1* QE over WT. Activity was normalized with respect to protein quantity and to *Renilla* activity from a cotransfected vector. Reporters used were Fes (18), lambda B1 (19), and GPAA1 sequences (3 copies) fused to the TK promoter. Sequences were as follows: Fes, GAGGAACCGC GGAATCAGGAAGAACTGGCCGCGC; GPAA1 WT, TGGGACC GGAAGTGC GGG; GPAA1 mutant, TGGGACCCCAAGTGC GGG; lambda B1 WT, TAAAAGGAAGTGAACCAAG; lambda B1 mutant, TAAACCAAGTGAACCAAG. The core ETS recognition sequence is in bold; mutant bases are underlined.

Cellular Proliferation Assays

Doxycycline (1,000 ng/mL) was added to exponentially growing cells. GFP⁺ cells were flow-sorted after 24 hours and plated in triplicate in 48-well plates, 90,000 cells/well, in 300 μ L of media containing doxycycline (1,000 ng/mL). The time of flow-sorting was considered to be "hour 0" (H0). The number of viable cells was determined by counts of Trypan blue-negative cells and the percentage of viable (Sytox) GFP⁺ cells was checked by flow cytometry at H48. Two independent experiments were performed in triplicate. Transduced BCWM.1 and MWCL-1 cells were analyzed similarly, in adapted culture conditions (75,000 or 150,000 cells/wells, respectively, in a 24-well plate). Three independent experiments were performed in triplicate.

Drug-Sensitivity Assays

OCI-Ly10 transduced cells were seeded at 300,000 cells/mL in triplicate in 300 μ L of fresh media, in a 48-well plate. JQ1 and lenalidomide were dissolved in DMSO and equal volumes were added to cells to reach the indicated final concentrations. OCI-Ly10 transduced cells were treated with doxycycline (1,000 ng/mL) and the indicated concentrations of JQ1 or lenalidomide for, respectively, 24 hours and 48 hours. Doxycycline, JQ1, and lenalidomide were added to the media at the same time (H0). BCWM.1 and MWCL-1 transduced cells were

seeded at 150,000 or 250,000 cells/well, respectively, in 500 μ L of fresh media, and analyzed similarly. Cell numbers were estimated as described above. Three independent experiments were performed in triplicate.

Flow Cytometry and Cell-Cycle Analyses

Cells analyzed by flow cytometry were stained in PBS 1 \times supplemented with 2% FBS for 30 minutes at 4°C with washing prior to FACS analysis. Antibodies were purchased from Becton Dickinson or eBioscience for human (h) and murine (m) antibodies. Viability of cells was confirmed by using the Sytox Blue (Invitrogen) viability marker. Cells were analyzed on a FACS CantoII with FACS Diva Software (BD Biosciences). For cell-cycle analyses, cells were fixed in 70% cold ethanol and stained with propidium iodide and 7-AAD. Quantification of different cycle phases was performed using Watson Pragmatic Algorithm (FlowJo). Alternatively, cells were labeled by the addition of 10 mmol/L BrdU (5-bromodeoxyuridine) to the culture medium for 1 hour before harvest. BrdU and 7-AAD were detected with the APC BrdU Detection Kit (BD Biosciences). Stained cells were analyzed with a FACS Canto II (BD Biosciences) and data analyses were performed with FlowJo 9.5.2 Software (TreeStar).

Chromatin Immunoprecipitation, Sequencing, and Analysis

ChIP protocol, adapted from the MagnaChIP Kit Protocol (Millipore) and ref. 50, and analyses are detailed in the Supplementary Material. ChIP-seq data for ELK1 (ENCFF556JBS), ETS1 (ENCFF332PGQ), GABPA (ENCFF116EXQ), and SPI1 (ENCFF002CHQ) in the GM12878 cell line were obtained from ENCODE. The first 500 peaks, unique to the transcription factor with the highest P_{adj} (as determined by HOMER merge peaks, v4.10.3), were selected and attributed to the nearest gene. Genes unique to a transcription factor were then identified (Supplementary Table S17) and used in GSEA.

Murine B-cell Transduction and Culture

All experiments were performed on 8- to 10-week-old C57BL/6 mice (The Jackson Laboratory) and procedures conformed to ethical principles and guidelines revised and approved by the Animal Care Committee of Gustave Roussy. Red cells were lysed from splenic suspensions for 5 minutes at 4°C. After washing (PBS 1 \times ; 2% FBS), CD43⁻ cells were separated using AutoMACS Pro Separator Instrument (Miltenyi Biotec) with anti-CD43 antibody-coated magnetic beads and comprised more than 95% B220⁺ cells (Supplementary Fig. S7C). CD43⁻ cells were activated with anti-IgM (7 μ g/mL) and anti-CD40 (10 μ g/mL) for 24 hours before retroviral transduction. Transduced B cells were then cultured in media containing anti-CD40 (10 μ g/mL). The number of viable cells was determined by counts of Trypan blue-negative cells and the percentage of GFP⁺ Sytox⁻ cells was determined by flow cytometry. Transduced cells were washed and seeded at 1×10^6 cells/mL in fresh complete RPMI without any mitogen supplement. The start of the mitogen-free cultures was designated as "H0." The number of GFP⁺ cells was monitored by counting Trypan blue-negative cells and by flow cytometry.

Disclosure of Potential Conflicts of Interest

F. Damm reports receiving a commercial research grant from Novartis. E. Verhoeven has ownership interest in a patent covering the BAEV glycoprotein pseudotyped lentiviral vectors for transduction of human B cells, T cells, and hematopoietic stem cells. V. Leblond has received speakers bureau honoraria from Roche, Janssen, Gilead, AbbVie, and Amgen, and is a consultant/advisory board member for AstraZeneca, Janssen, and Roche. V. Ribrag reports receiving commercial research grants from Servier and ArgenX; has received speakers bureau honoraria from Infinity, BMS, Pharmamar, NanoString, MSD, Roche, Gilead, Incyte, and Epizyme; and is a

consultant/advisory board member for Servier. P. Morel is a consultant/advisory board member for Janssen. P. Vyas reports receiving a commercial research grant from Celgene; has received speakers bureau honoraria from Celgene, Pfizer, Jazz, Astellas, Daiichi Sankyo, and Novartis; and has ownership interest (including stocks and patents) in Oxstem Oncology. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: D. Roos-Weil, F. Nguyen-Khac, O.A. Bernard
Development of methodology: D. Roos-Weil, V. Della-Valle, H. Ghamlouch, C. Hérate, E. Mylonas, F.-L. Cosset, E. Verhoeven, P. Vyas, T. Mercher, S. Aoufouchi, N. Droin, C. Guillouf, F. Nguyen-Khac

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): D. Roos-Weil, C. Decaudin, M. Armand, V. Della-Valle, H. Ghamlouch, V. Ropars, C. Hérate, D. Lara, E. Durot, R. Haddad, E. Mylonas, F. Damm, F. Pflumio, B. Stoilova, M. Metzner, V. Camara-Clayette, F.-L. Cosset, V. Leblond, V. Ribrag, P. Cornillet-Lefebvre, P. Rameau, N. Azar, F. Charlotte, P. Morel, J.-B. Charbonnier, P. Vyas, T. Mercher, N. Droin, C. Guillouf, F. Nguyen-Khac
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): D. Roos-Weil, V. Della-Valle, M.K. Diop, H. Ghamlouch, V. Ropars, C. Hérate, D. Lara, E. Mylonas, O. Elemento, P. Dessen, V. Ribrag, F. Charlotte, J.-B. Charbonnier, P. Vyas, T. Mercher, N. Droin, C. Guillouf, F. Nguyen-Khac, O.A. Bernard
Writing, review, and/or revision of the manuscript: D. Roos-Weil, V. Della-Valle, M.K. Diop, D. Lara, E. Durot, E. Mylonas, F. Damm, B. Stoilova, E. Verhoeven, V. Leblond, V. Ribrag, P. Cornillet-Lefebvre, N. Azar, F. Charlotte, P. Morel, J.-B. Charbonnier, P. Vyas, T. Mercher, S. Aoufouchi, N. Droin, C. Guillouf, F. Nguyen-Khac, O.A. Bernard
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): C. Guillouf

Study supervision: D. Roos-Weil, P. Vyas, F. Nguyen-Khac, O.A. Bernard

Acknowledgments

We thank the members of the Gustave Roussy platforms, including Patrick Gonin for excellent mouse care, Yann Lecluse for flow cytometry, Xavier Veaute and the CIGEX facility for protein production, and Patricia Kannouche for helpful discussions. We thank the Blueprint consortium for access to RNA-seq data, the Encode consortium for some ChIP-seq data, and W. Cook (La Trobe University, Melbourne, Australia) and R. DeKoter (Western University in London, Ontario) for cell lines. This work was supported by grants from INSERM, Institut National du Cancer (INCa), 2013-PLBIO-09, 2016-PLBIO-068, JTC 2014-143, INCa-DGOS-INSERM_12551, INCa-DGOS-Inserm_12560, Association Laurette Fugain (to O.A. Bernard and F. Nguyen-Khac), LNCC, Comité Val-d'Oise de la Ligue contre le cancer, infrastructure FRISBI ANR-10-INBS-05, Fondation ARC (SFI20111203530), GEFLUC, and FILO. P. Vyas acknowledges funding from the MRC Disease Team Awards (G1000729/94931 and MR/L008963/1), MRC Molecular Haematology Unit and the Oxford Partnership Comprehensive Biomedical Research Centre (NIHR BRC Funding scheme.oxfbr-2012-1). H. Ghamlouch is a recipient of successive grants from région Ile de France and Fondation de France (No 00067113). C. Hérate, D. Lara, and E. Durot were recipients of fellowships from region Ile de France, Fondation pour la Recherche Médicale, and Fondation ARC, respectively.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

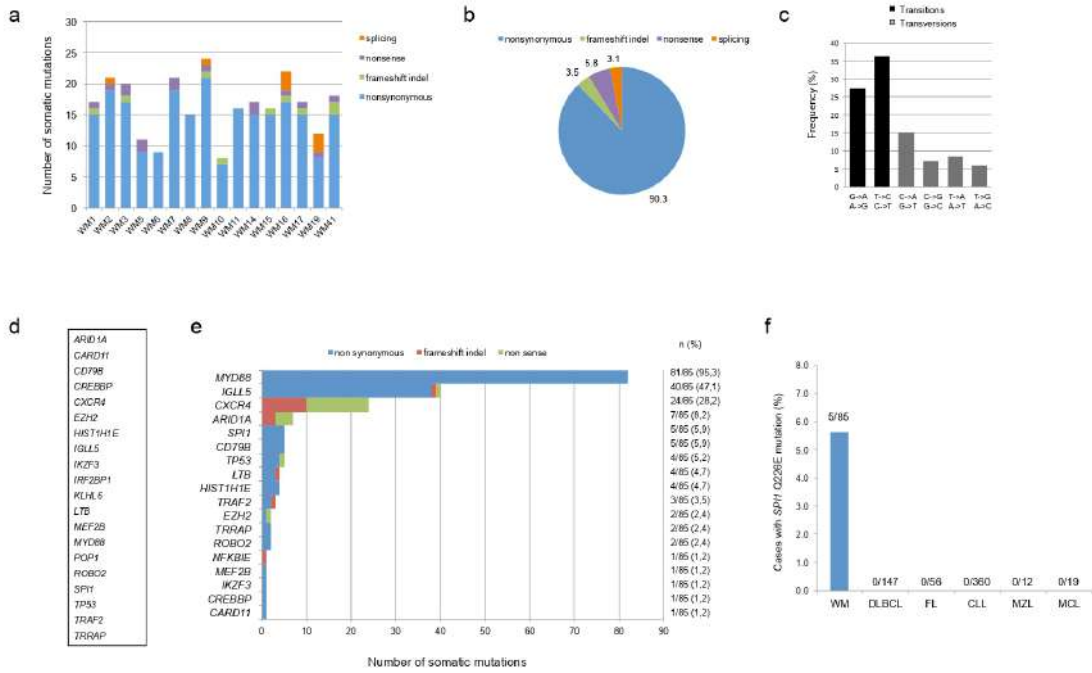
Received August 1, 2018; revised March 28, 2019; accepted April 18, 2019; published first April 24, 2019.

REFERENCES

- Basso K, Dalla-Favera R. Germinal centres and B cell lymphomagenesis. *Nat Rev Immunol* 2015;15:172–84.
- Garrett-Sinha LA. Review of Ets1 structure, function, and roles in immunity. *Cell Mol Life Sci* 2013;70:3373–90.
- Sharrocks AD. The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol* 2001;2:827–37.
- Hollenhorst PC, McIntosh LP, Graves BJ. Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu Rev Biochem* 2011;80:437–71.
- Scott EW, Simon MC, Anastasi J, Singh H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* 1994;265:1573–7.
- Cook WD, McCaw BJ, Herring C, John DL, Foote SJ, Nutt SL, et al. PU.1 is a suppressor of myeloid leukemia, inactivated in mice by gene deletion and mutation of its DNA binding domain. *Blood* 2004;104:3437–44.
- Rosenbauer F, Wagner K, Kutok JL, Iwasaki H, Le Beau MM, Okuno Y, et al. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat Genet* 2004;36:624–30.
- Carotta S, Willis SN, Hasbold J, Inouye M, Pang SH, Emslie D, et al. The transcription factors IRF8 and PU.1 negatively regulate plasma cell differentiation. *J Exp Med* 2014;211:2169–81.
- Willis SN, Tellier J, Liao Y, Trezise S, Light A, O'Donnell K, et al. Environmental sensing by mature B cells is controlled by the transcription factors PU.1 and Spi-B. *Nat Commun* 2017;8:1426.
- Mueller BU, Pabst T, Osato M, Asou N, Johansen LM, Minden MD, et al. Heterozygous PU.1 mutations are associated with acute myeloid leukemia. *Blood* 2002;100:998–1007.
- Lavallee VP, Baccelli I, Kros J, Wilhelm B, Barabe F, Gendron P, et al. The transcriptomic landscape and directed chemical interrogation of MLL-rearranged acute myeloid leukemias. *Nat Genet* 2015;47:1030–7.
- Seki M, Kimura S, Isobe T, Yoshida K, Ueno H, Nakajima-Takagi Y, et al. Recurrent SPI1 (PU.1) fusions in high-risk pediatric T cell acute lymphoblastic leukemia. *Nat Genet* 2017;49:1274–81.
- Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature* 2011;470:115–9.
- Treon SP, Xu L, Yang G, Zhou Y, Liu X, Cao Y, et al. MYD88 L265P somatic mutation in Waldenstrom's macroglobulinemia. *N Engl J Med* 2012;367:826–33.
- Reddy A, Zhang J, Davis NS, Moffitt AB, Love CL, Waldrop A, et al. Genetic and functional drivers of diffuse large B cell lymphoma. *Cell* 2017;171:481–94.
- Morel P, Duhamel A, Gobbi P, Dimopoulos MA, Dhodapkar MV, McCoy J, et al. International prognostic scoring system for Waldenstrom macroglobulinemia. *Blood* 2009;113:4163–70.
- Wei GH, Badis G, Berger MF, Kivioja T, Palin K, Enge M, et al. Genome-wide analysis of ETS-family DNA-binding *in vitro* and *in vivo*. *EMBO J* 2010;29:2147–60.
- Ray-Gallet D, Mao C, Tavittian A, Moreau-Gachelin F. DNA binding specificities of Spi-1/PU.1 and Spi-B transcription factors and identification of a Spi-1/Spi-B binding site in the *c-fes/c-fps* promoter. *Oncogene* 1995;11:303–13.
- Brass AL, Zhu AQ, Singh H. Assembly requirements of PU.1-Pip (IRF-4) activator complexes: inhibiting function *in vivo* using fused dimers. *EMBO J* 1999;18:977–91.
- Ernst J, Kellis M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat Protoc* 2017;12:2478–92.
- Shaffer AL, Wright G, Yang L, Powell J, Ngo V, Lamy L, et al. A library of gene expression signatures to illuminate normal and pathological lymphoid biology. *Immunol Rev* 2006;210:67–85.
- Laslo P, Pongubala JM, Lancki DW, Singh H. Gene regulatory networks directing myeloid and lymphoid cell fates within the immune system. *Semin Immunol* 2008;20:228–35.
- Wang JQ, Jeelall YS, Beutler B, Horikawa K, Goodnow CC. Consequences of the recurrent MYD88(L265P) somatic mutation for B cell tolerance. *J Exp Med* 2014;211:413–26.
- Ceribelli M, Kelly PN, Shaffer AL, Wright GW, Xiao W, Yang Y, et al. Blockade of oncogenic I kappa B kinase activity in diffuse large B-cell lymphoma by bromodomain and extraterminal domain protein inhibitors. *Proc Natl Acad Sci U S A* 2014;111:11365–70.
- Yang Y, Shaffer AL 3rd, Emre NC, Ceribelli M, Zhang M, Wright G, et al. Exploiting synthetic lethality for the therapy of ABC diffuse large B cell lymphoma. *Cancer Cell* 2012;21:723–37.
- Solomon LA, Li SK, Piskorz J, Xu LS, DeKoter RP. Genome-wide comparison of PU.1 and Spi-B binding sites in a mouse B lymphoma cell line. *BMC Genomics* 2015;16:76.
- Saelee P, Kearly A, Nutt SL, Garrett-Sinha LA. Genome-wide identification of target genes for the key B cell transcription factor Ets1. *Front Immunol* 2017;8:383.
- Cowley DO, Graves BJ. Phosphorylation represses Ets-1 DNA binding by reinforcing autoinhibition. *Genes Dev* 2000;14:366–76.
- Foulds CE, Nelson ML, Blaszcak AG, Graves BJ. Ras/mitogen-activated protein kinase signaling activates Ets-1 and Ets-2 by CBP/p300 recruitment. *Mol Cell Biol* 2004;24:10954–64.
- Hosokawa H, Ungerback J, Wang X, Matsumoto M, Nakayama KI, Cohen SM, et al. Transcription Factor PU.1 represses and activates gene expression in early T cells by redirecting partner transcription factor binding. *Immunity* 2018;49:782.
- Hong W, Kim AY, Ky S, Rakowski C, Seo SB, Chakravarti D, et al. Inhibition of CBP-mediated protein acetylation by the Ets family oncoprotein PU.1. *Mol Cell Biol* 2002;22:3729–43.
- de la Rica L, Rodriguez-Ubreva J, Garcia M, Islam AB, Urquiza JM, Hernando H, et al. PU.1 target genes undergo Tet-2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. *Genome Biol* 2013;14:R99.
- Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, Bahlo J, et al. Mutations driving CLL and their evolution in progression and relapse. *Nature* 2015;526:525–30.
- Damm F, Mylonas E, Cosson A, Yoshida K, Della Valle V, Mouly E, et al. Acquired initiating mutations in early hematopoietic cells of CLL patients. *Cancer Discov* 2014;4:1088–101.
- Zhang X, Choi PS, Francis JM, Gao GF, Campbell JD, Ramachandran A, et al. Somatic superenhancer duplications and hotspot mutations lead to oncogenic activation of the KLF5 transcription factor. *Cancer Discov* 2018;8:108–25.
- Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 2005;310:644–8.
- Lenz G, Nagel I, Siebert R, Roschke AV, Sanger W, Wright GW, et al. Aberrant immunoglobulin class switch recombination and switch translocations in activated B cell-like diffuse large B cell lymphoma. *J Exp Med* 2007;204:633–43.
- Bonetti P, Testoni M, Scandurra M, Ponzoni M, Piva R, Mensah AA, et al. Deregulation of ETS1 and FLI1 contributes to the pathogenesis of diffuse large B-cell lymphoma. *Blood* 2013;122:2233–41.
- Zhang XK, Moussa O, LaRue A, Bradshaw S, Molano I, Spyropoulos DD, et al. The transcription factor Flt-1 modulates marginal zone and follicular B cell development in mice. *J Immunol* 2008;181:1644–54.
- Xue HH, Bollenbacher-Reilly J, Wu Z, Spolski R, Jing X, Zhang YC, et al. The transcription factor GABP is a critical regulator of B lymphocyte development. *Immunity* 2007;26:421–31.
- DeKoter RP, Geadah M, Khoosal S, Xu LS, Thillainadesan G, Torchia J, et al. Regulation of follicular B cell differentiation by the related E26 transformation-specific transcription factors PU.1, Spi-B, and Spi-C. *J Immunol* 2010;185:7374–84.
- John SA, Clements JL, Russell LM, Garrett-Sinha LA. Ets-1 regulates plasma cell differentiation by interfering with the activity of the transcription factor Blimp-1. *J Biol Chem* 2008;283:951–62.
- Lu D, Nakagawa R, Lazzaro S, Staudacher P, Abreu-Goodger C, Henley T, et al. The miR-155-PU.1 axis acts on Pax5 to enable efficient terminal B cell differentiation. *J Exp Med* 2014;211:2183–98.

44. Owen RG, Treon SP, Al-Katib A, Fonseca R, Greipp PR, McMaster ML, et al. Clinicopathological definition of Waldenstrom's macroglobulinemia: consensus panel recommendations from the Second International Workshop on Waldenstrom's Macroglobulinemia. *Semin Oncol* 2003;30:110-5.
45. Nguyen-Khac F, Lambert J, Chapiro E, Grelier A, Mould S, Barin C, et al. Chromosomal aberrations and their prognostic value in a series of 174 untreated patients with Waldenstrom's macroglobulinemia. *Haematologica* 2013;98:649-54.
46. Quek L, Otto GW, Garnett C, Lhermitte L, Karamitros D, Stoilova B, et al. Genetically distinct leukemic stem cells in human CD34⁺ acute myeloid leukemia are arrested at a hemopoietic precursor-like stage. *J Exp Med* 2016;213:1513-35.
47. Malinge S, Ragu C, Della-Valle V, Pisani D, Constantinescu SN, Perez C, et al. Activating mutations in human acute megakaryoblastic leukemia. *Blood* 2008;112:4220-6.
48. Quivoron C, Couronne L, Della Valle V, Lopez CK, Plo I, Wagner-Ballon O, et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell* 2011;20:25-38.
49. Halber M, Tavittian A, Moreau-Gachelin F. The transcription factor Spi-1/PU.1 binds RNA and interferes with the RNA-binding protein p54nrb. *J Biol Chem* 1996;271:11177-81.
50. Ridinger-Saison M, Boeva V, Rimmel P, Kulakovskiy I, Gallais I, Levassieur B, et al. Spi-1/PU.1 activates transcription through clustered DNA occupancy in erythroleukemia. *Nucleic Acids Res* 2012;40:8927-41.

Supplementary Figure S1



Supplementary Figure S1.

(a) Number and type of somatic mutations in each WM patient analyzed by whole exome sequencing (WES, $n = 16$). Each column indicate the number of mutations per patient and each color-code represent a different type of mutation (orange: splicing; violet: nonsense; green: frameshift indel; blue: non synonymous).

(b) Repartition of the type of mutations among all somatic mutations identified by WES (orange: splicing; violet: nonsense; green: frameshift indel; blue: non synonymous).

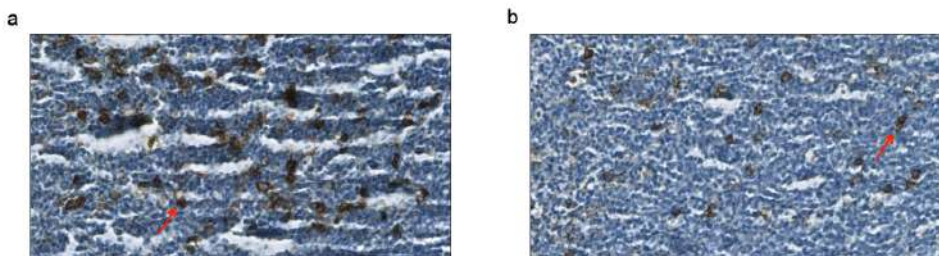
(c) Distribution of somatic coding-sequence transition (black) and transversion (grey) nucleotide interchanges among all mutations identified by WES.

(d) List of genes studied by targeted NGS.

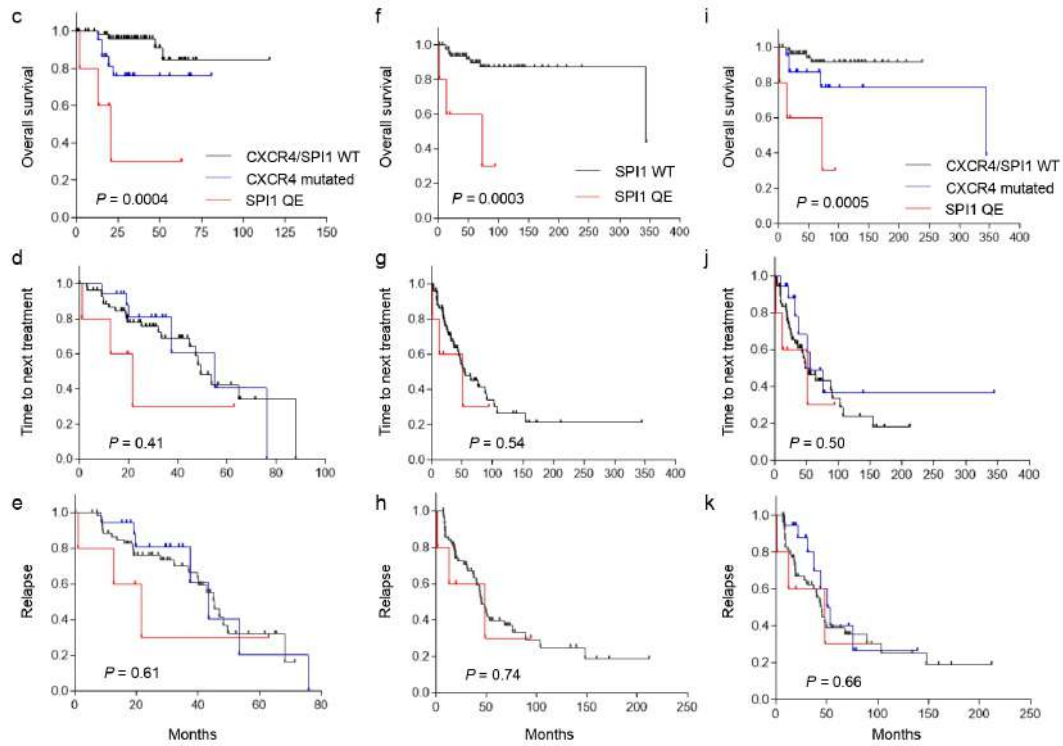
(e) Number and frequency of mutated genes identified by WES ($n = 16$) and targeted sequencing ($n = 69$). Each row represents a mutated gene in WM. The number and percentage of each mutated gene in the whole cohort are indicated on the right of the grid. Each color-code indicate a different type of mutation (violet: nonsense; green: frameshift indel; blue: non synonymous).

(f) Number and frequencies of *SPI1* Q226E mutations identified by NGS and/or Sanger sequencing, in WM and other B-cell lymphoproliferative disorders samples. The absolute numbers of mutations, out of the total analyzed samples, are shown above each bar. WM, Waldenström macroglobulinemia; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; CLL, chronic lymphocytic leukemia; MZL, marginal-zone lymphoma; MCL, mantle-cell lymphoma.

Supplementary Figure S2



Supplementary Figure S2

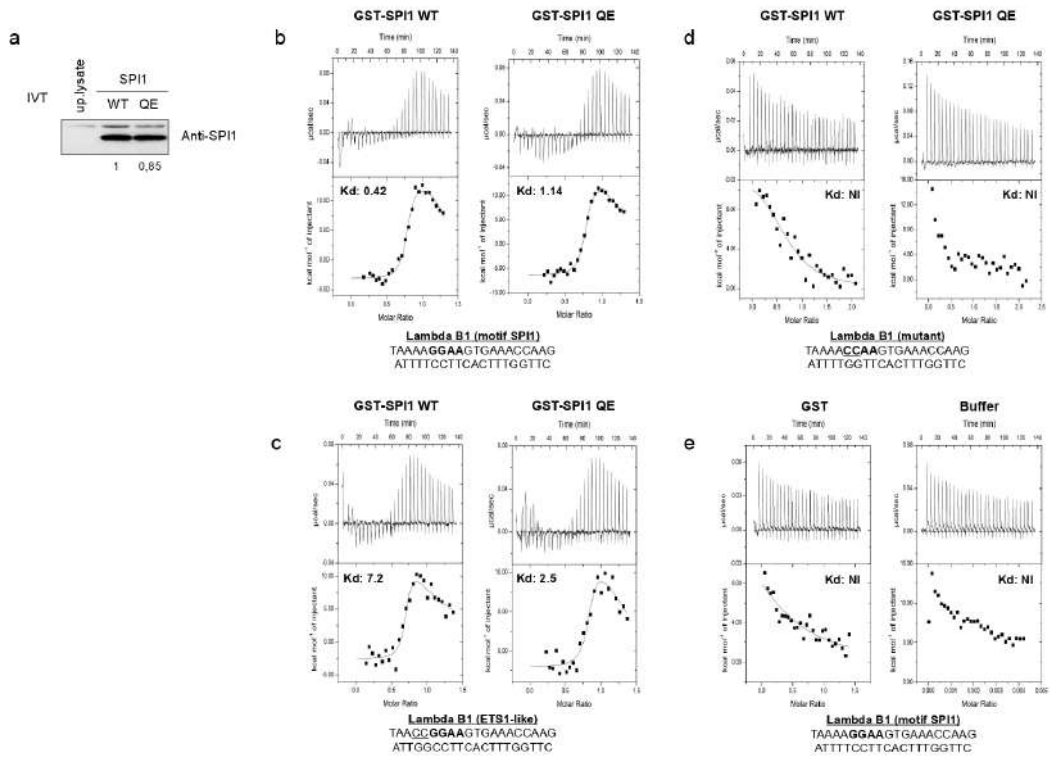


Supplementary Figure S2

(a-b) Immunohistochemical analysis of CD138 expression in bone marrow biopsy samples of *SPI1* WT (n = 8) and QE (n = 3) WM patients. Representative images of sections stained with a monoclonal antibody against CD138 show a higher proportion of CD138+ cells (0.25 mm/pixel at 40X magnification) for *SPI1* WT (left) compared to QE (right) samples. Red arrows indicate examples of CD138 positive cells.

(c-k) Kaplan-Meier survival curves of overall survival (c, f, i), time to next treatment (d, g, j) and relapse-free survival (e, h, k) for patients with (n = 5) and without (n = 80) *SPI1* Q226E mutations, for patients with *CXCR4* mutations (n = 22) and patients WT for both *SPI1* and *CXCR4* (n = 53). Survival times were calculated from dates of NGS (c, d, e) or WM diagnosis (f-k). P values are indicated and based on the log-rank test.

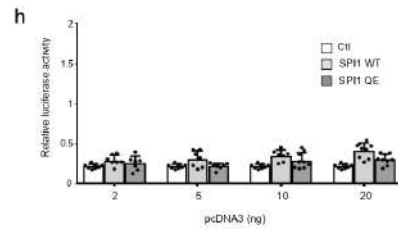
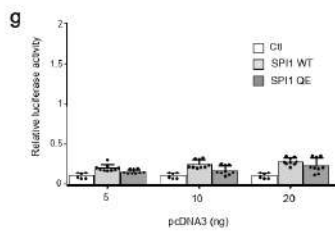
Supplementary Figure S3



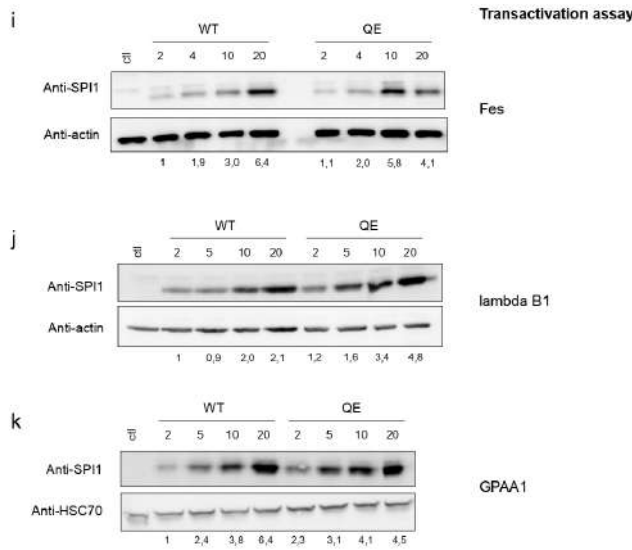
Supplementary Figure S3

f

Related to Figure S3	a	b	c	d
Model	2 sites	2 sites	2 sites	2 sites
Chi2	5.4*5	3.93*5	1.26*6	1.25*6
N1 (sites)	0.76 +/- 0.006	0.75 +/- 0.004	0.69 +/- 0.01	0.81 +/- 0.01
K1 (M-1)	2.37*9 +/- 5.08*8	8.79*8 +/- 1.32*8	1.38*8 +/- 5.10*7	3.93*8 +/- 1.42*8
ΔH1 (cal/mol)	-3194 +/- 271	-6407 +/- 230	-2605 +/- 400	-2119 +/- 410
ΔS1 (cal/mol/deg)	32.2	19.4	28.5	32.2
N2 (sites)	0.551 +/- 0	0.523 +/- 0	0.212 +/- 0	0.494 +/- 0
K2 (M1)	2.37*9 +/- 5.08*8	5.96*6 +/- 0	4.37*5 +/- 0	3.29*6 +/- 0
ΔH2 (cal/mol)	1.40*4 +/- 0	1.55*4 +/- 0	6.75*4 +/- 0	1.49*4 +/- 0
ΔS2 (cal/mol/deg)	80.2	83.2	252	80



Supplementary Figure S3



Supplementary Figure S3.

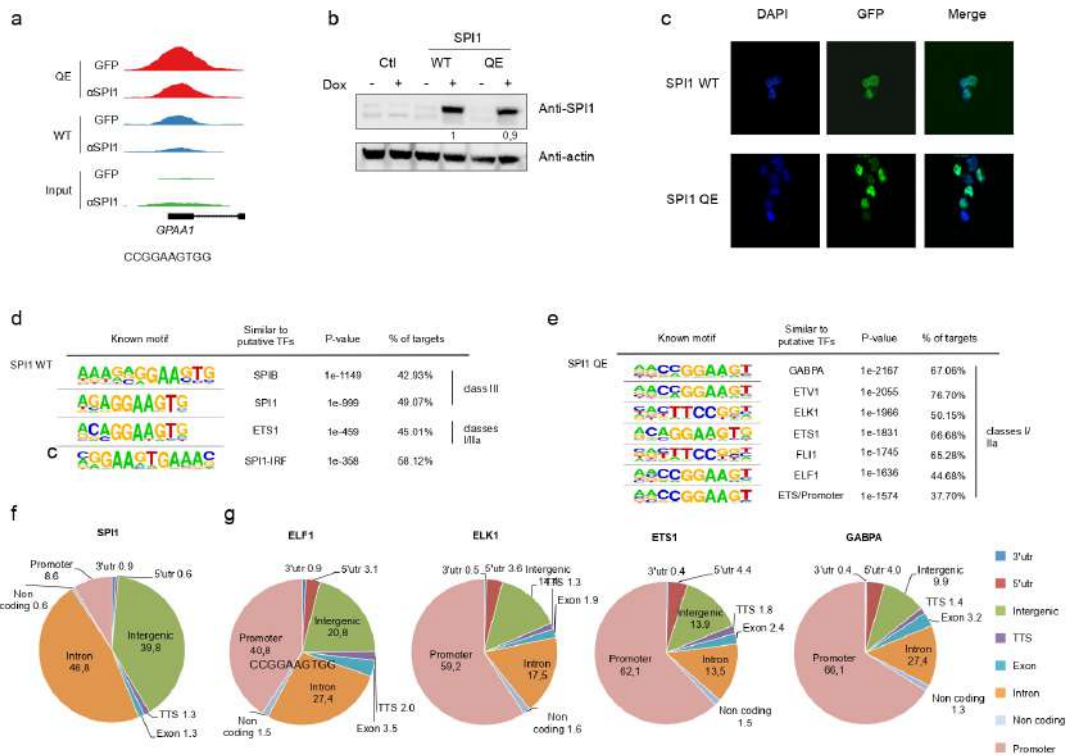
(a) Western blot evaluation of the quantities of *in vitro* produced SPI1 WT and QE proteins. Up, lysate, unprogrammed lysate. IVT, *in vitro* translation. Relative densitometric values are indicated below each band. SPI1 WT condition was set as 1.

(b-f) Summary of isothermal calorimetric data and representative isothermal curves. The used GST-proteins (SPI1 WT, left in b-d; SPI1 QE, right in b-d; GST, left and buffer, right in e) and lambda B1 probes (motif SPI1 (b, e), motif ETS1-like (c), mutant (d)) are indicated along with each dissociation constant (Kd, in nM). NI, no interaction. Details regarding reaction characteristics of each experiment are summarized in (f).

(g-h) Control of transactivation assays with mutant ETS-sites reporter vectors. HEK293T cells were transfected with increasing amounts of the indicated (control, SPI1 WT or QE) expression plasmids and fixed amounts of a luciferase reporter gene. Reporter genes contain mutant sequences of lambda B1 (TAAACC~~CA~~AGTGAAACCAAG) (g) and GPAA1 promoter (TGGGACC~~CC~~AAAGTGC~~GG~~) (h). Three independent experiments were performed (n = 6-9; mean ± s.d.). Relative luciferase activities were normalized with Renilla activity from a co-transfected vector. All errors bars represent standard deviation. Ctl, control; s.d, standard deviation.

(i-k) Western blots evaluation of the quantities of produced SPI1 and control (β-actin and HSC70) proteins upon transient transfection of 293T with the indicated vectors (Fes (i), lambda B1 (j) or GPAA1 (k)). Relative densitometric values were normalized against the levels of the internal control (β-actin or HSC70) conditions, set as 1 for SPI1 WT (2 ng) condition and are indicated below each band. Ctl, control.

Supplementary Figure S4



Supplementary Figure S4.

(a) Differential binding of SPI1 QE (red), SPI1 WT (blue) and input (green) on the *GPAA1* promoter in OCI-Ly10 transduced cells. The results of ChIP-seq data performed with GFP-Trap (top, GFP) and anti-SPI1 (bottom, α SPI1) antibodies are shown for each condition. The identified putative ETS binding site is represented at the bottom of the image.

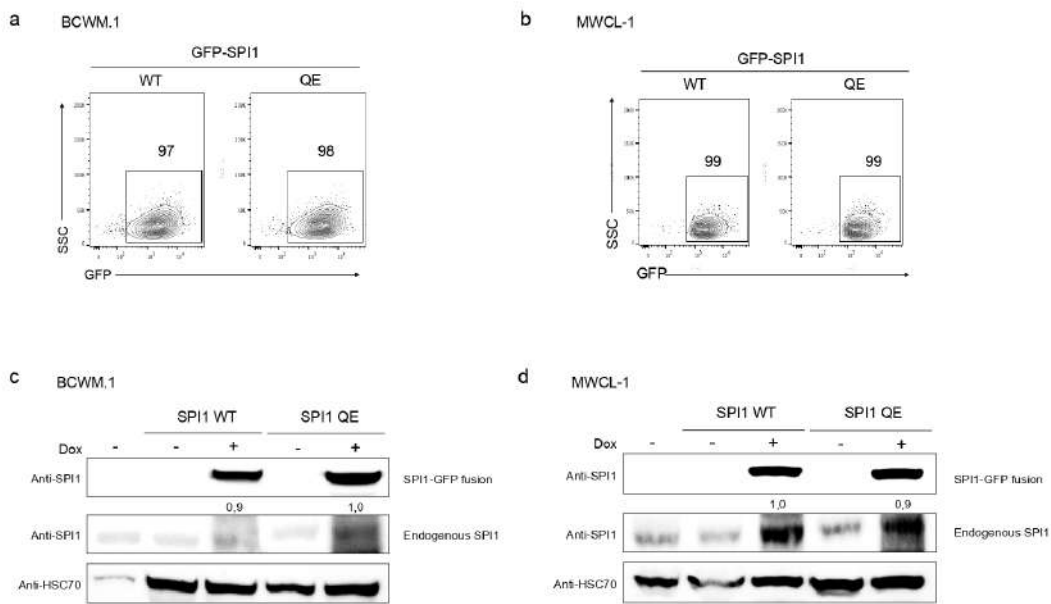
(b) Western blot evaluation of levels of SPI1 and β -actin proteins in OCI-Ly10 cells transduced with empty vector or doxycycline-inducible constructs encoding SPI1 WT or QE and incubated (+) or not (-) with doxycycline (1000 ng/mL) for 48h. Relative densitometric values were normalized against the levels of the internal control β -actin conditions and set as 1 for SPI1 WT condition. Ctl, control. Dox, doxycycline.

(c) OCI-Ly10 cells were first transduced by doxycycline-inducible constructs encoding SPI1 WT or QE and incubated with doxycycline for 48h, then prepared, and fluorescence was used to assess the expression and subcellular localization of GFP (green) using a fluorescence microscopy. Note that the SPI1 proteins are fused to GFP. DAPI-staining nuclei (blue) were also shown. The merge images confirmed that SPI1 WT and QE proteins localized in the nuclei. Images showed representative data from one of three individual experiments. Images were processed using ImageJ software.

(d-e) ChIP-seq analyses, performed with anti-SPI1 antibodies, in OCI-Ly10 cells transduced with doxycycline-inducible vectors expressing SPI1 WT or QE. Top of the list DNA motifs recognized by SPI1 WT (d) and QE (e) proteins. Homer known motif analysis was performed with a cutoff *P* value of 10^{-30} . The complete list of *de novo* and known motifs is available in **Supplementary Tables S6-9** and **S10-13** for ChIP performed respectively with anti-GFP-Trap and anti-SPI1 antibodies. Motifs, associated transcription factors, *P* values and percentages of peaks containing the motifs are indicated. TF, transcription factors.

(f-g) ChIP-seq analyses and genomic localization, defined by ENCODE annotations, of DNA sequences bound by SPI1 (class III) (f), and ELF1, ELK1, ETS1 and GABPA (classes I and IIa) (g), as indicated, in an EBV-transformed human B-cell line (GM12878). Different genomic regions and their respective percentage (in respect to the total number of peaks) are indicated and color-coded (color legends, left bottom).

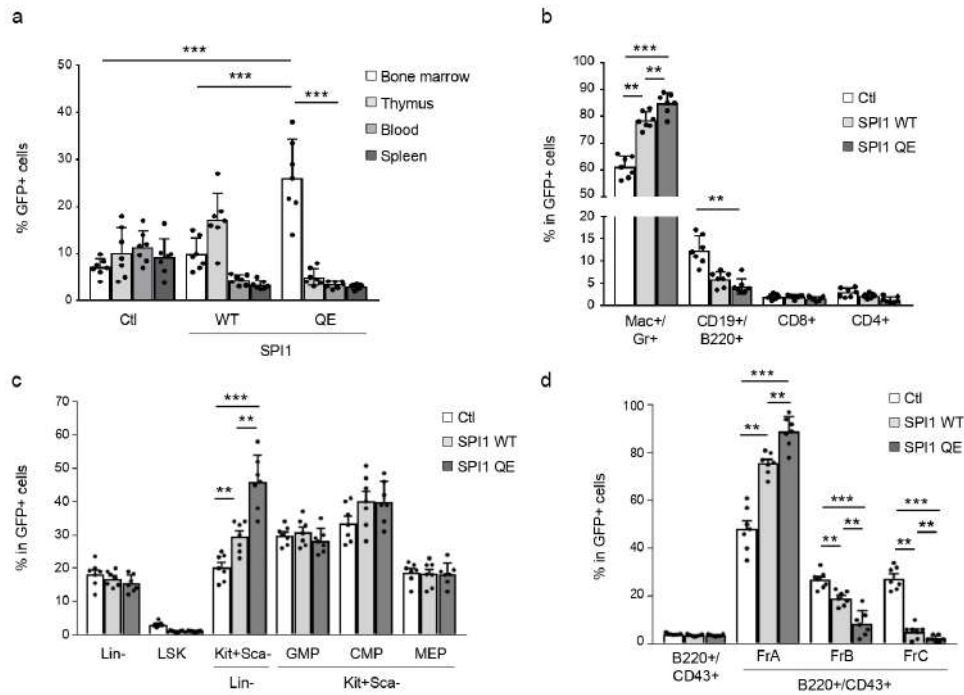
Supplementary Figure S5



Supplementary Figure S5.

(a-d) Analyses of SPI1 expression in WM-derived cells lines (BCWM.1 (a, c) and MWCL-1 (b, d) over expressing SPI1 WT or QE in a doxycycline (Dox)-inducible fashion. Both SPI1 WT and QE are expressed at comparable levels. Top (a, b): fluorescence of GFP-SPI1 fusion proteins evaluated by flow-cytometry. Percentages of GFP positive cells in each condition are indicated. Bottom (c, d): evaluation of SPI1 and HSC70 protein levels. Digits indicated relative SPI1 protein expression upon normalization with HSC70 signal.

Supplementary Figure S6



Supplementary Figure S6.

Analyses of bone marrow transplantation of progenitor cell transduced to express SPI1 WT or QE proteins.

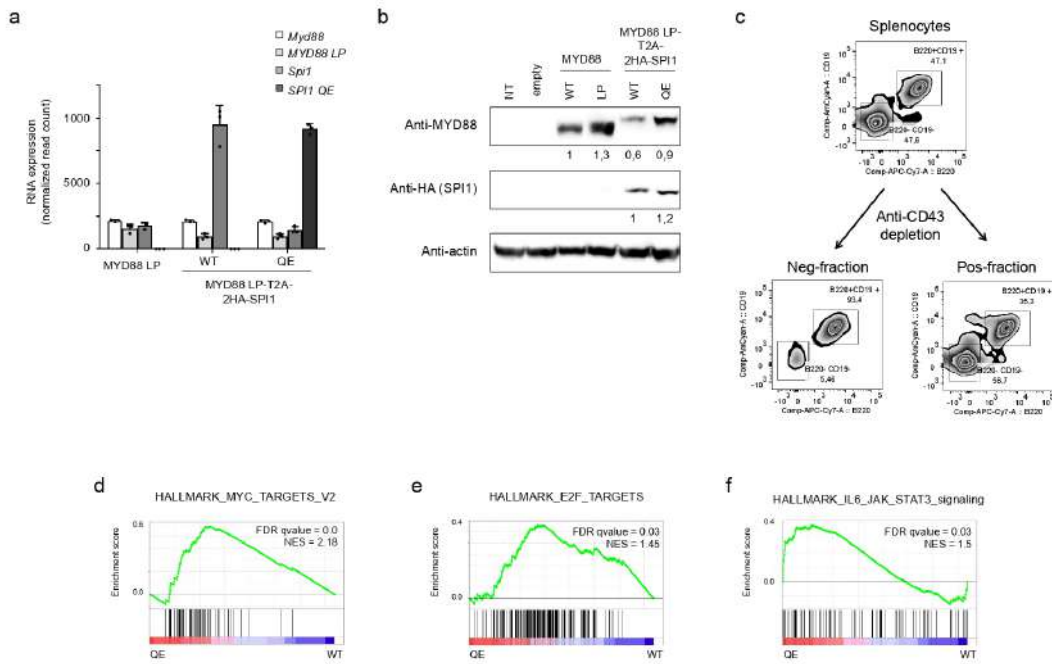
(a) Bone marrow transplantation (BMT) analyses of primary murine cells transduced by the indicated (control, SPI1 WT, SPI1 QE) retroviral vectors. The percentage of GFP positive cells in different murine hematopoietic organs is represented: bone marrow (white), thymus (light grey), blood (dark grey), spleen (black). Three independent experiments were performed ($n = 7$; mean \pm s.d.). Each dot corresponds to one mouse and all errors bars represent standard deviation. $*P \leq 0.05$, $**P \leq 0.01$, $***P \leq 0.001$. S.d: standard deviation. Flow-cytometry analyses have been performed at 5-7 months post-BMT. Ctl, control.

(b) Percentages of bone marrow mature myeloid (Mac⁺/Gr⁺), B- (CD19⁺/B220⁺) and T-cells (CD8⁺ and CD4⁺) in GFP positive fractions for the three conditions of transduction (control (white), SPI1 WT (light grey) and QE (dark grey)). Three independent experiments were performed ($n = 7$; mean \pm s.d.). Each dot corresponds to one mouse and all errors bars represent standard deviation. Only significant comparisons are indicated. $**P \leq 0.01$, $***P \leq 0.001$. S.d: standard deviation. Flow-cytometry analyses have been performed at 5-7 months post-BMT. Ctl, control.

(c) Percentages of bone marrow lineage negative and myeloid progenitors populations in GFP positive fractions for the three conditions of transduction (control (white), SPI1 WT (light grey) and QE (dark grey)). Three independent experiments were performed ($n = 7$; mean \pm s.d.). Each dot corresponds to one mouse and all errors bars represent standard deviation. Only significant comparisons are indicated. $**P \leq 0.01$, $***P \leq 0.001$. S.d: standard deviation. Flow-cytometry analyses have been performed at 5-7 months post-BMT. Ctl, control; Lin, lineage; LSK, Lin-Sca⁺Kit⁺; GMP, granulocyte-monocyte progenitors; CMP, common myeloid progenitors; MEP, megakaryocyte-erythrocyte progenitors.

(d) Percentages of bone marrow B-lymphoid progenitors populations in GFP positive fractions for the three conditions of transduction (control (white), SPI1 WT (light grey) and QE (dark grey)). Three independent experiments were performed ($n = 7$; mean \pm s.d.). Each dot corresponds to one mouse and all errors bars represent standard deviation. Only significant comparisons are indicated. $**P \leq 0.01$, $***P \leq 0.001$. S.d: standard deviation. Flow-cytometry analyses have been performed at 5-7 months post-BMT. Hardy's fractions (FrA, B and C) were gated in B220⁺/CD43⁺ cells and defined as previously described. Ctl, control; Fr, fraction.

Supplementary Figure S7



Supplementary Figure S7.

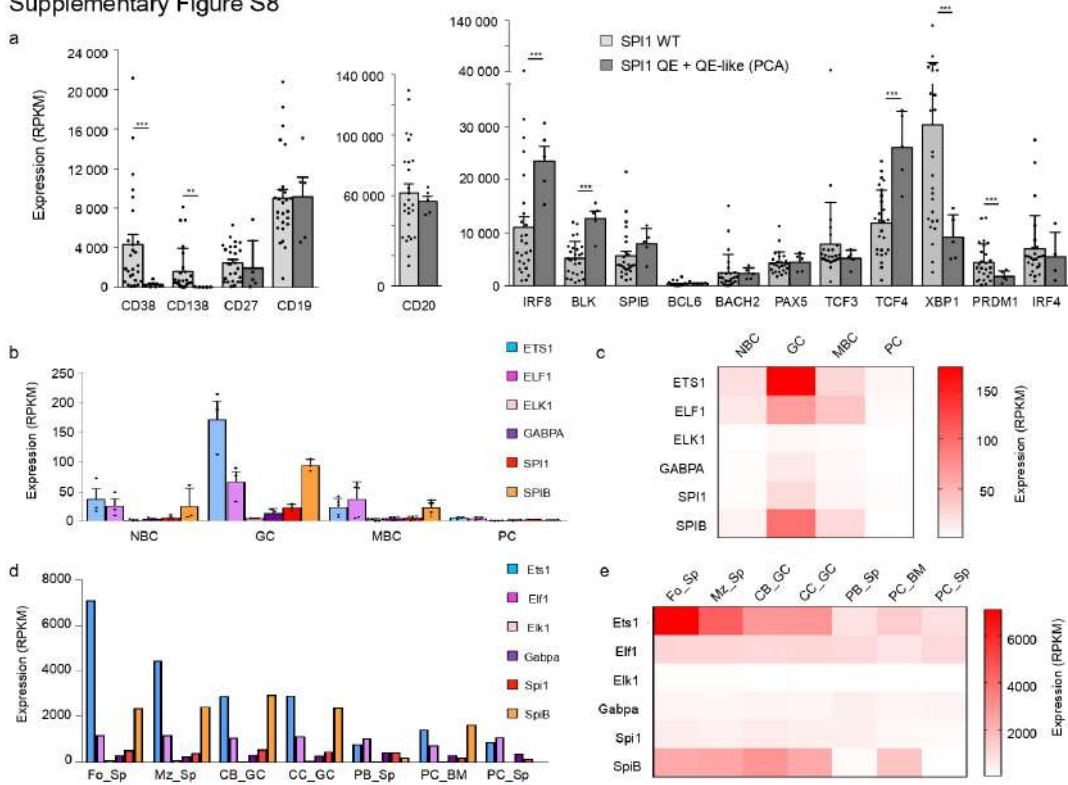
(a) Normalized read counts of WT and mutated *MYD88* and *SPI1* RNA in primary murine B-cells. Primary murine B-cells were transduced by retroviral vectors driving expression of MYD88 and SPI1 as WT or mutant species either alone or in combination (see Fig. 4a) and sorted on GFP expression.

(b) Expression levels of MYD88, HA-tagged SPI1 and β -actin in non-transduced Ba/F3 cells or Ba/F3 cells transduced by the indicated vectors. Relative densitometric values (shown for each band) were normalized against the levels of the internal control β -actin conditions. MYD88 WT was set as 1 for MYD88 expression and MYD88 LP-T2A-2HA-SPI1 WT for HA-tagged SPI1.

(c) Representative flow-cytometry analyses of primary murine B-cells before and after CD43 depletion. Cells were gated on the basis of CD19 and B220 expressions. Neg, negative; Pos, positive.

(d-f) GSEA enrichment plots from RNA-seq analyses obtained upon MYD88 LP and SPI1 WT or QE expression in murine B-cells showing enrichment in the QE condition for selected published dataset regarding proliferation, cell cycle, B-cell signaling and differentiation. FDR, false discovery rate. NES, normalized enrichment score.

Supplementary Figure S8



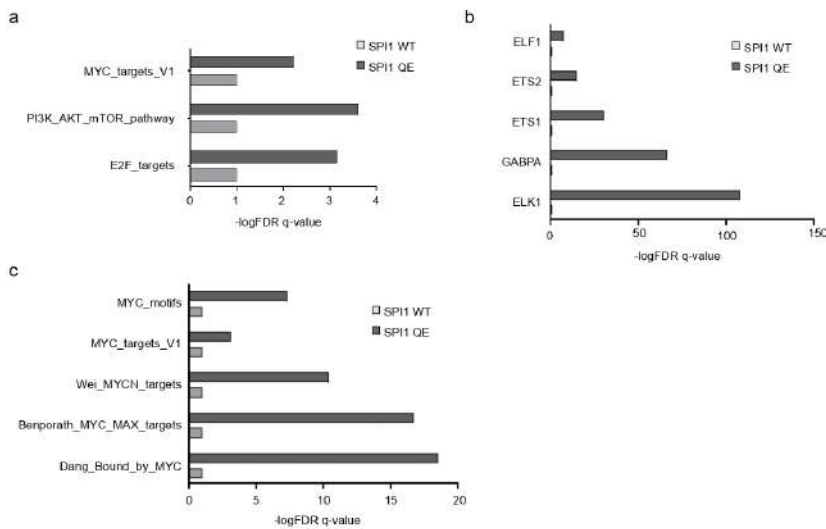
Supplementary Figure S8.

(a) Expressions of key genes controlling normal B-cell differentiation in *SPI1* WT ($n = 27$) compared to QE ($n = 3$) and QE-like ($n = 2$) WM RNA-seq samples. Error bars indicate standard deviation. Statistics were performed using Student *t* test. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$.

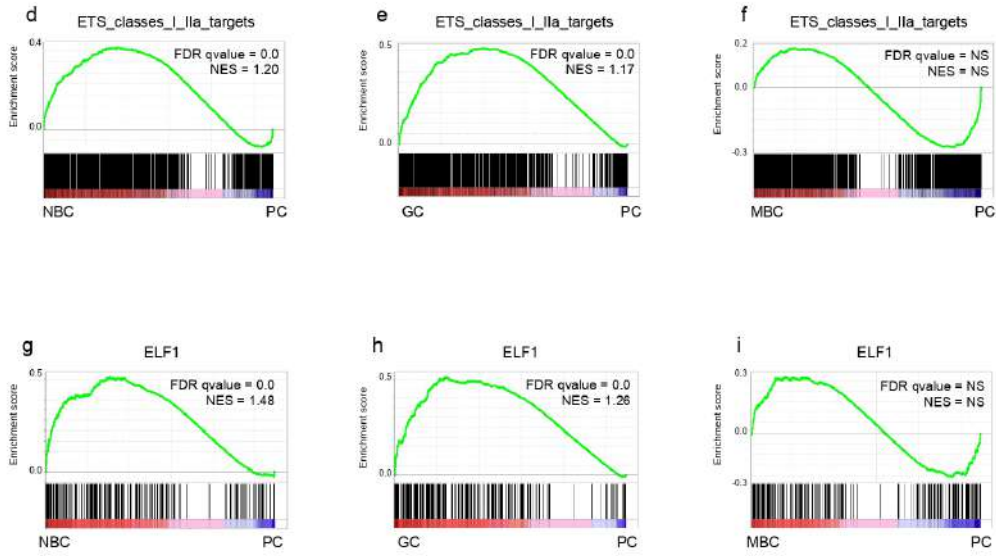
mRNA expressions (b) and representative heat map (c) of ETS classes I and IIa (ETS1, ELF1, ELK1, GABPA) and class III (SPI1, SPIB) proteins at different steps of human B-cell differentiation: naive B-cells (NBC, $n = 3$), germinal center (GC, $n = 3$), memory B-cells (MBC, $n = 5$) and plasma cells (PC, $n = 6$). These data were extracted from RNA-seq analyses of Blueprint databases. Error bars indicate standard deviation.

mRNA expressions (d) and representative heat map (e) of Ets classes I and IIa (Ets1, Elf1, Elk1, Gabpa) and class III (Spi1, Spib) proteins at different steps of murine B-cell differentiation: splenic follicular B-cells (Fo_Sp), splenic marginal B-cells (Mz_Sp), germinal center centroblasts (CB_GC), germinal center centrocytes (CC_GC), splenic plasmablasts (PB_Sp), bone marrow plasma cells (PC_BM) and splenic plasma cells (PC_Sp). These data were extracted from RNA-seq analyses of the ImmGen consortium (18)

Supplementary Figure S9



Supplementary Figure S9

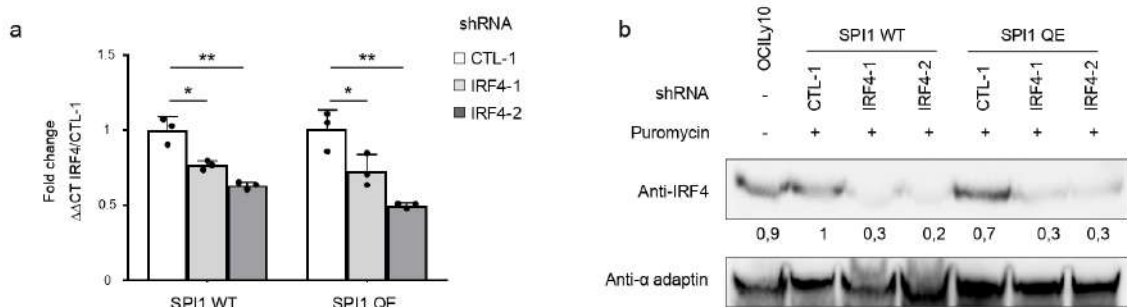


Supplementary Figure S9.

(a-c) GSEA enrichment in OCI-Ly10 active promoters regions (defined by ChromHMM). Histogram representations of MSigDB pathways significantly enriched in SPI1 QE (n = 3), with respect to WT (n = 3) condition, for main signaling pathways described in Figure 4a (a), ETS motifs (b) and MYC signatures (c). FDR q-value (-logFDR q-value) is represented. FDR, false discovery rate.

GSEA enrichment plots for ETS classes I and IIa targets (d-f) and ELF1 (g-i) comparing RNA-seq samples of naive B-cells (NBC, n = 8) (d, g), germinal center (GC, n = 3) (e, h) and memory B-cells (MBC, n = 5) (f, i) compared to plasma cells (PC, n = 6). FDR, false discovery rate. NES, normalized enrichment score.

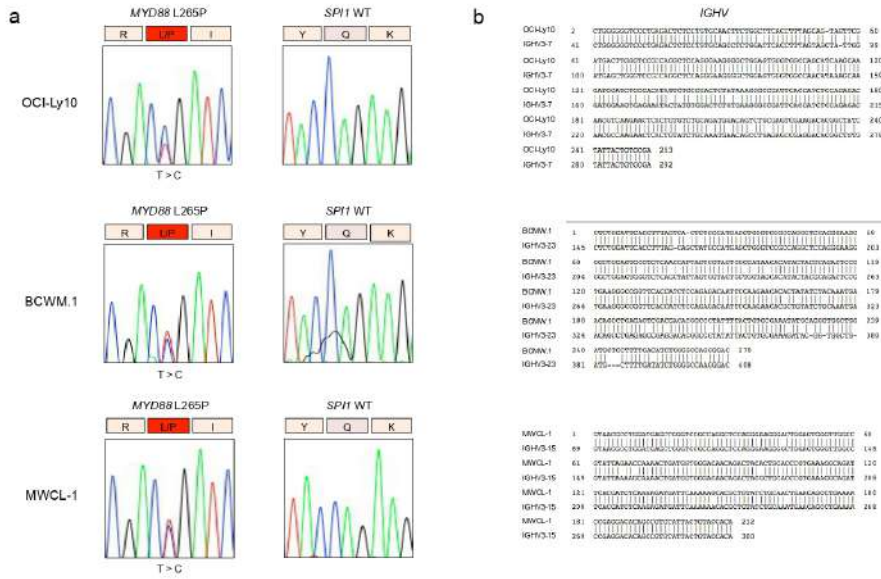
Supplementary Figure S10



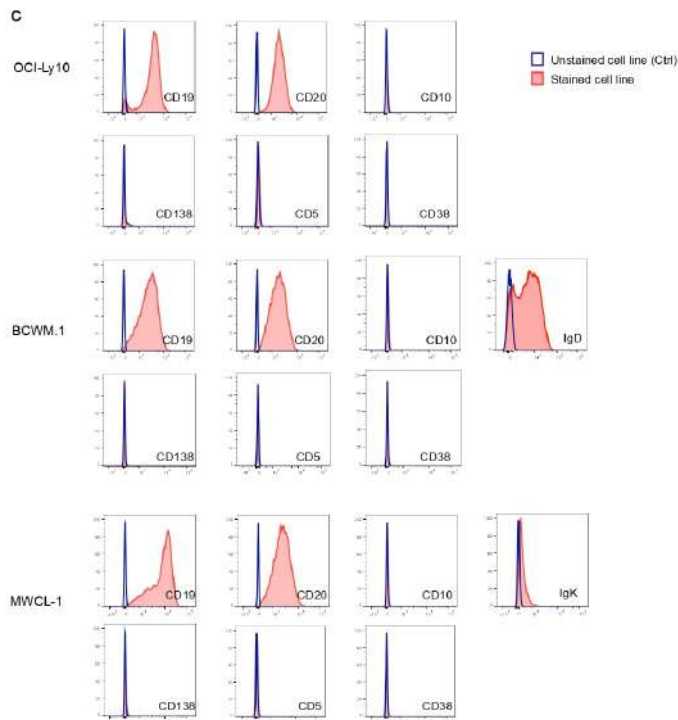
Supplementary Figure S10.

(a-b) Knock down of IRF4 expression by two shRNAs in OCI-Ly10 subclones over expressing SPI1 WT or QE. Evaluation of RNA (a) and proteins (b) levels of IRF4 and α-adaptin in OCI-Ly10 cells transduced (+) or not (-) with puromycin-inducible shRNAs (one control (ctl) and two targeting IRF4), by RT quantitative PCR and Western blotting respectively. Relative densitometric values were normalized against the levels of the internal control α-adaptin conditions and set as 1 for control SPI1 WT condition. Error bars indicate standard deviation. Statistics were performed using Student *t* test. **P* ≤ 0.05, ***P* ≤ 0.01, ****P* ≤ 0.001.

Supplementary Figure S11



Supplementary Figure S11



Supplementary Figure S11. Cell lines were checked for the presence of MYD88 L265P and SP11 QE mutations by Sanger sequencing (a), for IGHV mutational profile (b) and for expression of different B-cell surface markers by flow-cytometry (c).

2. Résultats supplémentaires

2.1 Contexte et objectifs

Les résultats sur les lignées cellulaires humaines et les cellules murines primaires présentés dans l'article 2 ont été obtenus en surexprimant SPI1. Ceci est un facteur limitant pour l'interprétation des effets biologiques du mutant car il est connu que la « dose » de SPI1 est d'une importance capitale pour l'engagement d'une cellule dans une voie de différenciation. Par exemple, au stade des progéniteurs multipotents, des taux élevés de SPI1 favoriseront l'engagement dans la lignée myéloïde au dépend de la lignée lymphoïde (15,346,348). Concernant la lignée lymphoïde B, il a été montré que la surexpression de SPIB, proche parent de SPI1, entraîne un blocage de la différenciation B en plasmocytes chez l'homme (475). Nous avons constaté le même effet en surexprimant SPI1 dans des lymphocytes B humains (données non montrées).

L'axe privilégié suite au travail publié dans l'article n°2 était donc d'obtenir des modèles avec une expression physiologique de l'allèle muté de *SPI1*.

L'aspect auquel je me suis plus particulièrement intéressée est l'étude de la coopération oncogénique entre MYD88 et SPI1, en utilisant deux approches. La première fait appel à un modèle murin développé dans l'équipe, knock-in pour la mutation de *Spi1*. Les lymphocytes B naïfs *Spi1*-mutés isolés à partir des rates de ces animaux ont ensuite été transduits avec un rétrovirus apportant la mutation de *MYD88*. La prolifération et la différenciation des cellules ont ensuite été analysées dans divers contextes. La seconde a consisté à modifier par CRISPR/Cas9 une lignée cellulaire humaine dérivée d'un patient ayant une MW (*MYD88* muté) afin d'y introduire la mutation de *SPI1*. Nous espérons ainsi identifier d'une part, les gènes et les mécanismes cellulaires dérégulés par la protéine SPI1 mutée au cours de la différenciation en plasmocytes, et d'autre part, les voies de signalisation activées par la coopération MYD88/SPI1.

2.2 Matériel et méthodes supplémentaires

2.1.1 Modèle murin SPI1QE-KI

Une lignée de souris avec un allèle knock-in pour la mutation Q226E (SPI1QE-KI) a été développée dans l'équipe (**Figure A**). La mutation SPI1QE se trouve sur le 5^{ème} et dernier exon de *SPI1*. La construction est telle que l'exon sauvage est floxé et un exon 5 alternatif avec la mutation se trouve en aval. En l'absence d'expression de la recombinaise Cre, le codon stop de l'exon sauvage arrête la traduction, l'exon muté n'est donc pas traduit. Lorsque la recombinaise Cre s'exprime, l'exon sauvage est excisé et l'exon muté est exprimé. Il n'y a donc dans tous les cas que deux copies du gène *Spi1* exprimées. Les souris SPI1QE-KI (C57bl/6) ont été croisées avec des souris Vav-Cre (C57bl/6),

aboutissant à l'expression de la mutation SPI1QE dans toutes les cellules hématopoïétiques. Toutes les expériences présentées dans cette partie ont été réalisées sur des animaux hétérozygotes pour la mutation SPI1QE (noté SPI1 WT/QE dans les résultats) et âgés de 10 à 12 semaines. Des animaux SPI WT/WT (noté SPI1 WT dans les résultats) du même âge ont été utilisés pour constituer les groupes contrôle.

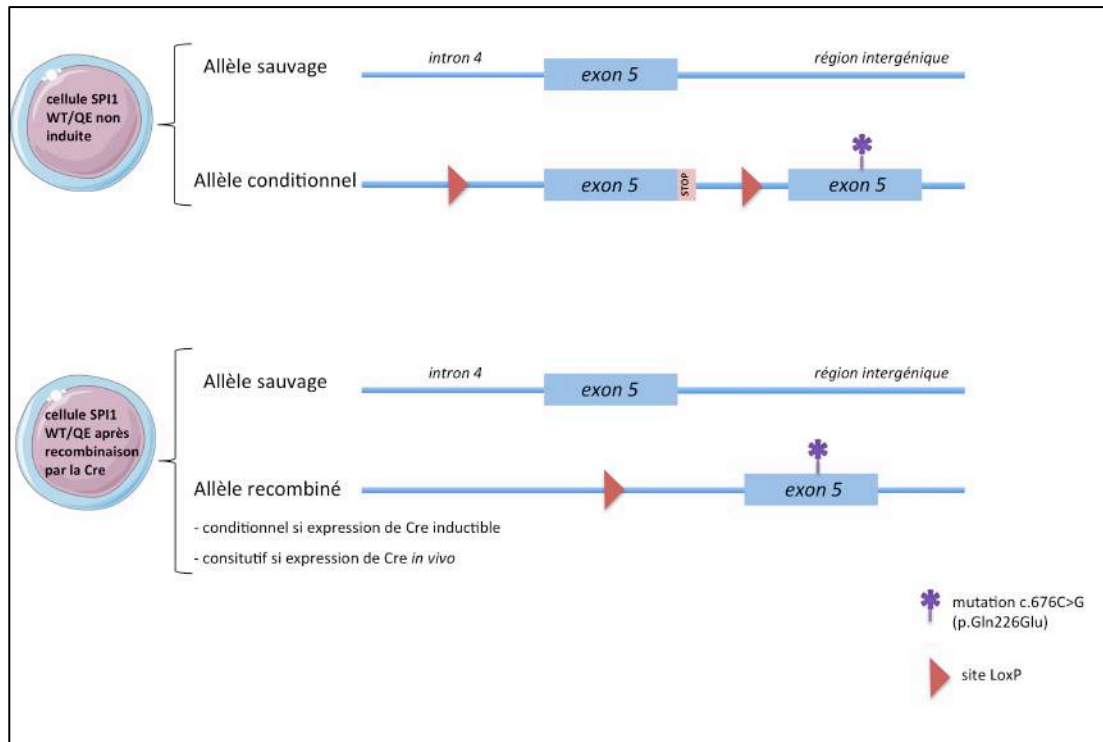


Figure A : Représentation schématique du transgène Spi1 Q226E dans les cellules hématopoïétiques

2.1.2 Stimulation LPS/IL2/IL5 et transduction rétrovirale

Les cellules B naïves CD43⁻ (isolées à partir de la rate de souris SPI1 WT/QE ou SPI1 WT selon le protocole décrit dans l'article n°1) ont été cultivées à raison de 3 millions de cellules par puits dans une plaque 6 puits et incubées pendant 24h dans le milieu de culture précédemment décrit complété par du LPS (20µg/ml), de l'IL-2 (20ng/ml) et de l'IL-5 (5ng/ml) avant transduction rétrovirale par les particules virales MSCV-MYD88LP-GFP ou MSCV-GFP. Les cellules B transduites ont ensuite été cultivées à raison de 200 000 cellules par puits dans une plaque 24 puits et incubées pendant 3 jours en présence de LPS (5 à 20 µg/ml selon les expériences), d'IL-2 (20ng/ml) et d'IL-5 (5ng/ml). La prolifération des cellules GFP⁺ a été suivie toutes les 24h, en même temps que l'expression du CD138, des marqueurs d'activation Fas et GL7, et du TLR4 (TLR4 conjugué à l'APC, Biolegend ; #145405).

2.3 Résultats

L'expression ectopique de MYD88LP dans les cellules SPI1 QE augmente l'expansion cellulaire et la différenciation B terminale.

Dans un premier temps, nous avons utilisé le modèle de culture *in vitro* utilisé dans la 1^{ère} partie du manuscrit, reposant sur l'utilisation d'une lignée de fibroblastes exprimant le CD40 et sécrétant du BAFF, couplée à l'ajout d'IL4 les quatre premiers jours de culture et d'IL21 les quatre suivants (**Figure 1**).

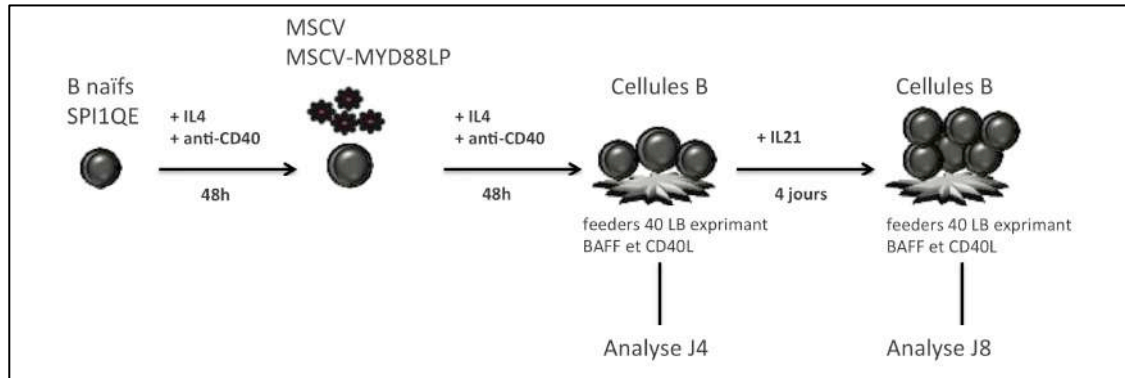


Figure 1. Représentation schématique du modèle *in vitro* CD40/IL4/IL21, adapté pour intégrer une étape de transduction rétrovirale

A J4, nous observons, selon les conditions, 5 à 10 fois plus de cellules GFP+ par rapport à J0. Le facteur d'expansion des cellules double-mutantes est comparable à celui des cellules SPI1QE-seul ou MYD88LP-seul. L'apport de la mutation MYD88 dans les cellules SPI1QE ne confère donc pas d'avantage prolifératif à ce stade de la culture par rapport aux cellules SPI1QE-seul ou par rapport aux cellules WT (**Figure 2A**).

A J8, nous observons, selon les conditions, 40 à 80 fois plus de cellules GFP+ par rapport à J4. Le facteur d'expansion des cellules double-mutantes est jusqu'à deux fois supérieur à toutes les autres conditions (**Figure 2B**).

Nous concluons donc qu'il y a un avantage de prolifération lorsque les mutations de SPI1 et de MYD88 sont présentes dans la même cellule.

A J4, après 4 jours de stimulation par CD40 et IL4, l'expression des marqueurs Fas et GL7 en cytométrie en flux permet d'identifier les BGC tandis que l'expression du CD138 permet d'identifier à la fois les plasmablastes (PB) et plasmocytes (PC) (**Figure 2C**). Nous constatons un pourcentage significativement plus élevé de BGC et moindre de cellules CD138+ dans la condition double-mutante par rapport à SPI1QE-seul (**Figure 2D et 2F**). Par rapport à des cellules WT transduites avec MYD88, la présence de la mutation de SPI1 n'entraîne pas d'augmentation de la proportion de BGC mais une augmentation de la proportion de cellules CD138+ (**Figure 2D et 2F**). La condition SPI1QE-seul entraîne une augmentation importante de la différenciation CD138+ dès J4, environ 3 fois supérieure à celle d'une cellule sauvage (**Figure 2D et 2E**).

A J8, après 4 jours de stimulation par CD40 et IL-21, l'expression des marqueurs Fas, GL7 et CD138 est de nouveau évaluée par cytométrie en flux (**Figure 2H**). Les cellules double-mutantes présentent significativement plus de cellules CD138+, en pourcentage et en valeur absolue, par rapport à toutes les autres conditions (**Figure 2I et 2J**), et le moins de BGC en proportion (**Figure 2K et 2L**).

Cela indique que sont des cellules engagées dans la différenciation plasmocytaire qui s'accumulent en contexte SPI1/MYD88 muté.

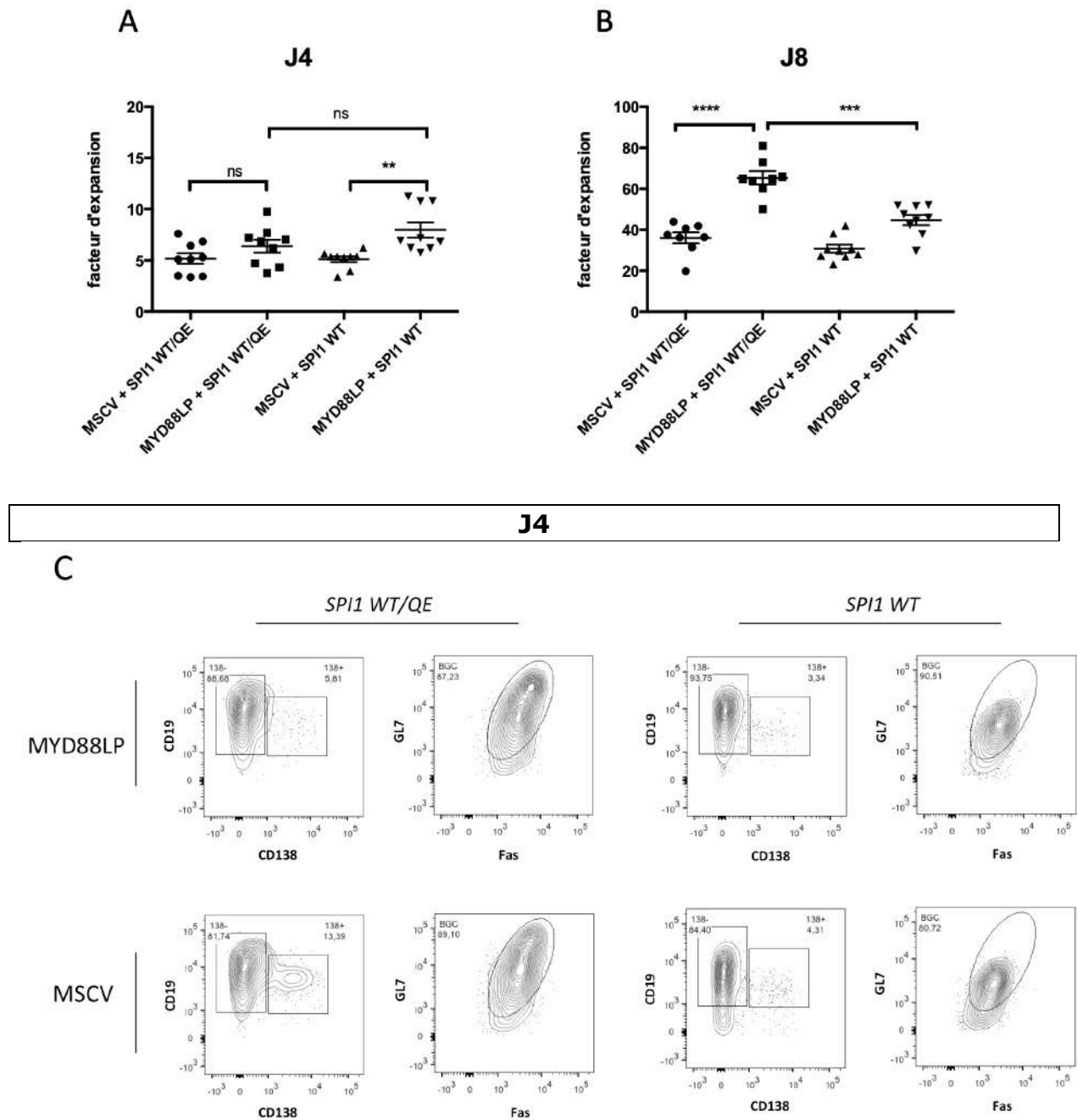
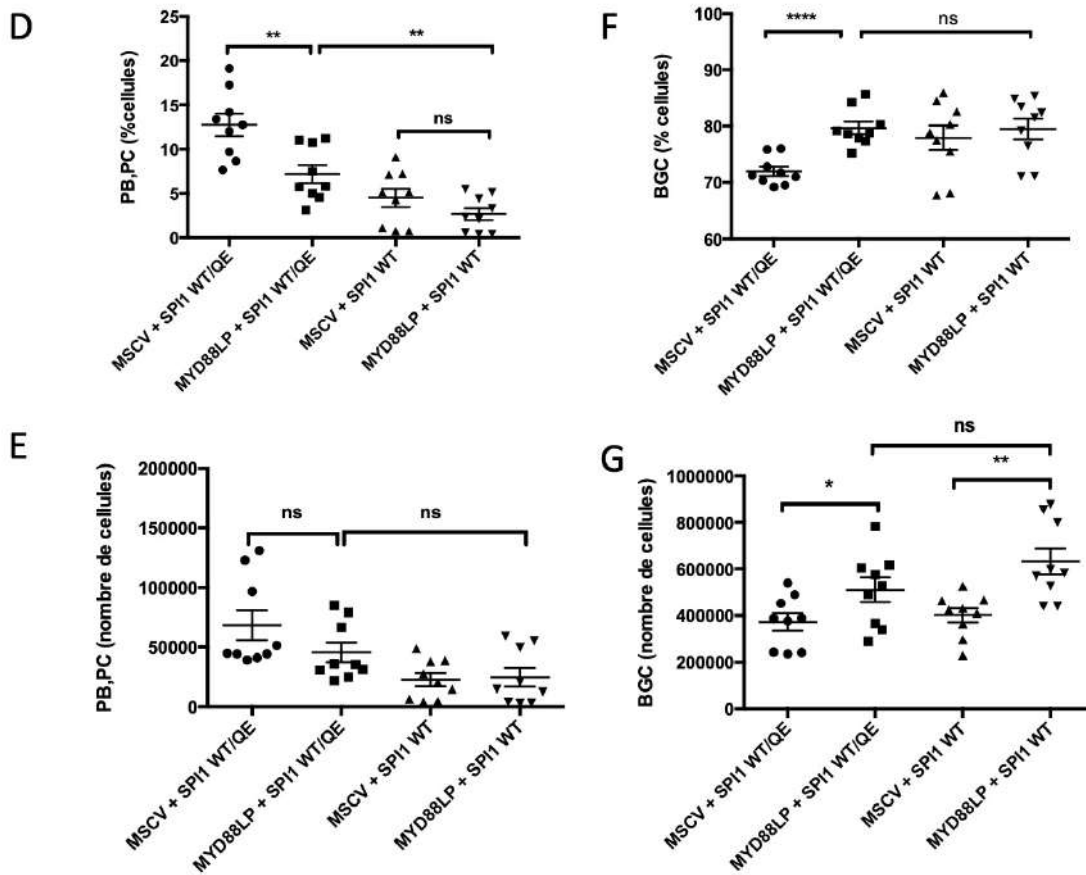


Figure 2. Stimulation des B naïfs spléniques par CD40/IL4/IL21

Après transduction par le vecteur apportant la mutation de MYD88 ou le vecteur contrôle, les cellules GFP+ ont été analysées à J4 et J8 par cytométrie en flux. **A**, Facteur d'expansion des cellules GFP+ entre J0 et J4. **B**, Facteur d'expansion des cellules GFP+ entre J4 et J8. **C**, Exemple de profil FACS à J4.



J8

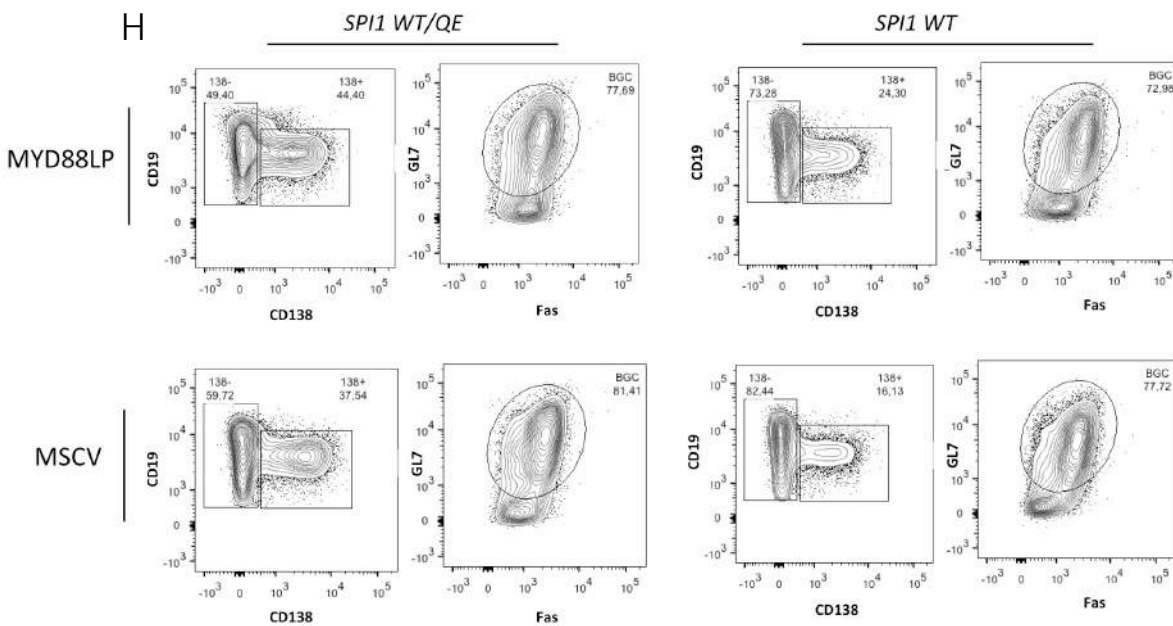


Figure 2. Stimulation des B naïfs spléniques par CD40/IL4/IL21 (suite)

D-G, Proportion (D et F) et valeur absolue (E et G) des plasmablasts (PB) et plasmocytes (PC) (CD19^{low} CD138⁺) et des BGC (CD19⁺GL7⁺Fas⁺) à J4. **H**, Exemple de profil FACS à J8.

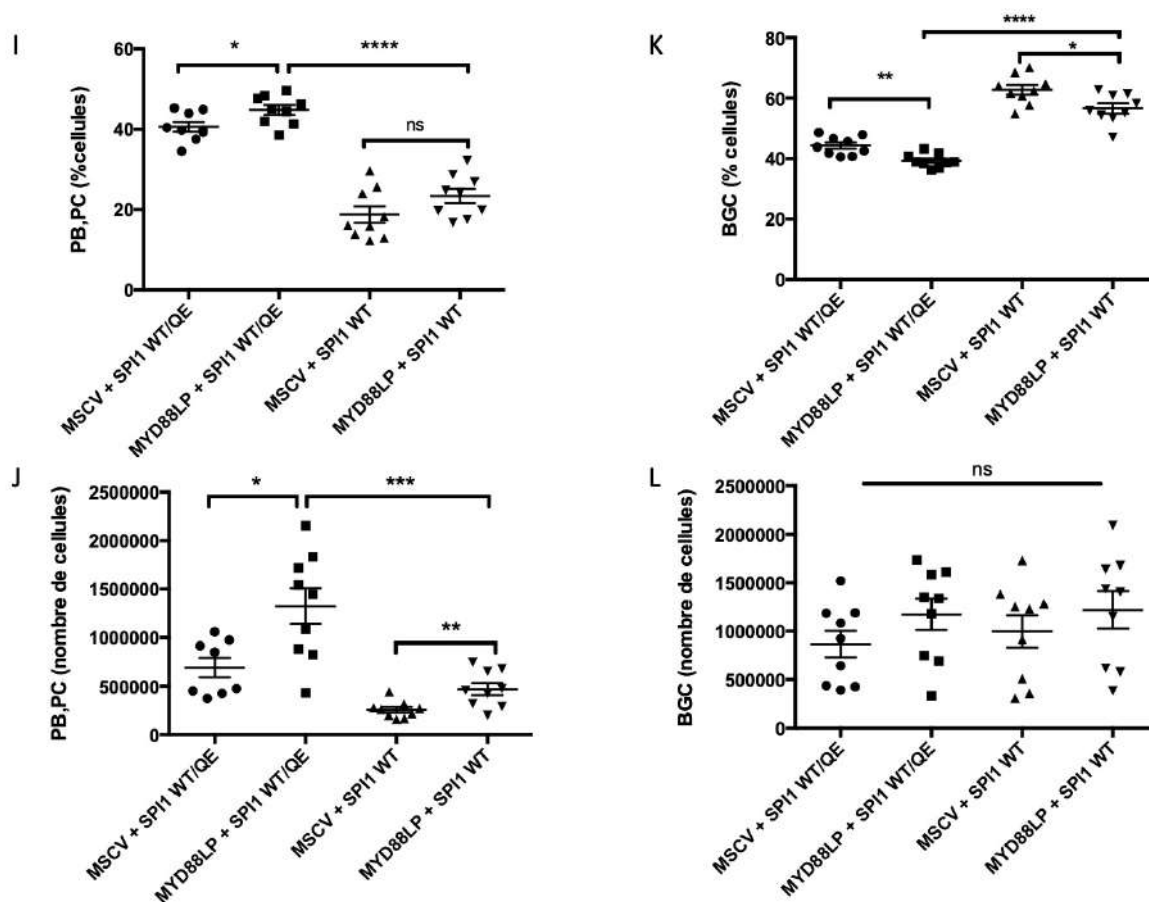


Figure 2. Stimulation des B naïfs spléniques par CD40/IL4/IL21 (suite)

I-L, Proportion (I et K) et valeur absolue (J et L) des plasmablastes (PB) et plasmocytes (PC), (CD19^{low} CD138+) et des BGC (CD19+GL7+Fas+) à J8.

Les points représentent 3 souris par génotype, analysées en triplicat lors de 2 expériences indépendantes. Toutes les p values ont été calculées à l'aide du t-test de Student non apparié, *p < 0,05, **p < 0,005, ***p < 0,0005 et ns : non significatif, dans toutes les expériences.

Nous avons testé une autre stimulation cellulaire, LPS/IL2/IL5 (**Figure 3**), qui a pour avantage d'engager la voie des TLR dont MYD88 est une protéine adaptatrice. Cette stimulation entraîne une expansion et une différenciation des cellules plus rapide que la stimulation CD40/IL4/IL21, dès les premières 24h de culture.

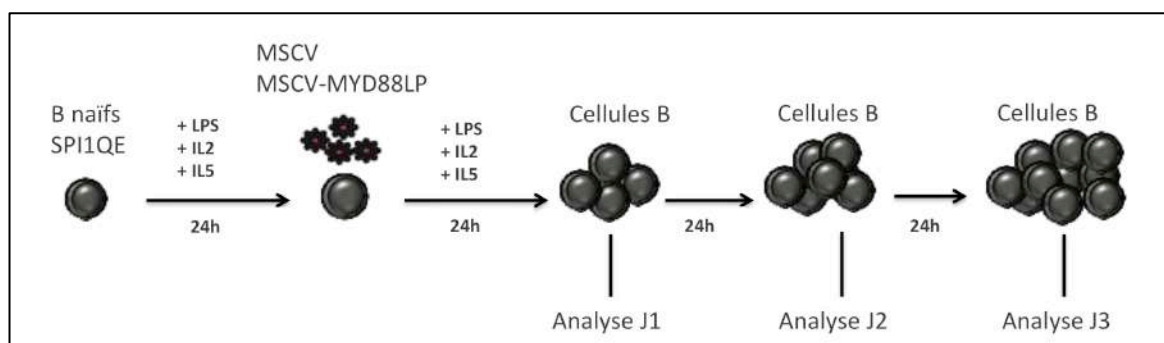


Figure 3. Représentation schématique du modèle *in vitro* LPS/IL2/IL5

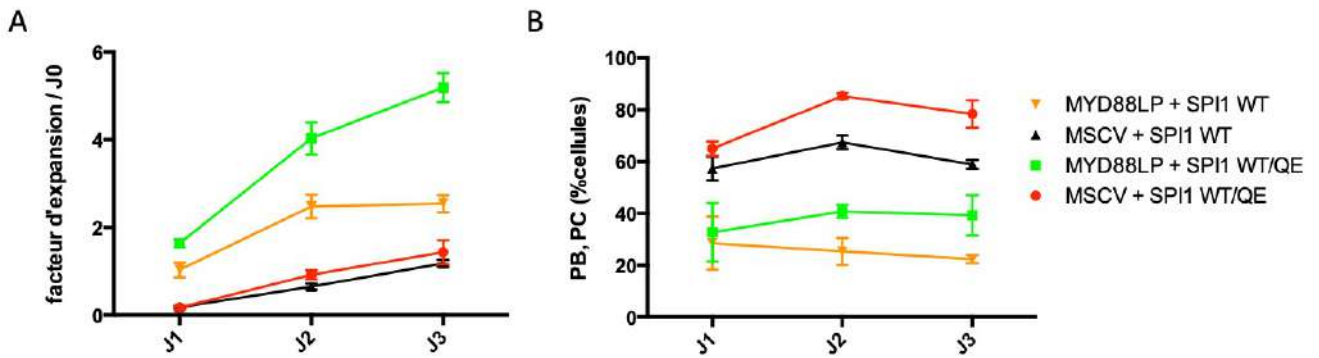


Figure 4. Stimulation des B naïfs spléniques par LPS/IL2/IL5

Après transduction par le vecteur apportant la mutation de MYD88 ou le vecteur contrôle, les cellules GFP+ ont été analysées à J1, J2 et J3 par cytométrie en flux. **A**, Cinétique de l'expansion des cellules GFP+ par rapport à J0. **B**, Cinétique de la proportion de PB/PC CD138+.

Ainsi, nous montrons une expansion des cellules double mutantes significativement supérieure à toutes les autres conditions (**Figure 4A**). La différenciation CD138+ est maximale pour les cellules SPI1QE-seul, et minimale pour les cellules MYD88-seul, à J1, J2 et J3 (**Figure 4B**). La transduction de MYD88 dans les cellules SPI1QE entraîne une diminution de la différenciation CD138+ par rapport aux cellules SPI1QE-seul; mais une augmentation de la différenciation par rapport aux cellules MYD88-seul (**Figure 4B-D**). En valeur absolue, étant donné l'avantage de prolifération important des double-mutants, les cellules CD138+ sont significativement plus nombreuses dans cette condition (**Figure 4D**).

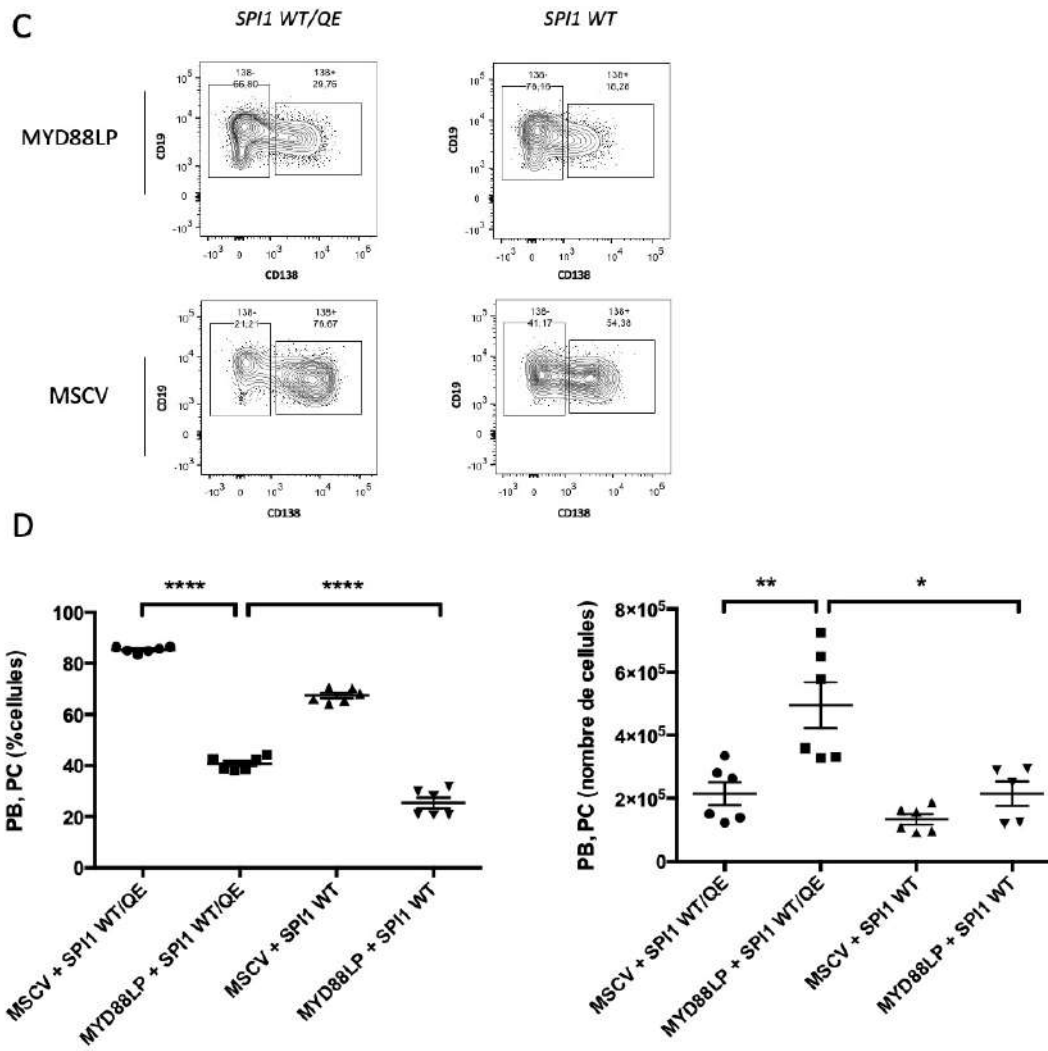


Figure 4. Stimulation des B naïfs spléniques par LPS/IL2/IL5 (suite)

C, Exemple de profil FACS à J2. **D**, Proportion (à gauche) et valeur absolue (à droite) de PB/PC (CD19^{low} CD138+). Les points représentent les triplicats à J2 de deux expériences indépendantes. Toutes les p values ont été calculées à l'aide du t-test de Student non apparié, *p < 0,05, **p < 0,005, ***p < 0,0005 et ns : non significatif, dans toutes les expériences.

Comme pour la stimulation CD40/IL4/IL21, la stimulation LPS/IL2/IL5 montre une coopération entre MYD88 et SPI1, visible sur l'expansion du nombre de cellules et sur l'engagement vers la différenciation plasmocytaire.

Expression du TLR4 sur les B activés et les plasmablastes/plasmocytes SPI1 QE

Nous avons ensuite cherché à documenter les bases de la coopération entre MYD88 et SPI1. Le gène *TLR4* est une cible connue de SPI1, et MYD88 est une de protéines adaptatrices de cette voie de signalisation (voir Figure 17, Introduction, page 70). Nous nous sommes donc interrogés sur la possibilité d'une régulation à la hausse de l'expression du TLR4 en présence de la mutation de SPI1. Par cytométrie en flux, nous avons étudié l'expression du TLR4 dans les cellules SPI1QE transduites avec le vecteur contrôle ou le vecteur apportant la mutation de MYD88, par rapport à des cellules sauvages.

Nous avons analysé la variation de l'expression du TLR4 en fonction du temps, sur les B activés ainsi que sur les PB/PC (**Figure 5A-C**). À l'état basal, le TLR4 est très faiblement exprimé par les cellules B naïves, son expression augmente après stimulation par LPS/IL2/IL5 (**Figure 5D-G**). Les transductions avec le vecteur contrôle montrent une expression de TLR4 en moyenne 20% à 40% supérieure dans les cellules SPI1QE par rapport aux WT dans les cellules CD138+. Dans les B activés, on note une augmentation de l'expression d'environ 20% à J1 et J2 (**Figure 5A, 5B, 5D et 5E**). A J3, les B activés WT ont une expression du TLR4 équivalente à celle des cellules SPI1QE (**Figure 5C**). Les transductions avec le vecteur exprimant la protéine MYD88 mutée ne permettent pas d'observer d'augmentation de l'expression du TLR4 dans les cellules SPI1QE (**Figure 5 F-G**).

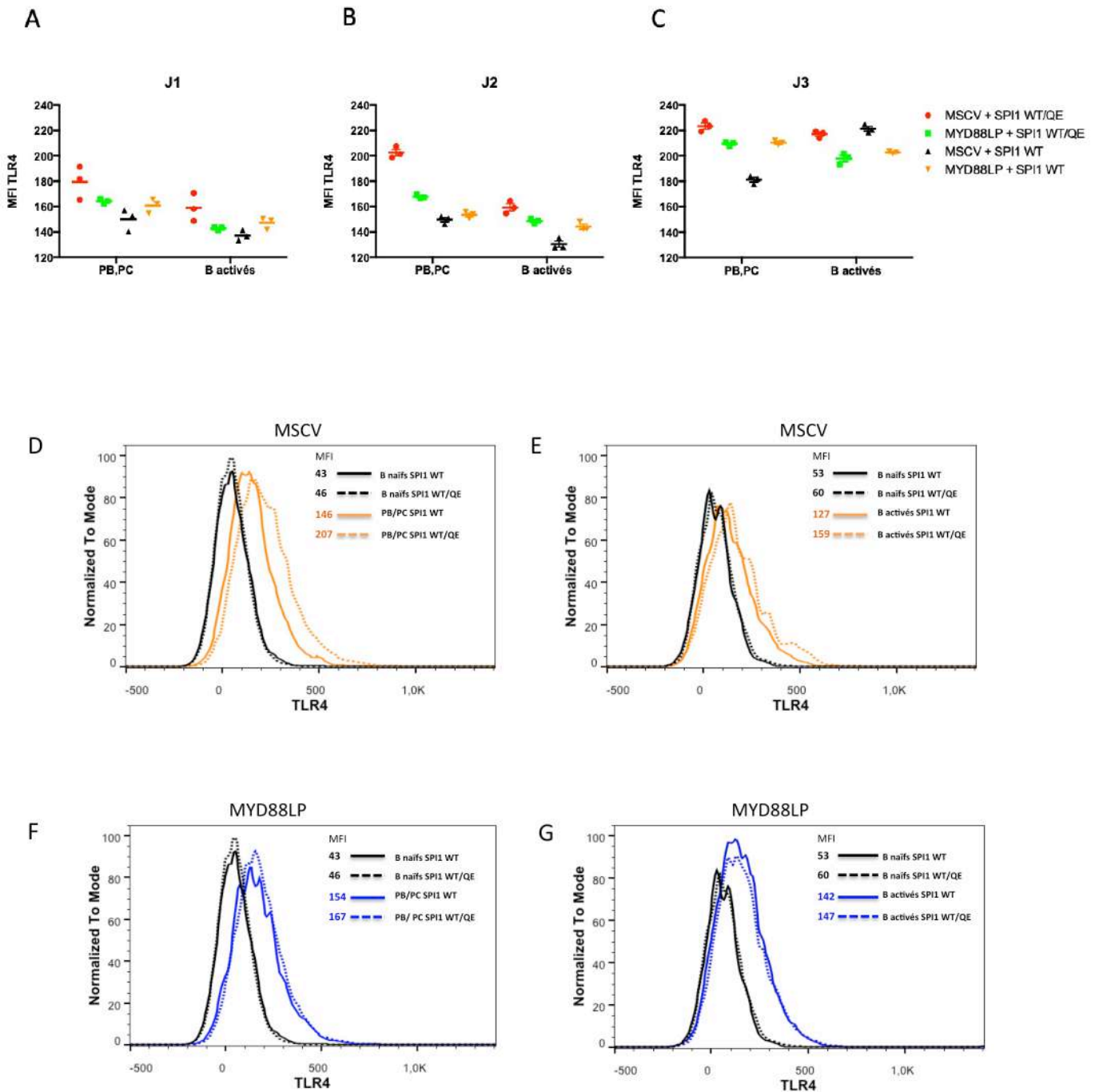
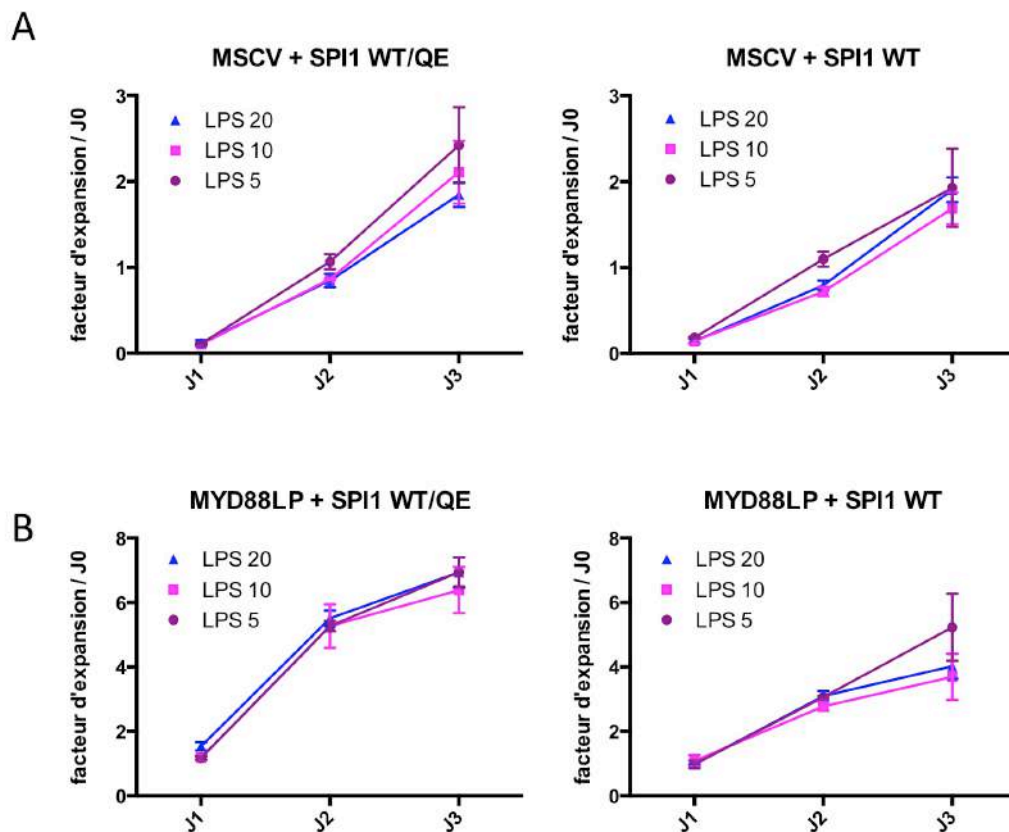


Figure 5. Expression du TLR4 sur les cellules B après stimulation par LPS/IL2/IL5

A-C. Variation de l'expression du TLR4 par cytométrie en flux sur les cellules CD138+ et sur les B activés à J1 (A) J2 (B) et J3 (C). **D-G.** Un exemple d'histogramme de cytométrie en flux à J2 est donné pour chaque vecteur et pour chaque type cellulaire.

Ayant mis en évidence une augmentation de l'expression du TLR4 dans les cellules avec la mutation de SPI1, nous avons voulu savoir si elles répondaient de manière plus importante que les cellules WT à une faible stimulation LPS. Une première expérience avec des doses 2 fois (10µg/ml) et 4 fois (5µg/ml) plus faibles qu'utilisée précédemment n'a pas montré de prolifération supérieure des cellules avec la mutation de SPI1 (**Figure 6A et 6B**). Nous avons reproduit cette expérience à des doses encore plus faibles, sans observer de différence, néanmoins les concentrations d'IL2 et IL5 ont été conservées à l'identique.

Figure 6. Effet de doses croissantes de LPS sur l'expansion cellulaire



A, Cinétique de l'expansion des cellules transduites avec le vecteur contrôle. **B**, Cinétique de l'expansion des cellules transduites avec le vecteur apportant la mutation de MYD88. La concentration de LPS dans le milieu de culture varie de 5µg/ml à 20µg/ml, les concentrations d'IL2 et d'IL5 sont gardées constantes. Toutes les p values calculées à l'aide du t-test de Student non apparié sont non significatives

L'augmentation de l'expression du TLR4 à la surface des cellules SPI1QE et leur sensibilité à la stimulation en réponse à de faibles doses de LPS est à vérifier dans de nouvelles expériences (voir Discussion p 165).

Etablissement d'une lignée humaine de MW SPI1-mutée par CRISPR-Cas9

Nous avons utilisé l'approche CRISPR/Cas9 dans la lignée cellulaire BCWM1, porteuse de la mutation *MYD88L265P*, pour introduire la mutation Q226E dans le locus endogène de *SPI1*. Le codon stop de *SPI1* est remplacé par l'ADNc de la GFP en fusion et en phase, afin d'obtenir l'expression physiologique de la protéine SPI1 mutée et fluorescente (**Figure 7**).

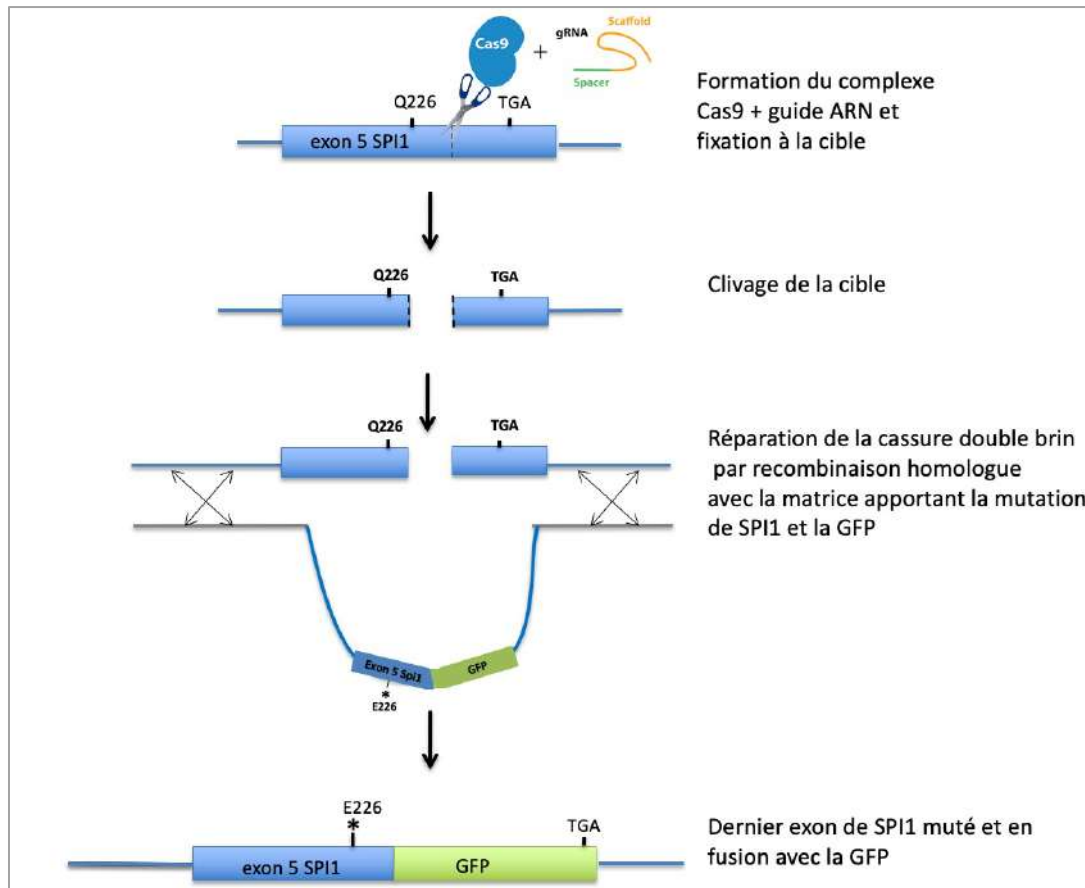


Figure 7. Stratégie utilisée pour la modification par CRISPR-Cas9 de la lignée BCWM1.

Le guide choisi induit un clivage par la Cas9 à proximité du site de la mutation. Une matrice de recombinaison homologue apporte la mutation et le cDNA de la GFP en fusion avec la fin de l'exon 5 de *SPI1*

Un guide ARN choisi à l'aide de l'outil CRISPOR (<http://crispor.tefor.net>) entraîne la coupure par la Cas9 à proximité immédiate du site de la mutation. Ce guide a ensuite été cloné dans le vecteur CRISPR-Cas9-Cherry (adapté du plasmide Addgene # 52961) permettant d'introduire à la fois la Cas9, le guide ARN et le traceur fluorescent cherry dans la cellule. Les séquences sont données ci-dessous.

1gRNA_ssmut_SPI_F	CACCCGCGCGCGCCATCTTCTGGT	Primer 1
2gRNA_ssmut_SPI_R	AAACACCAGAAGATGGCGCGCGCG	Primer 2

La matrice de recombinaison homologue a été construite par des approches de PCR, avec 4 fragments à assembler (**Figure 8**). Les fragments 1 et 4 constituent les bras d'homologie en 5' et 3', d'environ 800 pb chacun. Le fragment 2 correspond à l'exon 5 de SPI1, le primer sens apporte la mutation QE ainsi qu'une mutation silencieuse permettant de muter le PAM (Protospacer Adjacent Motif). Le fragment 3 correspond à la séquence de la GFP, en fusion avec SPI1 (linker de 3 AA entre la fin de SPI1 et le début de la GFP). Une fois les 4 fragments amplifiés séparément, des PCR supplémentaires ont permis d'assembler les fragments 2, 3 et 4. Du fait de la haute teneur en GC de certaines régions, de la présence de séquences répétées et de la probable formation de structures secondaires, nous n'avons pu assembler le fragment 1 aux fragments pré-assemblés 2, 3 et 4. Une stratégie In-Fusion® (Takara Bio) a dû être employée pour parvenir à finalement obtenir la séquence complète de la matrice de recombinaison homologue. Cette séquence a été clonée dans le vecteur PUC19 (Invitrogen).

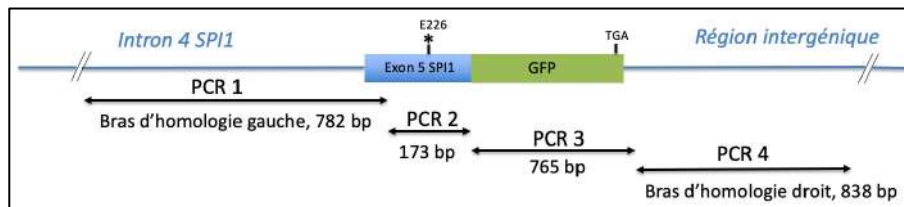


Figure 8 : Schéma de la construction de la matrice de recombinaison homologue

Les vecteurs contenant la Cas9 et le guide ARN d'une part, et la matrice de recombinaison homologue d'autre part ont été transfectés par électroporation dans la lignée BCWM1 (Amaxa Cell Line Nucleofector Kit V [VCA-1003]). Après 24h, les cellules ont été triées sur l'expression de la Cherry (permettant de sélectionner les cellules vivantes ayant reçu le vecteur contenant la Cas9 et le guide ARN). Celles-ci sont remises en culture pour 7 jours, date à laquelle les cellules ont été triées sur l'expression de la GFP (cellules ayant réparé la cassure avec la matrice de recombinaison homologue). L'expression de la GFP est faible, reflet de l'expression endogène de SPI1, et positive pour moins de 1% des cellules (**Figure 9**). Les cellules ont été triées une seconde fois sur l'expression de la GFP, quelques semaines plus tard, afin d'augmenter la pureté. Les cellules ont été gardées en bulk d'une part, et sous-clonées d'autre part.

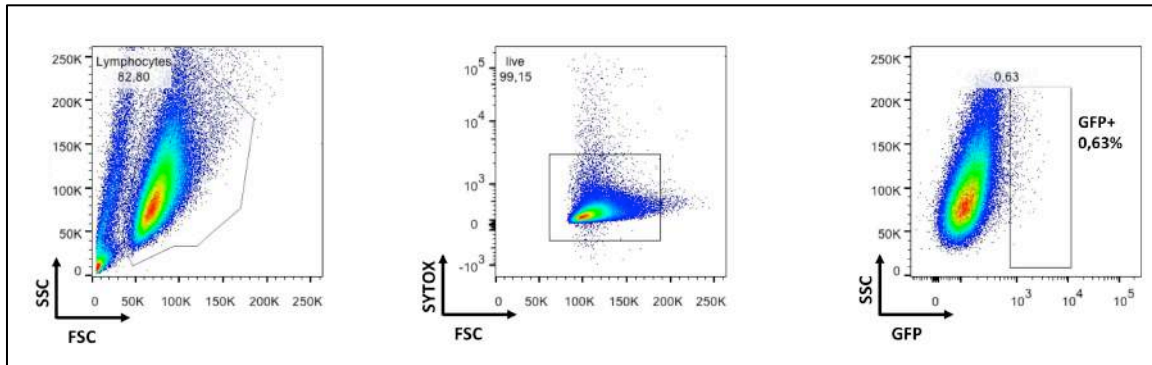


Figure 9. Tri cellulaire après transfection des cellules BCWM1

Profils de cytométrie en flux montrant le tri de la population GFP+, correspondant aux cellules BCWM1 ayant intégré la matrice de recombinaison homologue.

Pour valider l'efficacité du CRISPR, plusieurs approches ont été employées :

- 1) Validation par PCR de l'intégration de la matrice au bon locus, en utilisant un primer sur la GFP et un primer en dehors de la matrice (**Figure 10**)

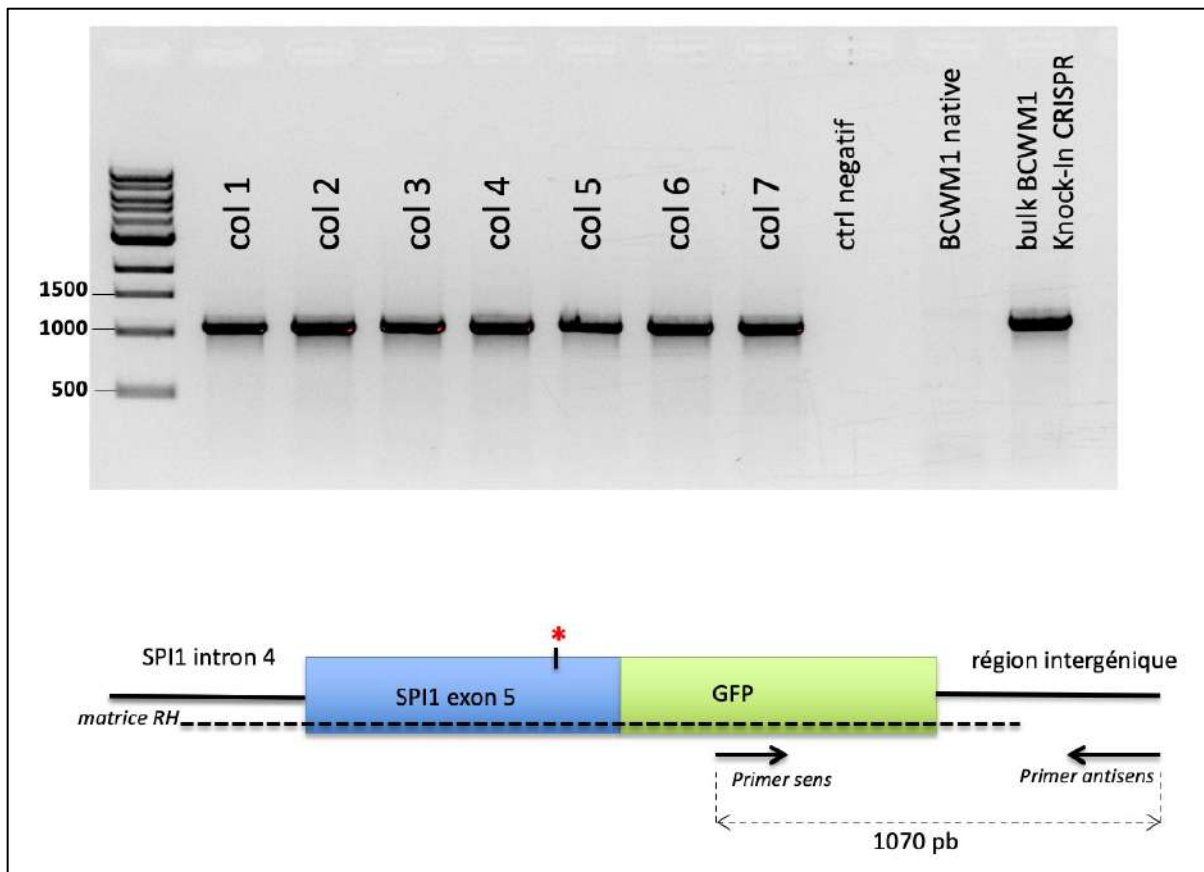


Figure 10. Vérification de l'intégration de la matrice par PCR.

Le primer sens se trouve dans la séquence de la GFP et le primer antisens dans la région intergénique en 3' de *SPI1*, en dehors de la matrice. Si la matrice a été intégrée, la taille attendue des amplicons est de 1070 paires de bases (pb). Analyse par migration sur gel à 1% d'agarose des produits de PCR obtenus à partir de cellules isolées (colonies 1 à 7), de la lignée parentale et la lignée Knock-In CRISPR.

2) Validation de la présence de la mutation par séquençage Sanger (**Figure 11**)

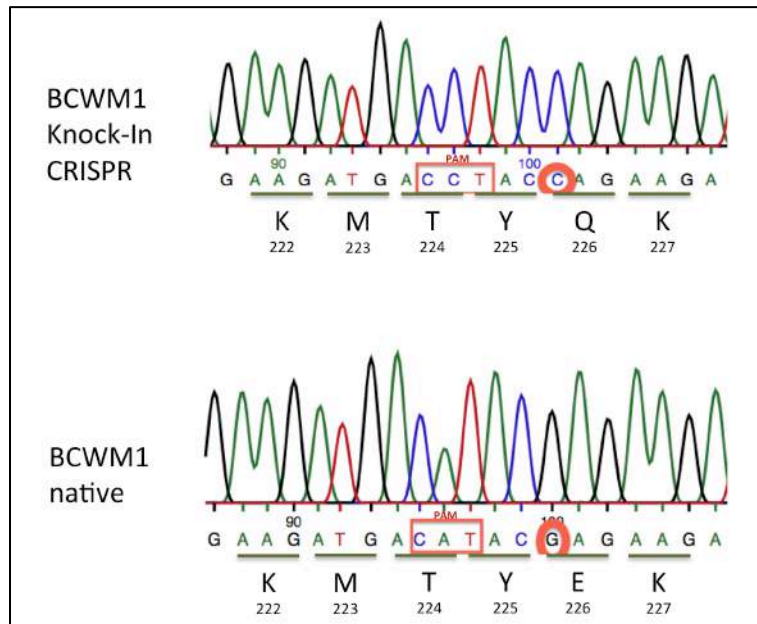


Figure 11. Vérification de la présence de la mutation SPI1 Q226E

Validation par séquençage de la présence de la mutation SPI1QE (entourée en rouge) dans la lignée BCWM1. Le PAM (encadré en rouge) a également été muté de façon silencieuse.

3) Expression de la GFP dans les cellules par cytométrie en flux (**Figure 12**)

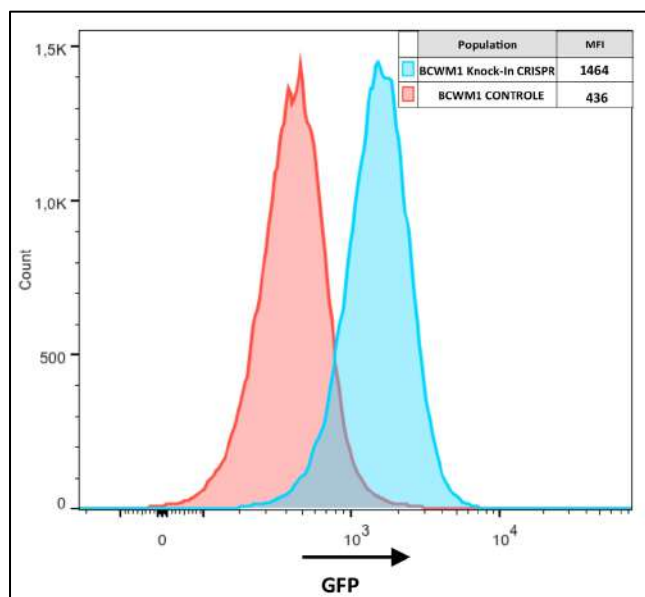


Figure 12. Détection de l'expression de SPI1-GFP endogène par cytométrie en flux

4) WB (*en cours*)

2.4 Discussion

Le travail développé dans l'article n°2 montre comment la mutation Q226E modifie la liaison à l'ADN de SPI1 sur ses sites habituels, dans les enhancers, au profit de site de liaison d'une autre classe de protéine ETS, dans les régions promotrices. Cela a pour conséquence d'activer des programmes transcriptionnels impliquant des voies de signalisation intracellulaires généralement favorisées par d'autres membres de la famille ETS. Nous avons mis évidence que cette mutation augmente la prolifération cellulaire des lignées de lymphome B (OCI-LY10, BCWM1 et MWCL1) et diminue la différenciation lymphoïde B terminale des cellules primaires humaines et murines.

Influence de la mutation SPI1QE sur la prolifération et la différenciation B terminale

Une des principales limites des modèles utilisés dans l'article n°2 est l'absence d'expression physiologique de SPI1QE. En effet, il a été montré que la surexpression de SPI1 dans la lignée myéloïde humaine K562 entraîne un arrêt de la croissance cellulaire (476), et sa surexpression dans les érythroleucémies murines induit le blocage de la différenciation (385). Concernant la lignée lymphoïde B, la surexpression de SPIB, proche parent de SPI1, entraîne un blocage de la différenciation B en plasmocytes chez l'homme (475). Nous avons constaté le même effet en surexprimant SPI1 dans des lymphocytes B humains (données non montrées). Cela explique probablement une partie des différences entre les données obtenues dans l'article n°2 et les résultats présentés dans cette section. En effet, la surexpression de SPI1QE dans les cellules primaires murines entraînait une diminution de la proportion de cellules CD138+, contrairement à ce que l'on observe avec le modèle murin knock-in pour la mutation, où l'augmentation de la proportion des cellules CD138+ après stimulation des B naïfs est significative, de façon reproductible. A noter également la différence entre les stimulations utilisées, anti-IgM et anti-CD40 uniquement avant l'étape de transduction, favorisant la survie mais pas la prolifération des cellules B, ou bien CD40/IL4/IL21 et LPS/IL2/IL5 pendant toute la durée de culture, stimulant leur prolifération et leur différenciation. Ces deux derniers types de stimulation entraînent toutefois une réponse cellulaire différente. La stimulation via le CD40 entraîne une réponse dose-dépendante : si la cellule reçoit un signal faible, les divisions seront plus lentes, la différenciation s'effectuera de manière précoce. En revanche un signal fort entraînera des divisions rapides, la différenciation sera différée pour favoriser l'expansion clonale. La stimulation via le LPS repose plutôt sur un mécanisme 'tout ou rien', si le seuil d'activation de la cellule est franchi, l'augmentation des doses n'entraîne pas plus de divisions ni plus de différenciation (188). En pratique, la stimulation LPS/IL2/IL5 entraîne une augmentation plus rapide du nombre de cellules et une différenciation plus importante en CD138+ que la stimulation CD40/IL4/IL21. Avec les cellules WT, on observe environ 20% de cellules CD138+ à J8 avec la stimulation CD40/IL4/IL21 tandis qu'on atteint 70% dès

J2 avec la stimulation LPS/IL2/IL5. Une autre différence notable entre ces deux modèles est la présence de feeders dans le modèle CD40/IL4/IL21, qui réduit de façon spectaculaire la mortalité cellulaire autrement observée dans les cultures liquides de cellules B (203). Néanmoins, dans les deux cas nous observons une augmentation de la prolifération dans la condition double mutante MYD88/SPI1 par rapport à toutes les autres conditions, ainsi qu'une augmentation de la proportion de cellules CD138+ par rapport aux cellules MYD88-seul.

Les gènes du cycle cellulaire font partie des cibles de SPI1. Cela a été montré dans les HSC, où SPI1 empêche leur division excessive, prévenant ainsi leur épuisement. Dans ces cellules, SPI1 se lie aux promoteurs et aux enhanceurs des inhibiteurs (*Gfi1*, *Cdkn1a*) et des activateurs (*Cdk1*, *E2f1* et *Cdc25a*) du cycle cellulaire (477). Les modèles avec knockdown de SPI1 montrent une augmentation de la fraction des HSC en division (phases S, G2 et M du cycle cellulaire), en lien avec l'augmentation de l'expression des activateurs du cycle cellulaire et la diminution de l'expression des inhibiteurs du cycle cellulaire, réversible en rétablissant des niveaux normaux de SPI1 (477). Les expériences de cycle cellulaire effectuées dans la lignée OCI-LY10 (article n°2) montrent que l'expression ectopique de SPI1 QE diminue le pourcentage de cellules en phase G1, au profit des phases G2-M et ce de façon supérieure à SPI1 WT. En stimulant par LPS/IL2/IL5 des B naïfs exprimant SPI1 QE de façon endogène, nous avons constaté qu'un plus grand nombre de cellules se trouvaient dans les divisions avancées en présence de la mutation SPI1 QE (données de CFSE non montrées), fournissant un mécanisme de l'augmentation de la différenciation en plasmablaste/plasmocyte.

La coopération observée entre MYD88 et SPI1 dans nos différents modèles reposerait donc sur la promotion de la croissance, assurée par la mutation activatrice de MYD88 auquel s'ajoute la capacité de différenciation plasmocytaire apporté par SPI1 QE.

Une limite à ce modèle est qu'il utilise la surexpression de MYD88 : un modèle plus physiologique pourra être apporté par le croisement des souris SPI1QE-KI avec des souris portant la mutation MYD88LP, permettant l'expression endogène des deux mutations.

Timing et coopération des mutations SPI1QE et MYD88LP

Une question clé pour comprendre la lymphomagénèse est de savoir quand et à quel niveau cellulaire les mutations se produisent. Pour tenter d'y répondre, nous avons effectué du tri cellulaire et le reséquençage ciblé d'une dizaine de patients atteints de MW. La mutation *MYD88* est détectable au stade pro-B (CD34+CD19+), sans pouvoir être identifiée avec certitude aux stades antérieurs. Elle est également trouvée dans les B matures non tumoraux mais pas dans les lymphocytes T (données non montrées). Cela suggère que la mutation apparaît à un stade antérieur au stade B mémoire (contrepartie physiologique supposée de la MW), pendant le développement B. La mutation de *SPI1*, quant à elle, n'a

pas été trouvée dans les lymphocytes T des patients, elle intervient donc probablement à stade ultérieur au progéniteur lymphoïde commun.

La mutation *MYD88* L265P n'est détectée que dans les hémopathies B, alors que *MYD88* joue un rôle dans la transduction du signal des TLR dans d'autres cellules hématopoïétiques (cellules dendritiques, macrophages). Une explication pourrait être que la mutation est sélectionnée lorsqu'elle intervient dans un progéniteur déjà engagé dans la lignée B car elle dépend de la signalisation des cellules B pour aboutir à la transformation tumorale. Cela est décrit par exemple pour la mutation *JAK2* V617F, trouvée exclusivement dans les hémopathies myéloïdes car elle nécessite une cellule co-exprimant un récepteur de cytokine homodimérique TPO, EPO ou GCSF (478). Selon un modèle similaire, l'activation constitutive de la voie *MYD88* ne pourrait entraîner un avantage sélectif que dans les cellules ayant un BCR. Dans ce sens, une signalisation active du BCR a été observée chez des patients atteints de MW et contribuerait à une signalisation pro-survie (479) et dans les DLBCL, les mutations activatrices de *MYD88* sont associées à des niveaux élevés de signalisation BCR chronique (154). Les mutations activatrices de *CD79A/B* sont pourtant présentes dans une minorité des ABC-DLBCL (20%) et des MW (10%). Parmi les autres mécanismes pouvant expliquer l'activation de la signalisation BCR dans ces pathologies, la synergie entre signalisation du BCR et des TLR est soutenue par plusieurs études, rapportant que *MYD88* muté active des composants du BCR (155,187). Par exemple, Munshi *et al.* (187) ont récemment montré dans des lignées de lymphomes B et des cellules de patients MW que la mutation L265P de *MYD88* augmente la phosphorylation de SYK (kinase en aval du BCR). Ceci est en accord avec une étude rapportant que dans les cellules B, la signalisation du TLR4 est dépendante du BCR et aboutit à l'activation de SYK (185). En outre, Phelan *et al.* (480) ont rapporté l'existence d'un supercomplexe *MYD88*-TLR9-BCR, co-localisant avec mTOR dans les endolysosomes et responsable de l'activation des voies NF- κ B et mTOR dans les DLBCL et la MW.

Plusieurs modèles murins de la mutation de *Myd88* ont été publiés. Knittel *et al.* (462) rapportent le développement de maladies lymphoprolifératives ou de DLBCL chez les souris, quel que soit le stade de différenciation B auquel la Cre s'exprime (CD19-Cre, AID-Cre ou CD21-Cre). Cela suggère que la mutation *Myd88* bloque la différenciation après le GC, et explique pourquoi elles sont enrichies dans les lymphomes post GC (ABC-DLBCL, WM, CLL mutées). En apportant la surexpression de *BCL2*, 100% des animaux développent un DLBCL. Sewastianik *et al.* (463) montrent également dans un modèle murin avec la mutation L265P le développement de syndrome lymphoprolifératifs B, pouvant se transformer en DLBCL après acquisition d'évènements secondaires. En revanche, Ouk *et al.* (481) rapportent la survenue de lymphomes lymphoplasmocytaires accompagnés de sécrétion d'IgM chez les souris dans lesquelles la mutation L252P (équivalent chez la souris

de L265P humaine) a été intégrée au niveau du locus Rosa26. Le profil d'expression du mutant MYD88 n'est pas le même entre les modèles, expliquant en partie les différences phénotypiques observées. Chez l'Homme, la mutation de *MYD88* est présente chez 60% des patients ayant un MGUS IgM, mais seulement 1,5 à 2 % des cas évoluent en MW chaque année (455,482), probablement en raison de mutations additionnelles coopérant avec celles de *MYD88*. Une étude de Wang *et al.* a notamment montré que la surexpression de MYD88L265P dans les cellules B primaires murines est rapidement contrebalancée par un rétrocontrôle négatif de la voie NF-κB, allant jusqu'à la mort cellulaire (461). Des mutations permettant de garder cette voie activée, comme celles inactivant *TNFAIP3* par exemple (inhibiteur de NF-κB), sont fréquemment associées à *MYD88L265P*. Dans le cas de SPI1QE, on peut également se demander si la coopération est soutenue par le maintien de l'activation de la voie NF-κB. En effet, Hu *et al.* (483) ont rapporté que l'expression de c-rel était régulée par les protéines ETS et ont décrit 3 sites de fixation de SPI1 et SPIB dans son promoteur.

Il serait intéressant de regarder l'activation des voies de signalisation NF-κB mais aussi d'autres voies impliquées dans la signalisation comme JAK-STAT3 (activée dans les DLBC-ABC MYD88-mutés) dans les cellules SPI1 QE-MYD88 par rapport aux cellules MYD88-seul issues du modèle de culture in vitro.

Influence de la mutation SPI1QE sur l'expression du TLR4

L'expression du TLR4 par les lymphocytes B chez l'homme est controversée. Elle serait absente des B naïfs mais inductible par la stimulation par son ligand le LPS (484,485). Plusieurs études rapportent son expression à la fois dans les cellules B de donneurs sains et de sujets atteints de lymphome du manteau (486) où il faisait partie des TLR avec l'expression protéique la plus forte, ou de myélome (487), où son transcrite est régulé à la hausse par rapport aux donneurs sains.

Deux études ont regardé l'expression des TLR dans la MW : Shrimpton *et al.* (488), qui montrent que l'ARNm du TLR4 est exprimé sur les B de donneurs sains et sur les B tumoraux, et Hunter *et al.* (454), qui montrent une régulation à la hausse de l'expression du TLR4 chez les patients *MYD88* muté-*CXCR4* WT par rapport aux double-mutants.

Le TLR4 est une cible connue de SPI1. Cela a été montré en premier dans les cellules myéloïdes (489), puis dans les lymphocytes B (371,490) où des sites de fixation de SPI1 dans le promoteur proximal de TLR4 ont été clairement identifiés (371). La stimulation du TLR4 par son ligand LPS permettant d'observer la prolifération ainsi que la différenciation CD138+ de cellules B *in vitro*, nous nous sommes donc interrogés sur la possibilité d'une régulation à la hausse de l'expression du TLR4 en présence de la mutation de SPI1 dans des cellules B activées.

Les cellules transduites avec le vecteur contrôle apportent des informations sur l'influence de la mutation de SPI1, indépendamment de MYD88. Il semble que les B SPI1QE activés augmentent leur expression du TLR4 plus rapidement que les cellules WT. En effet, il faut attendre J3 pour que les B activés WT aient un niveau d'expression équivalent aux cellules SPI1QE, tandis que les PB/PC WT ont toujours à J3 une expression moindre par rapport aux cellules mutées.

Avec la surexpression de MYD88LP, les différences d'expression ne sont plus visibles entre SPI1QE et SPIWT. Il est probable que l'activation massive de la voie des TLR qui en résulte rende la cellule indépendante de la fixation du ligand LPS, ce qui impliquerait qu'elle n'ait plus besoin de réguler à la hausse l'expression du TLR4. Si l'on confirme ces résultats sur de nouvelles expériences, plusieurs mécanismes pourront être évoqués, le premier serait une augmentation de la transcription du TLR4, dans le cas contraire il pourrait s'agir d'une diminution de l'internalisation du TLR4 (491).

Une approche indirecte pour confirmer l'augmentation d'expression du TLR4 sur les cellules SPI1 QE est de regarder si de faibles doses de LPS sont capables d'entraîner une réponse. Nous n'avons pas observé un tel effet, même avec le vecteur contrôle. Toutefois, il est nécessaire de reproduire ces expériences avec le LPS seul car il est probable que l'expansion des cellules indépendamment de la dose de LPS soit le reflet de l'action pro-proliférative de l'IL2 et IL5, conservées dans le milieu de culture. Il paraîtrait également intéressant de regarder l'effet de la stimulation d'autres TLR, le TLR9 en particulier, puisqu'il est impliqué dans le supercomplexe My-T-BCR évoqué plus haut. Cependant, cette stimulation n'avait pas été privilégiée dans un premier temps car elle ne donne pas de différenciation CD138+, en faisant donc un modèle moins pertinent pour notre étude de la différenciation B terminale.

Autres mutations des facteurs ETS ou de leurs sites de liaison

Les mutations des facteurs ETS sont relativement peu décrites dans les hémopathies B et le Whole Exome Sequencing (WES) n'a pas mis en évidence d'autres mutations récurrentes de protéines de cette famille dans notre cohorte de patients atteints de MW. Il peut exister, comme notre travail l'a montré, des mutations de FT capables de modifier leur affinité de liaison à l'ADN (ou à d'autres protéines) mais il peut aussi exister des mutations des régions régulatrices, changeant ou bien faisant apparaître des sites de liaison des FT, à l'instar des mutations au sein du promoteur du gène TERT, qui créent des sites de liaison de novo pour les facteurs de transcription ETS, entraînant la réactivation de la télomérase (335) ou encore des mutations dans des régions enhancers de PAX5, réduisant son expression, décrites dans la LLC (492) et le myélome (493). Dans cette perspective, nous sommes en train d'analyser les résultats de Whole genome Sequencing (WGS) sur une trentaine de

patients atteints de MW, les régions non codantes étant encore peu étudiées à ce jour, dans cette pathologie.

Etablissement d'une lignée humaine de MW SPI1-mutée par CRISPR-Cas9

Les expériences de ChIP-seq présentées dans l'article n°2 ont été réalisées sur une lignée humaine de DLBCL, OCI-LY10, dans laquelle une surexpression de SPI1 muté ou sauvage a été induite par transduction rétrovirale. Pour se replacer dans le contexte de la MW avec un niveau d'expression physiologique de la protéine SPI1 mutée, nous avons entrepris de modifier le génome de la lignée BCWM1 par CRISPR-Cas9. Cette lignée est dérivée d'un patient atteint de MW porteur de la mutation *MYD88L265P*. Nous l'avons choisie car ses caractéristiques de prolifération sont avantageuses pour toutes les étapes de culture cellulaire nécessaires, en particulier sa clonogénicité à l'échelle unicellulaire. Un élément de controverse vient du fait que cette lignée exprime la protéine membranaire de latence 1 (LMP1) de l'EBV, ce qui pousse certains auteurs à remettre en question la relation clonale entre la tumeur de départ et la lignée, au profit de l'immortalisation par l'EBV d'une cellule B « by-stander » (494,495). Néanmoins, la caractérisation phénotypique, la clonalité par l'étude des réarrangements VDJ, et la présence de la mutation somatique de *MYD88* montrent que cette lignée conserve les caractéristiques spécifiques de la maladie et en font donc un modèle pré-clinique acceptable.

Plusieurs stratégies ont été envisagées pour la modification de la lignée par CRISPR-Cas9. Intégrer la GFP dans la construction permet de faciliter les étapes de sélection des cellules, sachant que la réparation de la cassure double brin par recombinaison homologue, à l'aide de la matrice contenant la mutation SPI1QE, serait un évènement rare (496-498). La question principale était de savoir si la GFP devait être en fusion avec SPI1, ou bien séparée par une séquence IRES ou T2A. Étant donné les applications pour lesquelles nous souhaitons utiliser cette lignée, en particulier du ChIP-seq et de la co-IP, l'avantage certain de l'approche par fusion est que protéine sauvage et mutée peuvent facilement être distinguées par l'utilisation d'un anticorps anti-GFP. L'inconvénient est que l'ajout de la GFP modifie la taille de la protéine SPI1 et il n'est pas totalement exclu qu'elle puisse influencer sur sa liaison à l'ADN ou à d'autres partenaires. L'expérience que nous avons avec la fusion en N-ter de la GFP à SPI1 dans les cellules OCI-LY10 a montré qu'elle conservait sa localisation nucléaire, ses sites de fixation et son activité transcriptionnelle (article n°2). Il sera utile de vérifier qu'il en est de même pour la fusion de la GFP en C-ter, par exemple en comparant le ChIP-seq obtenu avec la lignée native et celui obtenu avec la lignée modifiée par CRISPR-Cas9. Une construction similaire, utilisant la Yellow Fluorescent Protein (YFP) au lieu de la GFP a été utilisée dans l'étude de Hoppe *et al.*, et ne modifiait ni l'expression ni la stabilité de SPI1 (499). En tout état de cause, l'impact éventuel de la

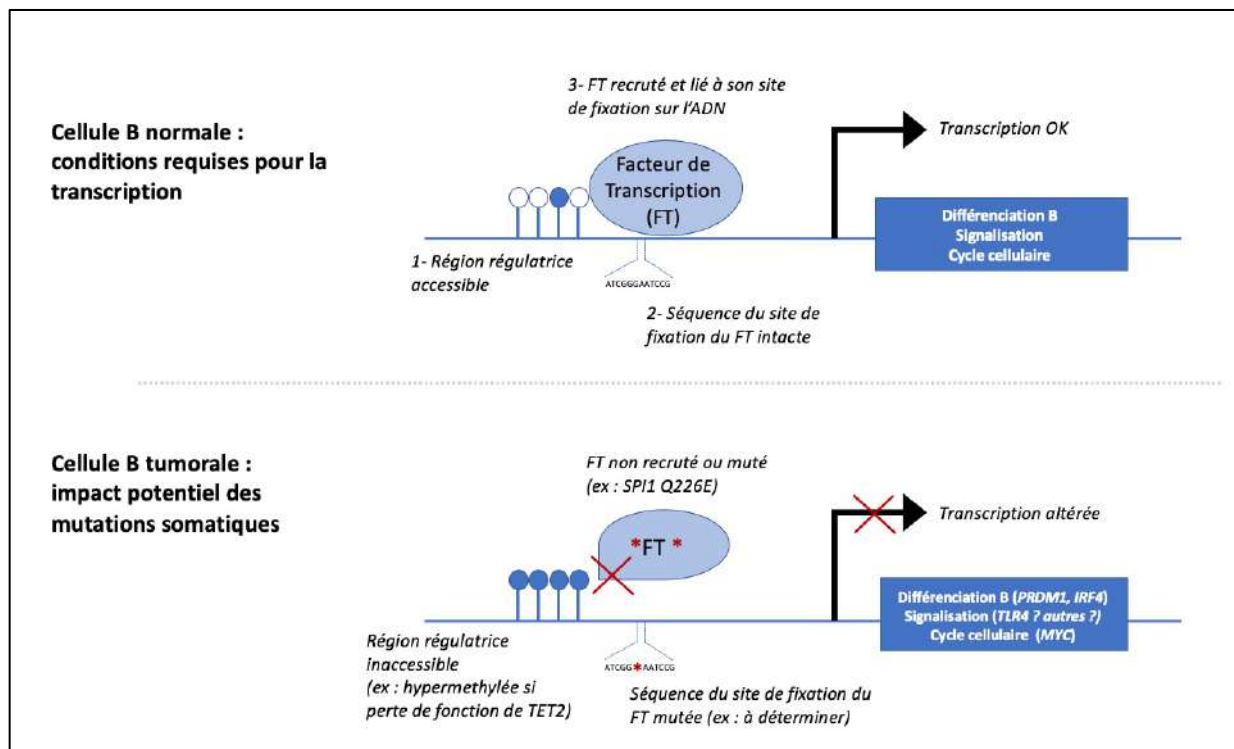
présence de la GFP sera contrôlé en comparant les profils des cellules SPI1QE-GFP et SPIWT-GFP.

Pour les prochaines étapes du travail, nous souhaitons exploiter l'avantage de la lignée cellulaire, donnant aisément accès à un grand nombre de cellules. Premièrement, nous souhaitons identifier par ChIP-seq les cibles de SPI1 muté dans ce contexte cellulaire, plus physiologique que la lignée de DLBCL OCI-LY10 surexprimant SPI1. Ces données pourront ensuite être reliées à la variation d'expression des gènes cibles évaluée par RNA-seq. Cela permettra de conforter les résultats de ChIP-seq obtenus sur la lignée OCI-LY10 montrant, entre autres, que SPI1QE adopte un comportement de protéines de classe I/IIa en termes de liaison à l'ADN mais aussi d'induction de programme transcriptionnel, le GSEA mettant en évidence un enrichissement des cibles de ETS1 et ELK1 dans la condition SPI1QE.

Les protéines ETS sont ubiquitaires, plusieurs membres de la famille étant exprimée dans le même tissu, ceci étant vrai pour de nombreux types cellulaires (314). Leur spécificité d'action est due, entre autres, à l'interaction avec des protéines partenaires (320,326). Nous voudrions donc savoir si la mutation, déplaçant la fixation de SPI1 sur d'autres sites ETS, pouvait entraîner l'occupation de ces sites mais sans interaction correcte avec les partenaires habituels. A titre d'exemple, ETS1 interagit physiquement avec BLIMP1, ce qui entraîne un blocage de l'activité de liaison à l'ADN de BLIMP1, une réduction de la capacité de BLIMP1 à réprimer les gènes cibles et donc à engager la différenciation B terminale. Il a été montré que d'autres protéines ETS, dont SPI1, étaient capables d'interagir avec BLIMP1 mais que seul ETS1 est capable d'inhiber sa liaison à l'ADN (84). Si dans les cellules B, SPI1 se trouve aux sites normalement liés par ETS1, mais sans pouvoir inhiber la liaison de BLIMP1 à l'ADN, l'absence de répression de l'activité transcriptionnelle de BLIMP1 entraînerait une différenciation plasmocytaire accrue, ce que l'on observe dans nos modèles. Ceci est en accord avec des études antérieures ayant rapporté que les souris dépourvues d'*Ets-1* présentent un nombre accru de plasmocytes et des titres sériques d'IgM plus élevés (81,82).

SPI1QE pourrait donc acquérir de nouvelles fonctions, via sa liaison nouvelle à des sites inhabituels, mais peut également subir une perte de fonction sur les sites habituels auxquels il se lie moins. Par exemple, SPI1 interagit avec IRF4 pour réguler la différenciation B terminale (68,69,376) et il est possible que la mutation QE entraîne une diminution de l'interaction entre les deux protéines au niveau des sites composites, due au *déplacement* de SPI1 vers d'autres sites. Cela pourrait être étudié grâce à des expériences de Co-IP, dans le cas où les partenaires sont ciblés, ou de façon plus globale par une approche couplée à de la spectrométrie de masse.

CONCLUSION et PERSPECTIVES



Mon travail de thèse permet d'illustrer certains aspects essentiels de la régulation de la différenciation B terminale :

- 1) L'accessibilité de la chromatine aux sites de fixation de FT clés. Elle dépend de plusieurs facteurs, dont la méthylation des cytosines et les modifications post-traductionnelles des histones.

Nous rapportons dans la partie 1 de ce travail qu'une conséquence notable de la perte de fonction de TET2 dans les cellules BGC est la perte focale d'hydroxyméthylation ainsi qu'une diminution de l'acétylation de la lysine K27 de l'histone 3 (H3K27Ac) au niveau des enhancers, corrélées à la répression transcriptionnelle des gènes qui interviennent dans la sortie des GC, comme *PRDM1*.

- 2) La capacité de liaison du FT à l'ADN. Certaines mutations des FT peuvent modifier l'affinité de liaison pour une séquence donnée.

Nous rapportons dans la partie 2 de ce travail que la mutation Q226E de *SPI1* modifie les caractéristiques de liaison à l'ADN de la protéine mutante, passant des séquences classiques reconnues par *SPI1* à des séquences reconnues par les protéines ETS de classes I/IIa, et d'une liaison à des régions enhancers à une liaison à des régions promotrices. La

liaison accrue du mutant de SPI1 aux régions promotrices active des programmes transcriptionnels impliquant des voies de signalisation intracellulaire généralement favorisées par d'autres membres de la famille ETS.

- 3) La reconnaissance du motif de fixation à l'ADN. Si la séquence normalement reconnue par le FT est mutée, la liaison peut s'en retrouver modifiée.

Cet aspect n'est pas développé dans le travail présenté ici mais fait l'objet d'un projet en cours, où une trentaine de cas de MW ont été séquencés par WGS. En effet, les régions non-codantes détiennent très vraisemblablement un certain nombre de réponses, les mutations dans les régions régulatrices pouvant faire apparaître ou disparaître des sites de liaison pour les FT. Ces données de WGS, associées aux données de RNA-seq et de méthylation pourront donner une image globale et complète des altérations (épi)génétiques dans la MW et être appliquées à d'autres contextes tumoraux.

REFERENCES

1. Orkin SH, Zon LI. Hematopoiesis: an evolving paradigm for stem cell biology. *Cell*. 2008 Feb 22;132(4):631–44.
2. Laurenti E, Göttgens B. From haematopoietic stem cells to complex differentiation landscapes. *Nature*. 2018 24;553(7689):418–26.
3. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The Human Transcription Factors. *Cell*. 2018 08;172(4):650–65.
4. Crane GM, Jeffery E, Morrison SJ. Adult haematopoietic stem cell niches. *Nat Rev Immunol*. 2017 Sep;17(9):573–90.
5. Clark MR, Mandal M, Ochiai K, Singh H. Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nat Rev Immunol*. 2014 Feb;14(2):69–80.
6. Wardemann H, Yurasov S, Schaefer A, Young JW, Meffre E, Nussenzweig MC. Predominant autoantibody production by early human B cell precursors. *Science*. 2003 Sep 5;301(5638):1374–7.
7. Melamed D, Nemazee D. Self-antigen does not accelerate immature B cell apoptosis, but stimulates receptor editing as a consequence of developmental arrest. *Proc Natl Acad Sci U S A*. 1997 Aug 19;94(17):9267–72.
8. Ottens K, Hinman RM, Barrios E, Skaug B, Davis LS, Li Q-Z, et al. Foxo3 Promotes Apoptosis of B Cell Receptor-Stimulated Immature B Cells, Thus Limiting the Window for Receptor Editing. *J Immunol Baltim Md 1950*. 2018 01;201(3):940–9.
9. Vilen BJ, Famiglietti SJ, Carbone AM, Kay BK, Cambier JC. B cell antigen receptor desensitization: disruption of receptor coupling to tyrosine kinase activation. *J Immunol Baltim Md 1950*. 1997 Jul 1;159(1):231–43.
10. Vilen BJ, Burke KM, Sleater M, Cambier JC. Transmodulation of BCR signaling by transduction-incompetent antigen receptors: implications for impaired signaling in anergic B cells. *J Immunol Baltim Md 1950*. 2002 May 1;168(9):4344–51.
11. Enders A, Bouillet P, Puthalakath H, Xu Y, Tarlinton DM, Strasser A. Loss of the proapoptotic BH3-only Bcl-2 family member Bim inhibits BCR stimulation-induced apoptosis and deletion of autoreactive B cells. *J Exp Med*. 2003 Oct 6;198(7):1119–26.
12. Norvell A, Mandik L, Monroe JG. Engagement of the antigen-receptor on immature murine B lymphocytes results in death by apoptosis. *J Immunol Baltim Md 1950*. 1995 May 1;154(9):4404–13.
13. Egawa T, Kawabata K, Kawamoto H, Amada K, Okamoto R, Fujii N, et al. The earliest stages of B cell development require a chemokine stromal cell-derived factor/pre-B cell growth-stimulating factor. *Immunity*. 2001 Aug;15(2):323–34.
14. Schwickert TA, Tagoh H, Gültekin S, Dakic A, Axelsson E, Minnich M, et al. Stage-specific control of early B cell development by the transcription factor Ikaros. *Nat Immunol*. 2014 Mar;15(3):283–93.
15. DeKoter RP, Schweitzer BL, Kamath MB, Jones D, Tagoh H, Bonifer C, et al. Regulation of the interleukin-7 receptor alpha promoter by the Ets transcription factors PU.1 and GA-binding protein in developing B cells. *J Biol Chem*. 2007 May 11;282(19):14194–204.
16. Heizmann B, Kastner P, Chan S. Ikaros is absolutely required for pre-B cell differentiation by attenuating IL-7 signals. *J Exp Med*. 2013 Dec 16;210(13):2823–32.
17. Kikuchi K, Lai AY, Hsu C-L, Kondo M. IL-7 receptor signaling is necessary for stage transition in adult B cell development through up-regulation of EBF. *J Exp Med*. 2005 Apr 18;201(8):1197–203.
18. O’Riordan M, Grosschedl R. Coordinate regulation of B cell differentiation by the

- transcription factors EBF and E2A. *Immunity*. 1999 Jul;11(1):21–31.
19. Fuxa M, Skok J, Souabni A, Salvagiotto G, Roldan E, Busslinger M. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev*. 2004 Feb 15;18(4):411–22.
 20. Bertolino E, Reddy K, Medina KL, Parganas E, Ihle J, Singh H. Regulation of interleukin 7-dependent immunoglobulin heavy-chain variable gene rearrangements by transcription factor STAT5. *Nat Immunol*. 2005 Aug;6(8):836–43.
 21. Lazorchak AS, Schlissel MS, Zhuang Y. E2A and IRF-4/Pip promote chromatin modification and transcription of the immunoglobulin kappa locus in pre-B cells. *Mol Cell Biol*. 2006 Feb;26(3):810–21.
 22. Nutt SL, Hodgkin PD, Tarlinton DM, Corcoran LM. The generation of antibody-secreting plasma cells. *Nat Rev Immunol*. 2015 Mar;15(3):160–71.
 23. Cobaleda C, Schebesta A, Delogu A, Busslinger M. Pax5: the guardian of B cell identity and function. *Nat Immunol*. 2007 May;8(5):463–70.
 24. Chung JB, Silverman M, Monroe JG. Transitional B cells: step by step towards immune competence. *Trends Immunol*. 2003 Jun;24(6):343–9.
 25. Hardy RR, Hayakawa K. B cell development pathways. *Annu Rev Immunol*. 2001;19:595–621.
 26. Cariappa A, Tang M, Parng C, Nebelitskiy E, Carroll M, Georgopoulos K, et al. The follicular versus marginal zone B lymphocyte cell fate decision is regulated by Aiolos, Btk, and CD21. *Immunity*. 2001 May;14(5):603–15.
 27. Ghosn EEB, Sadate-Ngatchou P, Yang Y, Herzenberg LA, Herzenberg LA. Distinct progenitors for B-1 and B-2 cells are present in adult mouse spleen. *Proc Natl Acad Sci U S A*. 2011 Feb 15;108(7):2879–84.
 28. Okada T, Moriyama S, Kitano M. Differentiation of germinal center B cells and follicular helper T cells as viewed by tracking Bcl6 expression dynamics. *Immunol Rev*. 2012 May;247(1):120–32.
 29. Choi YS, Yang JA, Crotty S. Dynamic regulation of Bcl6 in follicular helper CD4 T (Tfh) cells. *Curr Opin Immunol*. 2013 Jun;25(3):366–72.
 30. Kroenke MA, Eto D, Locci M, Cho M, Davidson T, Haddad EK, et al. Bcl6 and Maf cooperate to instruct human follicular helper CD4 T cell differentiation. *J Immunol Baltim Md 1950*. 2012 Apr 15;188(8):3734–44.
 31. Qi H, Chen X, Chu C, Lu P, Xu H, Yan J. Follicular T-helper cells: controlled localization and cellular interactions. *Immunol Cell Biol*. 2014 Jan;92(1):28–33.
 32. Goodnow CC, Vinuesa CG, Randall KL, Mackay F, Brink R. Control systems and decision making for antibody production. *Nat Immunol*. 2010 Aug;11(8):681–8.
 33. Shah HB, Smith K, Wren JD, Webb CF, Ballard JD, Bourn RL, et al. Insights From Analysis of Human Antigen-Specific Memory B Cell Repertoires. *Front Immunol*. 2019 Jan 15;9:3064.
 34. Shinnakasu R, Inoue T, Kometani K, Moriyama S, Adachi Y, Nakayama M, et al. Regulated selection of germinal-center cells into the memory B cell compartment. *Nat Immunol*. 2016;17(7):861–9.
 35. Camacho SA, Kosco-Vilbois MH, Berek C. The dynamic structure of the germinal center. *Immunol Today*. 1998 Nov;19(11):511–4.
 36. Mesin L, Ersching J, Victora GD. Germinal Center B Cell Dynamics. *Immunity*. 2016 Sep;45(3):471–82.
 37. Holmes AB, Corinaldesi C, Shen Q, Kumar R, Compagno N, Wang Z, et al. Single-cell analysis of germinal-center B cells informs on lymphoma cell of origin and outcome. *J Exp Med*. 2020 Oct 5;217(10):e20200483.
 38. Milpied P, Cervera-Marzal I, Mollichella M-L, Tesson B, Brisou G, Traverse-Glehen

- A, et al. Human germinal center transcriptional programs are de-synchronized in B cell lymphoma. *Nat Immunol*. 2018 Sep;19(9):1013–24.
39. Hauser AE, Junt T, Mempel TR, Sneddon MW, Kleinstein SH, Henrickson SE, et al. Definition of germinal-center B cell migration in vivo reveals predominant intrazonal circulation patterns. *Immunity*. 2007 May;26(5):655–67.
 40. Schwickert TA, Lindquist RL, Shakhar G, Livshits G, Skokos D, Kosco-Vilbois MH, et al. In vivo imaging of germinal centres reveals a dynamic open structure. *Nature*. 2007 Mar 1;446(7131):83–7.
 41. Hauser AE, Shlomchik MJ, Haberman AM. In vivo imaging studies shed light on germinal-centre development. *Nat Rev Immunol*. 2007;7(7):499–504.
 42. Allen CDC, Okada T, Tang HL, Cyster JG. Imaging of germinal center selection events during affinity maturation. *Science*. 2007 Jan 26;315(5811):528–31.
 43. Allen CDC, Ansel KM, Low C, Lesley R, Tamamura H, Fujii N, et al. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nat Immunol*. 2004 Sep;5(9):943–52.
 44. Tas JMJ, Mesin L, Pasqual G, Targ S, Jacobsen JT, Mano YM, et al. Visualizing antibody affinity maturation in germinal centers. *Science*. 2016 04;351(6277):1048–54.
 45. Klein U, Tu Y, Stolovitzky GA, Keller JL, Haddad J, Miljkovic V, et al. Transcriptional analysis of the B cell germinal center reaction. *Proc Natl Acad Sci U S A*. 2003 Mar 4;100(5):2639–44.
 46. Shaffer AL, Rosenwald A, Hurt EM, Giltneane JM, Lam LT, Pickeral OK, et al. Signatures of the Immune Response. *Immunity*. 2001 Sep;15(3):375–85.
 47. Phan RT, Dalla-Favera R. The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. *Nature*. 2004 Dec 2;432(7017):635–9.
 48. Hu BT, Lee SC, Marin E, Ryan DH, Insel RA. Telomerase is up-regulated in human germinal center B cells in vivo and can be re-expressed in memory B cells activated in vitro. *J Immunol Baltim Md 1950*. 1997 Aug 1;159(3):1068–71.
 49. Weller S, Faili A, Garcia C, Braun MC, Le Deist F, de Saint Basile G, et al. CD40-CD40L independent Ig gene hypermutation suggests a second B cell diversification pathway in humans. *Proc Natl Acad Sci*. 2001 Jan 30;98(3):1166–70.
 50. William J, Euler C, Christensen S, Shlomchik MJ. Evolution of autoantibody responses via somatic hypermutation outside of germinal centers. *Science*. 2002 Sep 20;297(5589):2066–70.
 51. Papavasiliou FN, Schatz DG. Cell-cycle-regulated DNA double-stranded breaks in somatic hypermutation of immunoglobulin genes. *Nature*. 2000 Nov 9;408(6809):216–21.
 52. Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. *Cell*. 2000 Sep 1;102(5):553–63.
 53. Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, et al. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). *Cell*. 2000 Sep 1;102(5):565–75.
 54. Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem*. 2007;76:1–22.
 55. Kuppers R. B cells under influence: transformation of B cells by Epstein-Barr virus. *Nat Rev Immunol*. 2003 Oct;3(10):801–12.
 56. Pasqualucci L, Guglielmino R, Houldsworth J, Mohr J, Aoufouchi S, Polakiewicz R, et al. Expression of the AID protein in normal and neoplastic B cells. *Blood*. 2004 Nov 15;104(10):3318–25.
 57. Toellner KM, Gulbranson-Judge A, Taylor DR, Sze DM, MacLennan IC. Immunoglobulin switch transcript production in vivo related to the site and time of antigen-

- specific B cell activation. *J Exp Med*. 1996 May 1;183(5):2303–12.
58. Roco JA, Mesin L, Binder SC, Nefzger C, Gonzalez-Figueroa P, Canete PF, et al. Class-Switch Recombination Occurs Infrequently in Germinal Centers. *Immunity*. 2019 Aug;51(2):337–350.e7.
59. Dent AL, Shaffer AL, Yu X, Allman D, Staudt LM. Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science*. 1997 Apr 25;276(5312):589–92.
60. Ye BH, Cattoretti G, Shen Q, Zhang J, Hawe N, de Waard R, et al. The BCL-6 proto-oncogene controls germinal-centre formation and Th2-type inflammation. *Nat Genet*. 1997 Jun;16(2):161–70.
61. Basso K, Dalla-Favera R. Roles of BCL6 in normal and transformed germinal center B cells. *Immunol Rev*. 2012 May;247(1):172–83.
62. Huang C, Geng H, Boss I, Wang L, Melnick A. Cooperative transcriptional repression by BCL6 and BACH2 in germinal center B-cell differentiation. *Blood*. 2014 Feb 13;123(7):1012–20.
63. Igarashi K, Ochiai K, Itoh-Nakadai A, Muto A. Orchestration of plasma cell differentiation by Bach2 and its gene regulatory network. *Immunol Rev*. 2014 Sep;261(1):116–25.
64. Hipp N, Symington H, Pastoret C, Caron G, Monvoisin C, Tarte K, et al. IL-2 imprints human naive B cell fate towards plasma cell through ERK/ELK1-mediated BACH2 repression. *Nat Commun*. 2017 Dec;8(1):1443.
65. Shapiro-Shelef M, Lin K-I, McHeyzer-Williams LJ, Liao J, McHeyzer-Williams MG, Calame K. Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. *Immunity*. 2003 Oct;19(4):607–20.
66. Reimold AM, Iwakoshi NN, Manis J, Vallabhajosyula P, Szomolanyi-Tsuda E, Gravalles EM, et al. Plasma cell differentiation requires the transcription factor XBP-1. *Nature*. 2001 Jul 19;412(6844):300–7.
67. Shaffer AL, Shapiro-Shelef M, Iwakoshi NN, Lee A-H, Qian S-B, Zhao H, et al. XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity*. 2004 Jul;21(1):81–93.
68. Klein U, Casola S, Cattoretti G, Shen Q, Lia M, Mo T, et al. Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nat Immunol*. 2006 Jul;7(7):773–82.
69. Sciammas R, Shaffer AL, Schatz JH, Zhao H, Staudt LM, Singh H. Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. *Immunity*. 2006 Aug;25(2):225–36.
70. Delogu A, Schebesta A, Sun Q, Aschenbrenner K, Perlot T, Busslinger M. Gene repression by Pax5 in B cells is essential for blood cell homeostasis and is reversed in plasma cells. *Immunity*. 2006 Mar;24(3):269–81.
71. Nera K-P, Kohonen P, Narvi E, Peippo A, Mustonen L, Terho P, et al. Loss of Pax5 promotes plasma cell differentiation. *Immunity*. 2006 Mar;24(3):283–93.
72. Kallies A, Hasbold J, Fairfax K, Pridans C, Emslie D, McKenzie BS, et al. Initiation of plasma-cell differentiation is independent of the transcription factor Blimp-1. *Immunity*. 2007 May;26(5):555–66.
73. De Silva NS, Klein U. Dynamics of B cells in germinal centres. *Nat Rev Immunol*. 2015 Mar;15(3):137–48.
74. Dominguez-Sola D, Victora GD, Ying CY, Phan RT, Saito M, Nussenzweig MC, et al. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. *Nat Immunol*. 2012 Nov;13(11):1083–91.
75. Calado DP, Sasaki Y, Godinho SA, Pellerin A, Köchert K, Sleckman BP, et al. The cell-

- cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat Immunol.* 2012 Nov;13(11):1092–100.
76. Lin KI, Lin Y, Calame K. Repression of c-myc is necessary but not sufficient for terminal differentiation of B lymphocytes in vitro. *Mol Cell Biol.* 2000 Dec;20(23):8684–95.
77. Nie Z, Hu G, Wei G, Cui K, Yamane A, Resch W, et al. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell.* 2012 Sep 28;151(1):68–79.
78. Carotta S, Willis SN, Hasbold J, Inouye M, Pang SHM, Emslie D, et al. The transcription factors IRF8 and PU.1 negatively regulate plasma cell differentiation. *J Exp Med.* 2014 Oct 20;211(11):2169–81.
79. Feng J, Wang H, Shin D-M, Masiuk M, Qi C-F, Morse HC. IFN regulatory factor 8 restricts the size of the marginal zone and follicular B cell pools. *J Immunol Baltim Md 1950.* 2011 Feb 1;186(3):1458–66.
80. Ochiai K, Maienschein-Cline M, Simonetti G, Chen J, Rosenthal R, Brink R, et al. Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity.* 2013 May 23;38(5):918–29.
81. Bories JC, Willerford DM, Grévin D, Davidson L, Camus A, Martin P, et al. Increased T-cell apoptosis and terminal B-cell differentiation induced by inactivation of the Ets-1 proto-oncogene. *Nature.* 1995 Oct 19;377(6550):635–8.
82. Barton K, Muthusamy N, Fischer C, Ting CN, Walunas TL, Lanier LL, et al. The Ets-1 transcription factor is required for the development of natural killer cells in mice. *Immunity.* 1998 Oct;9(4):555–63.
83. Wang D, John SA, Clements JL, Percy DH, Barton KP, Garrett-Sinha LA. Ets-1 deficiency leads to altered B cell differentiation, hyperresponsiveness to TLR9 and autoimmune disease. *Int Immunol.* 2005 Sep;17(9):1179–91.
84. John SA, Clements JL, Russell LM, Garrett-Sinha LA. Ets-1 regulates plasma cell differentiation by interfering with the activity of the transcription factor Blimp-1. *J Biol Chem.* 2008 Jan 11;283(2):951–62.
85. Ochiai K, Muto A, Tanaka H, Takahashi S, Igarashi K. Regulation of the plasma cell transcription factor Blimp-1 gene by Bach2 and Bcl6. *Int Immunol.* 2008 Mar;20(3):453–60.
86. Muto A, Ochiai K, Kimura Y, Itoh-Nakadai A, Calame KL, Ikebe D, et al. Bach2 represses plasma cell gene regulatory network in B cells to promote antibody class switch. *EMBO J.* 2010 Dec 1;29(23):4048–61.
87. Tellier J, Nutt SL. Standing out from the crowd: How to identify plasma cells. *Eur J Immunol.* 2017 Aug;47(8):1276–9.
88. Soro PG, Morales-A P, Martínez-M JA, Morales-A S, Copín SG, Marcos MA, et al. Differential involvement of the transcription factor Blimp-1 in T cell-independent and -dependent B cell differentiation to plasma cells. *J Immunol Baltim Md 1950.* 1999 Jul 15;163(2):611–7.
89. Kallies A, Hasbold J, Tarlinton DM, Dietrich W, Corcoran LM, Hodgkin PD, et al. Plasma cell ontogeny defined by quantitative changes in blimp-1 expression. *J Exp Med.* 2004 Oct 18;200(8):967–77.
90. Lin Y, Wong K, Calame K. Repression of c-myc transcription by Blimp-1, an inducer of terminal B cell differentiation. *Science.* 1997 Apr 25;276(5312):596–9.
91. Ise W, Fujii K, Shiroguchi K, Ito A, Kometani K, Takeda K, et al. T Follicular Helper Cell-Germinal Center B Cell Interaction Strength Regulates Entry into Plasma Cell or Recycling Germinal Center Cell Fate. *Immunity.* 2018 Apr;48(4):702–715.e4.
92. Weill J-C, Weller S, Reynaud C-A. Human Marginal Zone B Cells. *Annu Rev Immunol.* 2009 Apr;27(1):267–85.
93. Weisel F, Shlomchik M. Memory B Cells of Mice and Humans. *Annu Rev Immunol.* 2017 26;35:255–84.

94. Liu YJ, Zhang J, Lane PJ, Chan EY, MacLennan IC. Sites of specific B cell activation in primary and secondary responses to T cell-dependent and T cell-independent antigens. *Eur J Immunol.* 1991 Dec;21(12):2951–62.
95. Toyama H, Okada S, Hatano M, Takahashi Y, Takeda N, Ichii H, et al. Memory B cells without somatic hypermutation are generated from Bcl6-deficient B cells. *Immunity.* 2002 Sep;17(3):329–39.
96. Takemori T, Kaji T, Takahashi Y, Shimoda M, Rajewsky K. Generation of memory B cells inside and outside germinal centers: Highlights. *Eur J Immunol.* 2014 May;44(5):1258–64.
97. Wang Y, Shi J, Yan J, Xiao Z, Hou X, Lu P, et al. Germinal-center development of memory B cells driven by IL-9 from follicular helper T cells. *Nat Immunol.* 2017 Aug;18(8):921–30.
98. Takatsuka S, Yamada H, Haniuda K, Saruwatari H, Ichihashi M, Renauld J-C, et al. IL-9 receptor signaling in memory B cells regulates humoral recall responses. *Nat Immunol.* 2018 Sep;19(9):1025–34.
99. Scheeren FA, Naspetti M, Diehl S, Schotte R, Nagasawa M, Wijnands E, et al. STAT5 regulates the self-renewal capacity and differentiation of human memory B cells and controls Bcl-6 expression. *Nat Immunol.* 2005 Mar;6(3):303–13.
100. Reljic R, Wagner SD, Peakman LJ, Fearon DT. Suppression of signal transducer and activator of transcription 3-dependent B lymphocyte terminal differentiation by BCL-6. *J Exp Med.* 2000 Dec 18;192(12):1841–8.
101. Kabashima K, Haynes NM, Xu Y, Nutt SL, Allende ML, Proia RL, et al. Plasma cell S1P1 expression determines secondary lymphoid organ retention versus bone marrow tropism. *J Exp Med.* 2006 Nov 27;203(12):2683–90.
102. Odendahl M, Mei H, Hoyer BF, Jacobi AM, Hansen A, Muehlinghaus G, et al. Generation of migratory antigen-specific plasma blasts and mobilization of resident plasma cells in a secondary immune response. *Blood.* 2005 Feb 15;105(4):1614–21.
103. Good-Jacobson KL, O'Donnell K, Belz GT, Nutt SL, Tarlinton DM. c-Myb is required for plasma cell migration to bone marrow after immunization or infection. *J Exp Med.* 2015 Jun 29;212(7):1001–9.
104. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell.* 1999 Oct 29;99(3):247–57.
105. Hermann A, Goyal R, Jeltsch A. The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem.* 2004 Nov 12;279(46):48350–9.
106. Choy JS, Wei S, Lee JY, Tan S, Chu S, Lee T-H. DNA Methylation Increases Nucleosome Compaction and Rigidity. *J Am Chem Soc.* 2010 Feb 17;132(6):1782–3.
107. Kulis M, Merkel A, Heath S, Queirós AC, Schuyler RP, Castellano G, et al. Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat Genet.* 2015 Jul;47(7):746–56.
108. Ji H, Ehrlich LIR, Seita J, Murakami P, Doi A, Lindau P, et al. Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature.* 2010 Sep 16;467(7313):338–42.
109. Oakes CC, Seifert M, Assenov Y, Gu L, Przekopowicz M, Ruppert AS, et al. DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. *Nat Genet.* 2016 Mar;48(3):253–64.
110. Kulis M, Heath S, Bibikova M, Queirós AC, Navarro A, Clot G, et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nat Genet.* 2012 Nov;44(11):1236–42.

111. Barwick BG, Scharer CD, Bally APR, Boss JM. Plasma cell differentiation is coupled to division-dependent DNA hypomethylation and gene regulation. *Nat Immunol.* 2016 Oct;17(10):1216–25.
112. Orlanski S, Labi V, Reizel Y, Spiro A, Lichtenstein M, Levin-Klein R, et al. Tissue-specific DNA demethylation is required for proper B-cell differentiation and function. *Proc Natl Acad Sci U S A.* 2016 May 3;113(18):5018–23.
113. Lai AY, Mav D, Shah R, Grimm SA, Phadke D, Hatzi K, et al. DNA methylation profiling in human B cells reveals immune regulatory elements and epigenetic plasticity at Alu elements during B-cell activation. *Genome Res.* 2013 Dec;23(12):2030–41.
114. Caron G, Hussein M, Kulis M, Delaloy C, Chatonnet F, Pignarre A, et al. Cell-Cycle-Dependent Reconfiguration of the DNA Methylome during Terminal Differentiation of Human B Cells into Plasma Cells. *Cell Rep.* 2015 Nov 3;13(5):1059–71.
115. Shaknovich R, Cerchietti L, Tsikitas L, Kormaksson M, De S, Figueroa ME, et al. DNA methyltransferase 1 and DNA methylation patterning contribute to germinal center B-cell differentiation. *Blood.* 2011 Sep 29;118(13):3559–69.
116. Bostick M, Kim JK, Estève P-O, Clark A, Pradhan S, Jacobsen SE. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science.* 2007 Sep 21;317(5845):1760–4.
117. Chen C, Zhai S, Zhang L, Chen J, Long X, Qin J, et al. Uhrf1 regulates germinal center B cell expansion and affinity maturation to control viral infection. *J Exp Med.* 2018 07;215(5):1437–48.
118. Barwick BG, Scharer CD, Martinez RJ, Price MJ, Wein AN, Haines RR, et al. B cell activation and plasma cell differentiation are inhibited by de novo DNA methylation. *Nat Commun.* 2018 15;9(1):1900.
119. Oakes CC, Martin-Subero JI. Insight into origins, mechanisms, and utility of DNA methylation in B-cell malignancies. *Blood.* 2018 06;132(10):999–1006.
120. Giacomelli B, Zhao Q, Ruppert AS, Agyeman A, Weigel C, Wu Y-Z, et al. Developmental subtypes assessed by DNA methylation-iPLEX forecast the natural history of chronic lymphocytic leukemia. *Blood.* 2019 22;134(8):688–98.
121. Queirós AC, Beekman R, Vilarrasa-Blasi R, Duran-Ferrer M, Clot G, Merkel A, et al. Decoding the DNA Methylome of Mantle Cell Lymphoma in the Light of the Entire B Cell Lineage. *Cancer Cell.* 2016 Nov 14;30(5):806–21.
122. Reddy A, Zhang J, Davis NS, Moffitt AB, Love CL, Waldrop A, et al. Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. *Cell.* 2017 Oct 5;171(2):481-494.e15.
123. Béguelin W, Popovic R, Teater M, Jiang Y, Bunting KL, Rosen M, et al. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell.* 2013 May 13;23(5):677–92.
124. Caganova M, Carrisi C, Varano G, Mainoldi F, Zanardi F, Germain P-L, et al. Germinal center dysregulation by histone methyltransferase EZH2 promotes lymphomagenesis. *J Clin Invest.* 2013 Dec;123(12):5009–22.
125. Ortega-Molina A, Boss IW, Canela A, Pan H, Jiang Y, Zhao C, et al. The histone lysine methyltransferase KMT2D sustains a gene expression program that represses B cell lymphoma development. *Nat Med.* 2015 Oct;21(10):1199–208.
126. Zhang J, Dominguez-Sola D, Hussein S, Lee J-E, Holmes AB, Bansal M, et al. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat Med.* 2015 Oct;21(10):1190–8.
127. Jiang Y, Ortega-Molina A, Geng H, Ying H-Y, Hatzi K, Parsa S, et al. CREBBP Inactivation Promotes the Development of HDAC3-Dependent Lymphomas. *Cancer Discov.* 2017;7(1):38–53.
128. Zhang J, Vlasevska S, Wells VA, Nataraj S, Holmes AB, Duval R, et al. The CREBBP

- Acetyltransferase Is a Haploinsufficient Tumor Suppressor in B-cell Lymphoma. *Cancer Discov.* 2017;7(3):322–37.
129. Heise N, De Silva NS, Silva K, Carette A, Simonetti G, Pasparakis M, et al. Germinal center B cell maintenance and differentiation are controlled by distinct NF- κ B transcription factor subunits. *J Exp Med.* 2014 Sep 22;211(10):2103–18.
 130. Jacque E, Schweighoffer E, Visekruna A, Papoutsopoulou S, Janzen J, Zillwood R, et al. IKK-induced NF- κ B1 p105 proteolysis is critical for B cell antibody responses to T cell-dependent antigen. *J Exp Med.* 2014 Sep 22;211(10):2085–101.
 131. Arpin C, Déchanet J, Van Kooten C, Merville P, Grouard G, Brière F, et al. Generation of memory B cells and plasma cells in vitro. *Science.* 1995 May 5;268(5211):720–2.
 132. Basso K, Dalla-Favera R. Germinal centres and B cell lymphomagenesis. *Nat Rev Immunol.* 2015 Mar;15(3):172–84.
 133. Weltgesundheitsorganisation. WHO classification of tumours of haematopoietic and lymphoid tissues. Revised 4th edition. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, et al., editors. Lyon: International Agency for Research on Cancer; 2017. 585 p. (World Health Organization classification of tumours).
 134. Stevenson F, Sahota S, Zhu D, Ottensmeier C, Chapman C, Oscier D, et al. Insight into the origin and clonal history of B-cell tumors as revealed by analysis of immunoglobulin variable region genes. *Immunol Rev.* 1998 Apr;162:247–59.
 135. Küppers R, Klein U, Hansmann ML, Rajewsky K. Cellular origin of human B-cell lymphomas. *N Engl J Med.* 1999 Nov 11;341(20):1520–9.
 136. Kim SS, Ruiz VE, Carroll JD, Moss SF. *Helicobacter pylori* in the pathogenesis of gastric cancer and gastric lymphoma. *Cancer Lett.* 2011 Jun 28;305(2):228–38.
 137. Dammacco F, Gatti P, Sansonno D. Hepatitis C virus infection, mixed cryoglobulinemia, and non-Hodgkin's lymphoma: an emerging picture. *Leuk Lymphoma.* 1998 Nov;31(5–6):463–76.
 138. Michot J-M, Canioni D, Driss H, Alric L, Cacoub P, Suarez F, et al. Antiviral therapy is associated with a better survival in patients with hepatitis C virus and B-cell non-Hodgkin lymphomas, ANRS HC-13 lympho-C study. *Am J Hematol.* 2015 Mar;90(3):197–203.
 139. Sakamuro D, Furukawa T, Takegami T. Hepatitis C virus nonstructural protein NS3 transforms NIH 3T3 cells. *J Virol.* 1995 Jun;69(6):3893–6.
 140. Machida K, Cheng KT-N, Sung VM-H, Shimodaira S, Lindsay KL, Levine AM, et al. Hepatitis C virus induces a mutator phenotype: enhanced mutations of immunoglobulin and protooncogenes. *Proc Natl Acad Sci U S A.* 2004 Mar 23;101(12):4262–7.
 141. Armitage RJ, Fanslow WC, Strockbine L, Sato TA, Clifford KN, Macduff BM, et al. Molecular and biological characterization of a murine ligand for CD40. *Nature.* 1992 May 7;357(6373):80–2.
 142. Noelle RJ, Roy M, Shepherd DM, Stamenkovic I, Ledbetter JA, Aruffo A. A 39-kDa protein on activated helper T cells binds CD40 and transduces the signal for cognate activation of B cells. *Proc Natl Acad Sci U S A.* 1992 Jul 15;89(14):6550–4.
 143. Gray D, Siepmann K, Wohlleben G. CD40 ligation in B cell activation, isotype switching and memory development. *Semin Immunol.* 1994 Oct;6(5):303–10.
 144. Turner ML, Corcoran LM, Brink R, Hodgkin PD. High-affinity B cell receptor ligation by cognate antigen induces cytokine-independent isotype switching. *J Immunol Baltim Md 1950.* 2010 Jun 15;184(12):6592–9.
 145. Lenschow DJ, Sperling AI, Cooke MP, Freeman G, Rhee L, Decker DC, et al. Differential up-regulation of the B7-1 and B7-2 costimulatory molecules after Ig receptor engagement by antigen. *J Immunol Baltim Md 1950.* 1994 Sep 1;153(5):1990–7.
 146. Tarlinton DM. Evolution in miniature: selection, survival and distribution of antigen reactive cells in the germinal centre. *Immunol Cell Biol.* 2008 Feb;86(2):133–8.

147. Dalla-Favera R, Pasqualucci L. Molecular Pathogenesis of B Cell Lymphomas. In: *Molecular Biology of B Cells* [Internet]. Elsevier; 2015 [cited 2020 Aug 20]. p. 399–416. Available from: <https://linkinghub.elsevier.com/retrieve/pii/B9780123979339000229>
148. Clark MR, Tanaka A, Powers SE, Veselits M. Receptors, subcellular compartments and the regulation of peripheral B cell responses: the illuminating state of anergy. *Mol Immunol*. 2011 Jun;48(11):1281–6.
149. Dal Porto JM, Gauld SB, Merrell KT, Mills D, Pugh-Bernard AE, Cambier J. B cell antigen receptor signaling 101. *Mol Immunol*. 2004 Jul;41(6–7):599–613.
150. Thome M, Charton JE, Pelzer C, Hailfinger S. Antigen receptor signaling to NF-kappaB via CARMA1, BCL10, and MALT1. *Cold Spring Harb Perspect Biol*. 2010 Sep;2(9):a003004.
151. Johnson GL, Lapadat R. Mitogen-activated protein kinase pathways mediated by ERK, JNK, and p38 protein kinases. *Science*. 2002 Dec 6;298(5600):1911–2.
152. Kurosaki T, Hikida M. Tyrosine kinases and their substrates in B lymphocytes. *Immunol Rev*. 2009 Mar;228(1):132–48.
153. Young RM, Staudt LM. Targeting pathological B cell receptor signalling in lymphoid malignancies. *Nat Rev Drug Discov*. 2013 Mar;12(3):229–43.
154. Davis RE, Ngo VN, Lenz G, Tolar P, Young RM, Romesser PB, et al. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature*. 2010 Jan 7;463(7277):88–92.
155. Wilson WH, Young RM, Schmitz R, Yang Y, Pittaluga S, Wright G, et al. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat Med*. 2015 Aug;21(8):922–6.
156. Mackay F, Schneider P. Cracking the BAFF code. *Nat Rev Immunol*. 2009 Jul;9(7):491–502.
157. Warnatz K, Salzer U, Rizzi M, Fischer B, Gutenberger S, Böhm J, et al. B-cell activating factor receptor deficiency is associated with an adult-onset antibody deficiency syndrome in humans. *Proc Natl Acad Sci U S A*. 2009 Aug 18;106(33):13945–50.
158. Almaden JV, Liu YC, Yang E, Otero DC, Birnbaum H, Davis-Turak J, et al. B-cell survival and development controlled by the coordination of NF-κB family members RelB and cRel. *Blood*. 2016 Mar 10;127(10):1276–86.
159. He B, Santamaria R, Xu W, Cols M, Chen K, Puga I, et al. The transmembrane activator TACI triggers immunoglobulin class switching by activating B cells through the adaptor MyD88. *Nat Immunol*. 2010 Sep;11(9):836–45.
160. Salzer U, Chapel HM, Webster ADB, Pan-Hammarström Q, Schmitt-Graeff A, Schlesier M, et al. Mutations in TNFRSF13B encoding TACI are associated with common variable immunodeficiency in humans. *Nat Genet*. 2005 Aug;37(8):820–8.
161. Kawai T, Akira S. Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. *Immunity*. 2011 May 27;34(5):637–50.
162. Ntoufa S, Vilia MG, Stamatopoulos K, Ghia P, Muzio M. Toll-like receptors signaling: A complex network for NF-κB activation in B-cell lymphoid malignancies. *Semin Cancer Biol*. 2016;39:15–25.
163. Rawlings DJ, Schwartz MA, Jackson SW, Meyer-Bahlburg A. Integration of B cell responses through Toll-like receptors and antigen receptors. *Nat Rev Immunol*. 2012 Mar 16;12(4):282–94.
164. Takeda K, Akira S. TLR signaling pathways. *Semin Immunol*. 2004 Feb;16(1):3–9.
165. Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim K-H, et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature*. 2011 Feb 3;470(7332):115–9.
166. Treon SP, Xu L, Yang G, Zhou Y, Liu X, Cao Y, et al. MYD88 L265P somatic mutation in Waldenström’s macroglobulinemia. *N Engl J Med*. 2012 Aug 30;367(9):826–33.
167. Yang G, Zhou Y, Liu X, Xu L, Cao Y, Manning RJ, et al. A mutation in MYD88

- (L265P) supports the survival of lymphoplasmacytic cells by activation of Bruton tyrosine kinase in Waldenström macroglobulinemia. *Blood*. 2013 Aug 15;122(7):1222–32.
168. Staudt LM. Oncogenic activation of NF-kappaB. *Cold Spring Harb Perspect Biol*. 2010 Jun;2(6):a000109.
169. Zarnegar BJ, Wang Y, Mahoney DJ, Dempsey PW, Cheung HH, He J, et al. Noncanonical NF-kappaB activation requires coordinated assembly of a regulatory complex of the adaptors cIAP1, cIAP2, TRAF2 and TRAF3 and the kinase NIK. *Nat Immunol*. 2008 Dec;9(12):1371–8.
170. Gyrd-Hansen M, Darding M, Miasari M, Santoro MM, Zender L, Xue W, et al. IAPs contain an evolutionarily conserved ubiquitin-binding domain that regulates NF-kappaB as well as cell survival and oncogenesis. *Nat Cell Biol*. 2008 Nov;10(11):1309–17.
171. Zhang Q, Lenardo MJ, Baltimore D. 30 Years of NF- κ B: A Blossoming of Relevance to Human Pathobiology. *Cell*. 2017 Jan 12;168(1–2):37–57.
172. Kaileh M, Sen R. NF- κ B function in B lymphocytes. *Immunol Rev*. 2012 Mar;246(1):254–71.
173. Saito M, Gao J, Basso K, Kitagawa Y, Smith PM, Bhagat G, et al. A signaling pathway mediating downregulation of BCL6 in germinal center B cells is blocked by BCL6 gene alterations in B cell lymphoma. *Cancer Cell*. 2007 Sep;12(3):280–92.
174. Tavares RM, Turer EE, Liu CL, Advincula R, Scapini P, Rhee L, et al. The ubiquitin modifying enzyme A20 restricts B cell survival and prevents autoimmunity. *Immunity*. 2010 Aug 27;33(2):181–91.
175. Küppers R. Molecular biology of Hodgkin lymphoma. *Hematol Am Soc Hematol Educ Program*. 2009;491–6.
176. Schmitz R, Hansmann M-L, Bohle V, Martin-Subero JI, Hartmann S, Mechtersheimer G, et al. TNFAIP3 (A20) is a tumor suppressor gene in Hodgkin lymphoma and primary mediastinal B cell lymphoma. *J Exp Med*. 2009 May 11;206(5):981–9.
177. Honma K, Tsuzuki S, Nakagawa M, Tagawa H, Nakamura S, Morishima Y, et al. TNFAIP3/A20 functions as a novel tumor suppressor gene in several subtypes of non-Hodgkin lymphomas. *Blood*. 2009 Sep 17;114(12):2467–75.
178. Manier S, Salem KZ, Park J, Landau DA, Getz G, Ghobrial IM. Genomic complexity of multiple myeloma and its clinical implications. *Nat Rev Clin Oncol*. 2017;14(2):100–13.
179. Damm F, Mylonas E, Cosson A, Yoshida K, Della Valle V, Mouly E, et al. Acquired initiating mutations in early hematopoietic cells of CLL patients. *Cancer Discov*. 2014 Sep;4(9):1088–101.
180. Della-Valle V, Roos-Weil D, Scourzic L, Mouly E, Aid Z, Darwiche W, et al. Nfkb-deficiency leads to increased susceptibility to develop B-cell lymphoproliferative disorders in aged mice. *Blood Cancer J*. 2020 13;10(3):38.
181. Rossi D, Deaglio S, Dominguez-Sola D, Rasi S, Vaisitti T, Agostinelli C, et al. Alteration of BIRC3 and multiple other NF- κ B pathway genes in splenic marginal zone lymphoma. *Blood*. 2011 Nov 3;118(18):4930–4.
182. Rossi D, Gaidano G. The clinical implications of gene mutations in chronic lymphocytic leukaemia. *Br J Cancer*. 2016 Apr 12;114(8):849–54.
183. Pone EJ, Zhang J, Mai T, White CA, Li G, Sakakura JK, et al. BCR-signalling synergizes with TLR-signalling for induction of AID and immunoglobulin class-switching through the non-canonical NF- κ B pathway. *Nat Commun*. 2012 Jan;3(1):767.
184. Freeman SA, Jaumouillé V, Choi K, Hsu BE, Wong HS, Abraham L, et al. Toll-like receptor ligands sensitize B-cell receptor signalling by reducing actin-dependent spatial confinement of the receptor. *Nat Commun*. 2015 May;6(1):6168.
185. Schweighoffer E, Nys J, Vanes L, Smithers N, Tybulewicz VLJ. TLR4 signals in B lymphocytes are transduced via the B cell antigen receptor and SYK. *J Exp Med*. 2017 May

1;214(5):1269–80.

186. Rui L, Healy JI, Blasioli J, Goodnow CC. ERK signaling is a molecular switch integrating opposing inputs from B cell receptor and T cell cytokines to control TLR4-driven plasma cell differentiation. *J Immunol Baltim Md 1950*. 2006 Oct 15;177(8):5337–46.

187. Munshi M, Liu X, Chen JG, Xu L, Tsakmaklis N, Demos MG, et al. SYK is activated by mutated MYD88 and drives pro-survival signaling in MYD88 driven B-cell lymphomas. *Blood Cancer J*. 2020 Jan;10(1):12.

188. Hawkins ED, Turner ML, Wellard CJ, Zhou JHS, Dowling MR, Hodgkin PD. Quantal and graded stimulation of B lymphocytes as alternative strategies for regulating adaptive immune responses. *Nat Commun*. 2013 Dec;4(1):2406.

189. Renshaw BR, Fanslow WC, Armitage RJ, Campbell KA, Liggitt D, Wright B, et al. Humoral immune responses in CD40 ligand-deficient mice. *J Exp Med*. 1994 Nov 1;180(5):1889–900.

190. Kawabe T, Naka T, Yoshida K, Tanaka T, Fujiwara H, Suematsu S, et al. The immune responses in CD40-deficient mice: impaired immunoglobulin class switching and germinal center formation. *Immunity*. 1994 Jun;1(3):167–78.

191. Choe J, Kim HS, Armitage RJ, Choi YS. The functional role of B cell antigen receptor stimulation and IL-4 in the generation of human memory B cells from germinal center B cells. *J Immunol Baltim Md 1950*. 1997 Oct 15;159(8):3757–66.

192. Ozaki K, Spolski R, Feng CG, Qi C-F, Cheng J, Sher A, et al. A critical role for IL-21 in regulating immunoglobulin production. *Science*. 2002 Nov 22;298(5598):1630–4.

193. Ozaki K, Spolski R, Ettinger R, Kim H-P, Wang G, Qi C-F, et al. Regulation of B cell differentiation and plasma cell generation by IL-21, a novel inducer of Blimp-1 and Bcl-6. *J Immunol Baltim Md 1950*. 2004 Nov 1;173(9):5361–71.

194. Zotos D, Coquet JM, Zhang Y, Light A, D’Costa K, Kallies A, et al. IL-21 regulates germinal center B cell differentiation and proliferation through a B cell-intrinsic mechanism. *J Exp Med*. 2010 Feb 15;207(2):365–78.

195. Moens L, Tangye SG. Cytokine-Mediated Regulation of Plasma Cell Generation: IL-21 Takes Center Stage. *Front Immunol [Internet]*. 2014 [cited 2020 Apr 28];5. Available from: <http://journal.frontiersin.org/article/10.3389/fimmu.2014.00065/abstract>

196. Schilizzi BM, Boonstra R, The TH, de Leij LF. Effect of B-cell receptor engagement on CD40-stimulated B cells. *Immunology*. 1997 Nov;92(3):346–53.

197. Hoshino K, Takeuchi O, Kawai T, Sanjo H, Ogawa T, Takeda Y, et al. Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product. *J Immunol Baltim Md 1950*. 1999 Apr 1;162(7):3749–52.

198. Mingari MC, Gerosa F, Carra G, Accolla RS, Moretta A, Zubler RH, et al. Human interleukin-2 promotes proliferation of activated B cells via surface receptors similar to those of activated T cells. *Nature*. 1984 Dec;312(5995):641–3.

199. Le Gallou S, Caron G, Delaloy C, Rossille D, Tarte K, Fest T. IL-2 Requirement for Human Plasma Cell Generation: Coupling Differentiation and Proliferation by Enhancing MAPK–ERK Signaling. *J Immunol*. 2012 Jul 1;189(1):161–73.

200. Itoh K, Inoue T, Ito K, Hirohata S. The Interplay of Interleukin-10 (IL-10) and Interleukin-2 (IL-2) in Humoral Immune Responses: IL-10 Synergizes with IL-2 to Enhance Responses of Human B Lymphocytes in a Mechanism Which Is Different from Upregulation of CD25 Expression. *Cell Immunol*. 1994 Sep;157(2):478–88.

201. Horikawa K, Takatsu K. Interleukin-5 regulates genes involved in B-cell terminal maturation. *Immunology*. 2006 Aug;118(4):497–508.

202. Baptista BJA, Granato A, Canto FB, Montalvão F, Tostes L, de Matos Guedes HL, et al. TLR9 Signaling Suppresses the Canonical Plasma Cell Differentiation Program in Follicular B Cells. *Front Immunol*. 2018;9:2281.

203. Nojima T, Haniuda K, Moutai T, Matsudaira M, Mizokawa S, Shiratori I, et al. In-vitro derived germinal centre B cells differentially generate memory B or plasma cells in vivo. *Nat Commun.* 2011 Sep;2(1):465.
204. Haniuda K, Nojima T, Kitamura D. In Vitro-Induced Germinal Center B Cell Culture System. *Methods Mol Biol Clifton NJ.* 2017;1623:125–33.
205. Bestor TH, Coxon A. Cytosine methylation: the pros and cons of DNA methylation. *Curr Biol CB.* 1993 Jun 1;3(6):384–6.
206. Cedar H, Bergman Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet.* 2009 May;10(5):295–304.
207. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet.* 2008 Jun;9(6):465–76.
208. Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet.* 2017 Sep;18(9):517–34.
209. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science.* 2009 May 15;324(5929):930–5.
210. Iyer LM, Tahiliani M, Rao A, Aravind L. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle Georget Tex.* 2009 Jun 1;8(11):1698–710.
211. Loenarz C, Schofield CJ. Oxygenase catalyzed 5-methylcytosine hydroxylation. *Chem Biol.* 2009 Jun 26;16(6):580–3.
212. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science.* 2011 Sep 2;333(6047):1300–3.
213. He Y-F, Li B-Z, Li Z, Liu P, Wang Y, Tang Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science.* 2011 Sep 2;333(6047):1303–7.
214. Spruijt CG, Gnerlich F, Smits AH, Pfaffeneder T, Jansen PWTC, Bauer C, et al. Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives. *Cell.* 2013 Feb;152(5):1146–59.
215. Kinde B, Gabel HW, Gilbert CS, Griffith EC, Greenberg ME. Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2. *Proc Natl Acad Sci U S A.* 2015 Jun 2;112(22):6800–6.
216. Ecsedi S, Rodríguez-Aguilera J, Hernandez-Vargas H. 5-Hydroxymethylcytosine (5hmC), or How to Identify Your Favorite Cell. *Epigenomes.* 2018 Jan 30;2(1):3.
217. Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi YG. TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res.* 2010 Dec;20(12):1390–3.
218. Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, et al. Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell.* 2011 May 20;42(4):451–64.
219. Kosmider O, Gelsi-Boyer V, Cheok M, Grabar S, Della-Valle V, Picard F, et al. TET2 mutation is an independent favorable prognostic factor in myelodysplastic syndromes (MDSs). *Blood.* 2009 Oct 8;114(15):3285–91.
220. Wang M, Zhang S, Chuang S-S, Ashton-Key M, Ochoa E, Bolli N, et al. Angioimmunoblastic T cell lymphoma: novel molecular insights by mutation profiling. *Oncotarget.* 2017 Mar 14;8(11):17763–70.
221. Lorschach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR. TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia.* 2003 Mar;17(3):637–41.

222. Abdel-Wahab O, Mullally A, Hedvat C, Garcia-Manero G, Patel J, Wadleigh M, et al. Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood*. 2009 Jul 2;114(1):144–7.
223. Albano F, Anelli L, Zagaria A, Coccaro N, Minervini A, Rossi AR, et al. Decreased TET2 gene expression during chronic myeloid leukemia progression. *Leuk Res*. 2011 Nov;35(11):e220-222.
224. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*. 2010 Dec 14;18(6):553–67.
225. Weissmann S, Alpermann T, Grossmann V, Kowarsch A, Nadarajah N, Eder C, et al. Landscape of TET2 mutations in acute myeloid leukemia. *Leukemia*. 2012 May;26(5):934–42.
226. Chou W-C, Chou S-C, Liu C-Y, Chen C-Y, Hou H-A, Kuo Y-Y, et al. TET2 mutation is an unfavorable prognostic factor in acute myeloid leukemia patients with intermediate-risk cytogenetics. *Blood*. 2011 Oct 6;118(14):3803–10.
227. Li W, Liu M. Distribution of 5-hydroxymethylcytosine in different human tissues. *J Nucleic Acids*. 2011;2011:870726.
228. Haffner MC, Chaux A, Meeker AK, Esopi DM, Gerber J, Pellakuru LG, et al. Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell compartments and in human cancers. *Oncotarget*. 2011 Aug;2(8):627–37.
229. Jin S-G, Jiang Y, Qiu R, Rauch TA, Wang Y, Schackert G, et al. 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. *Cancer Res*. 2011 Dec 15;71(24):7360–5.
230. Mayer W, Niveleau A, Walter J, Fundele R, Haaf T. Demethylation of the zygotic paternal genome. *Nature*. 2000 Feb 3;403(6769):501–2.
231. Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, et al. Active demethylation of the paternal genome in the mouse zygote. *Curr Biol CB*. 2000 Apr 20;10(8):475–8.
232. Iqbal K, Jin S-G, Pfeifer GP, Szabó PE. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci U S A*. 2011 Mar 1;108(9):3642–7.
233. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, et al. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun*. 2011;2:241.
234. Inoue A, Zhang Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science*. 2011 Oct 14;334(6053):194.
235. Gu T-P, Guo F, Yang H, Wu H-P, Xu G-F, Liu W, et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*. 2011 Sep 4;477(7366):606–10.
236. Hon GC, Song C-X, Du T, Jin F, Selvaraj S, Lee AY, et al. 5mC Oxidation by Tet2 Modulates Enhancer Activity and Timing of Transcriptome Reprogramming during Differentiation. *Mol Cell*. 2014 Oct;56(2):286–97.
237. Huang Y, Chavez L, Chang X, Wang X, Pastor WA, Kang J, et al. Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc Natl Acad Sci*. 2014 Jan 28;111(4):1361–6.
238. Tsagaratou A, Aijo T, Lio C-WJ, Yue X, Huang Y, Jacobsen SE, et al. Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation. *Proc Natl Acad Sci*. 2014 Aug 12;111(32):E3306–15.
239. de la Rica L, Rodríguez-Ubreva J, García M, Islam ABMMK, Urquiza JM, Hernando H, et al. PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated methylation in monocyte-to-osteoclast differentiation. *Genome Biol*. 2013;14(9):R99.
240. Lio C-W, Zhang J, González-Avalos E, Hogan PG, Chang X, Rao A. Tet2 and Tet3

cooperate with B-lineage transcription factors to regulate DNA modification and chromatin accessibility. *eLife*. 2016 21;5.

241. Costa Y, Ding J, Theunissen TW, Faiola F, Hore TA, Shliaha PV, et al. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature*. 2013 Mar;495(7441):370–4.

242. An J, González-Avalos E, Chawla A, Jeong M, López-Moyado IF, Li W, et al. Acute loss of TET function results in aggressive myeloid cancer in mice. *Nat Commun*. 2015 Dec;6(1):10071.

243. Ko M, An J, Pastor WA, Koralov SB, Rajewsky K, Rao A. TET proteins and 5-methylcytosine oxidation in hematological cancers. *Immunol Rev*. 2015 Jan;263(1):6–21.

244. Lio C-WJ, Shukla V, Samaniego-Castruita D, González-Avalos E, Chakraborty A, Yue X, et al. TET enzymes augment activation-induced deaminase (AID) expression via 5-hydroxymethylcytosine modifications at the Aicda superenhancer. *Sci Immunol*. 2019 26;4(34).

245. Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, et al. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell*. 2011 Feb 4;8(2):200–13.

246. Song SJ, Ito K, Ala U, Kats L, Webster K, Sun SM, et al. The oncogenic microRNA miR-22 targets the TET2 tumor suppressor to promote hematopoietic stem cell self-renewal and transformation. *Cell Stem Cell*. 2013 Jul 3;13(1):87–101.

247. Ko M, An J, Bandukwala HS, Chavez L, Aijö T, Pastor WA, et al. Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature*. 2013 May 2;497(7447):122–6.

248. Wang Y, Zhang Y. Regulation of TET protein stability by calpains. *Cell Rep*. 2014 Jan 30;6(2):278–84.

249. Yin R, Mao S-Q, Zhao B, Chong Z, Yang Y, Zhao C, et al. Ascorbic acid enhances Tet-mediated 5-methylcytosine oxidation and promotes DNA demethylation in mammals. *J Am Chem Soc*. 2013 Jul 17;135(28):10396–403.

250. Blaschke K, Ebata KT, Karimi MM, Zepeda-Martínez JA, Goyal P, Mahapatra S, et al. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature*. 2013 Aug 8;500(7461):222–6.

251. Minor EA, Court BL, Young JI, Wang G. Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. *J Biol Chem*. 2013 May 10;288(19):13669–74.

252. Yue X, Rao A. TET-family dioxygenases and the TET activator Vitamin C in immune responses and cancer. *Blood*. 2020 Jul 30;blood.2019004158.

253. Park S. The effects of high concentrations of vitamin C on cancer cells. *Nutrients*. 2013 Sep 9;5(9):3496–505.

254. Park CH. Biological nature of the effect of ascorbic acids on the growth of human leukemic cells. *Cancer Res*. 1985 Aug;45(8):3969–73.

255. Park CH, Kimler BF, Bodensteiner D, Lynch SR, Hassanein RS. In vitro growth modulation by L-ascorbic acid of colony-forming cells from bone marrow of patients with myelodysplastic syndromes. *Cancer Res*. 1992 Aug 15;52(16):4458–66.

256. Agathocleous M, Meacham CE, Burgess RJ, Piskounova E, Zhao Z, Crane GM, et al. Ascorbate regulates haematopoietic stem cell function and leukaemogenesis. *Nature*. 2017 Sep;549(7673):476–81.

257. Cimmino L, Dolgalev I, Wang Y, Yoshimi A, Martin GH, Wang J, et al. Restoration of TET2 Function Blocks Aberrant Self-Renewal and Leukemia Progression. *Cell*. 2017 Sep;170(6):1079-1095.e20.

258. Wang Y, Xiao M, Chen X, Chen L, Xu Y, Lv L, et al. WT1 recruits TET2 to regulate

- its target gene expression and suppress leukemia cell proliferation. *Mol Cell*. 2015 Feb 19;57(4):662–73.
259. Xiong J, Zhang Z, Chen J, Huang H, Xu Y, Ding X, et al. Cooperative Action between SALL4A and TET Proteins in Stepwise Oxidation of 5-Methylcytosine. *Mol Cell*. 2016 01;64(5):913–25.
260. Chen Q, Chen Y, Bian C, Fujiki R, Yu X. TET2 promotes histone O-GlcNAcylation during gene transcription. *Nature*. 2013 Jan 24;493(7433):561–4.
261. Deplus R, Delatte B, Schwinn MK, Defrance M, Méndez J, Murphy N, et al. TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J*. 2013 Mar 6;32(5):645–55.
262. Ichiyama K, Chen T, Wang X, Yan X, Kim B-S, Tanaka S, et al. The Methylcytosine Dioxygenase Tet2 Promotes DNA Demethylation and Activation of Cytokine Gene Expression in T Cells. *Immunity*. 2015 Apr;42(4):613–26.
263. Quesada V, Conde L, Villamor N, Ordóñez GR, Jares P, Bassaganyas L, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet*. 2011 Dec 11;44(1):47–52.
264. Dolnik A, Engelmann JC, Scharfenberger-Schmeer M, Mauch J, Kelkenberg-Schade S, Haldemann B, et al. Commonly altered genomic regions in acute myeloid leukemia are enriched for somatic mutations involved in chromatin remodeling and splicing. *Blood*. 2012 Nov 1;120(18):e83–92.
265. Kalender Atak Z, De Keersmaecker K, Gianfelici V, Geerdens E, Vandepoel R, Pauwels D, et al. High accuracy mutation detection in leukemia on a selected panel of cancer genes. *PloS One*. 2012;7(6):e38463.
266. Palomero T, Couronné L, Khiabani H, Kim M-Y, Ambesi-Impiombato A, Perez-Garcia A, et al. Recurrent mutations in epigenetic regulators, RHOA and FYN kinase in peripheral T cell lymphomas. *Nat Genet*. 2014 Feb;46(2):166–70.
267. Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Massé A, et al. Mutation in TET2 in myeloid cancers. *N Engl J Med*. 2009 May 28;360(22):2289–301.
268. Langemeijer SMC, Kuiper RP, Berends M, Knops R, Aslanyan MG, Massop M, et al. Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat Genet*. 2009 Jul;41(7):838–42.
269. Tefferi A, Pardanani A, Lim K-H, Abdel-Wahab O, Lasho TL, Patel J, et al. TET2 mutations and their clinical correlates in polycythemia vera, essential thrombocythemia and myelofibrosis. *Leukemia*. 2009 May;23(5):905–11.
270. Tefferi A, Lim K-H, Abdel-Wahab O, Lasho TL, Patel J, Patnaik MM, et al. Detection of mutant TET2 in myeloid malignancies other than myeloproliferative neoplasms: CMML, MDS, MDS/MPN and AML. *Leukemia*. 2009 Jul;23(7):1343–5.
271. Grossmann V, Kohlmann A, Eder C, Haferlach C, Kern W, Cross NCP, et al. Molecular profiling of chronic myelomonocytic leukemia reveals diverse mutations in >80% of patients with TET2 and EZH2 being of high prognostic relevance. *Leukemia*. 2011 May;25(5):877–9.
272. Nibourel O, Kosmider O, Cheok M, Boissel N, Renneville A, Philippe N, et al. Incidence and prognostic value of TET2 alterations in de novo acute myeloid leukemia achieving complete remission. *Blood*. 2010 Aug 19;116(7):1132–5.
273. Ortmann CA, Kent DG, Nangalia J, Silber Y, Wedge DC, Grinfeld J, et al. Effect of Mutation Order on Myeloproliferative Neoplasms. *N Engl J Med*. 2015 Feb 12;372(7):601–12.
274. Bejar R, Stevenson K, Abdel-Wahab O, Galili N, Nilsson B, Garcia-Manero G, et al. Clinical effect of point mutations in myelodysplastic syndromes. *N Engl J Med*. 2011 Jun 30;364(26):2496–506.
275. Itzykson R, Kosmider O, Renneville A, Gelsi-Boyer V, Meggendorfer M, Morabito M, et al. Prognostic score including gene mutations in chronic myelomonocytic leukemia. *J Clin*

- Oncol Off J Am Soc Clin Oncol. 2013 Jul 1;31(19):2428–36.
276. Asmar F, Punj V, Christensen J, Pedersen MT, Pedersen A, Nielsen AB, et al. Genome-wide profiling identifies a DNA methylation signature that associates with TET2 mutations in diffuse large B-cell lymphoma. *Haematologica*. 2013 Dec;98(12):1912–20.
277. Meissner B, Kridel R, Lim RS, Rogic S, Tse K, Scott DW, et al. The E3 ubiquitin ligase UBR5 is recurrently mutated in mantle cell lymphoma. *Blood*. 2013 Apr 18;121(16):3161–4.
278. Lemonnier F, Couronné L, Parrens M, Jaïs J-P, Travert M, Lamant L, et al. Recurrent TET2 mutations in peripheral T-cell lymphomas correlate with TFH-like features and adverse clinical parameters. *Blood*. 2012 Aug 16;120(7):1466–9.
279. Couronné L, Bastard C, Bernard OA. TET2 and DNMT3A mutations in human T-cell lymphoma. *N Engl J Med*. 2012 Jan 5;366(1):95–6.
280. Huang Y, Rao A. Connections between TET proteins and aberrant DNA modification in cancer. *Trends Genet TIG*. 2014 Oct;30(10):464–74.
281. Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*. 2010 Dec 9;468(7325):839–43.
282. Abdel-Wahab O, Gao J, Adli M, Dey A, Trimarchi T, Chung YR, et al. Deletion of *Asx11* results in myelodysplasia and severe developmental defects in vivo. *J Exp Med*. 2013 Nov 18;210(12):2641–59.
283. Muto T, Sashida G, Oshima M, Wendt GR, Mochizuki-Kashio M, Nagata Y, et al. Concurrent loss of *Ezh2* and *Tet2* cooperates in the pathogenesis of myelodysplastic disorders. *J Exp Med*. 2013 Nov 18;210(12):2627–39.
284. Haferlach T, Nagata Y, Grossmann V, Okuno Y, Bacher U, Nagae G, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia*. 2014 Feb;28(2):241–7.
285. Sakata-Yanagimoto M, Enami T, Yoshida K, Shiraishi Y, Ishii R, Miyake Y, et al. Somatic RHOA mutation in angioimmunoblastic T cell lymphoma. *Nat Genet*. 2014 Feb;46(2):171–5.
286. Moran-Crusio K, Reavie L, Shih A, Abdel-Wahab O, Ndiaye-Lobry D, Lobry C, et al. *Tet2* loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell*. 2011 Jul 12;20(1):11–24.
287. Ko M, Bandukwala HS, An J, Lamperti ED, Thompson EC, Hastie R, et al. Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc Natl Acad Sci U S A*. 2011 Aug 30;108(35):14566–71.
288. Quivoron C, Couronné L, Della Valle V, Lopez CK, Plo I, Wagner-Ballon O, et al. TET2 inactivation results in pleiotropic hematopoietic abnormalities in mouse and is a recurrent event during human lymphomagenesis. *Cancer Cell*. 2011 Jul 12;20(1):25–38.
289. Solary E, Bernard OA, Tefferi A, Fuks F, Vainchenker W. The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases. *Leukemia*. 2014 Mar;28(3):485–96.
290. Pronier E, Almire C, Mokrani H, Vasanthakumar A, Simon A, da Costa Reis Monte Mor B, et al. Inhibition of TET2-mediated conversion of 5-methylcytosine to 5-hydroxymethylcytosine disturbs erythroid and granulomonocytic differentiation of human hematopoietic progenitors. *Blood*. 2011 Sep 1;118(9):2551–5.
291. Li Z, Cai X, Cai C-L, Wang J, Zhang W, Petersen BE, et al. Deletion of *Tet2* in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood*. 2011 Oct 27;118(17):4509–18.
292. Kaasinen E, Kuismin O, Rajamäki K, Ristolainen H, Aavikko M, Kondelin J, et al. Impact of constitutional TET2 haploinsufficiency on molecular and clinical phenotype in humans. *Nat Commun*. 2019 Dec;10(1):1252.

293. Duployez N, Goursaud L, Fenwarth L, Bories C, Marceau-Renaut A, Boyer T, et al. Familial myeloid malignancies with germline TET2 mutation. *Leukemia*. 2020 May;34(5):1450–3.
294. Stremenova Spegarova J, Lawless D, Mohamad SMB, Engelhardt KR, Doody G, Shrimpton J, et al. Germline TET2 loss of function causes childhood immunodeficiency and lymphoma. *Blood*. 2020 Aug 27;136(9):1055–66.
295. Schoeler K, Aufschnaiter A, Messner S, Derudder E, Herzog S, Villunger A, et al. TET enzymes control antibody production and shape the mutational landscape in germinal centre B cells. *FEBS J*. 2019 May 23;
296. Tsagaratou A, Lio C-WJ, Yue X, Rao A. TET Methylcytosine Oxidases in T Cell and B Cell Development and Function. *Front Immunol*. 2017;8:220.
297. Mouly E, Ghamlouch H, Della-Valle V, Scourzic L, Quivoron C, Roos-Weil D, et al. B-cell tumor development in Tet2-deficient mice. *Blood Adv*. 2018 Mar 27;2(6):703–14.
298. Tanaka S, Ise W, Inoue T, Ito A, Ono C, Shima Y, et al. Tet2 and Tet3 in B cells are required to repress CD86 and prevent autoimmunity. *Nat Immunol*. 2020 Aug;21(8):950–61.
299. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010 May 28;38(4):576–89.
300. Ghisletti S, Barozzi I, Mietton F, Polletti S, De Santa F, Venturini E, et al. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity*. 2010 Mar 26;32(3):317–28.
301. Guilhamon P, Eskandarpour M, Halai D, Wilson GA, Feber A, Teschendorff AE, et al. Meta-analysis of IDH-mutant cancers identifies EBF1 as an interaction partner for TET2. *Nat Commun*. 2013;4:2166.
302. Schmitz R, Wright GW, Huang DW, Johnson CA, Phelan JD, Wang JQ, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *N Engl J Med*. 2018 12;378(15):1396–407.
303. Buscarlet M, Provost S, Zada YF, Barhdadi A, Bourgoin V, Lépine G, et al. DNMT3A and TET2 dominate clonal hematopoiesis and demonstrate benign phenotypes and different genetic predispositions. *Blood*. 2017 10;130(6):753–62.
304. Gutierrez-Hartmann A, Duval DL, Bradford AP. ETS transcription factors in endocrine systems. *Trends Endocrinol Metab*. 2007 Jun;18(4):150–8.
305. Sizemore GM, Pitarresi JR, Balakrishnan S, Ostrowski MC. The ETS family of oncogenic transcription factors in solid tumours. *Nat Rev Cancer*. 2017;17(6):337–51.
306. Hollenhorst PC, McIntosh LP, Graves BJ. Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu Rev Biochem*. 2011;80:437–71.
307. Oh S, Shin S, Janknecht R. ETV1, 4 and 5: an oncogenic subfamily of ETS transcription factors. *Biochim Biophys Acta*. 2012 Aug;1826(1):1–12.
308. Sharrocks AD. The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol*. 2001 Nov;2(11):827–37.
309. Ma Y, Zhang P, Wang F, Yang J, Yang Z, Qin H. The relationship between early embryo development and tumorigenesis. *J Cell Mol Med*. 2010 Dec;14(12):2697–701.
310. Adamo P, Lodomery MR. The oncogene ERG: a key factor in prostate cancer. *Oncogene*. 2016 Jan 28;35(4):403–14.
311. Li Y, Luo H, Liu T, Zacksenhaus E, Ben-David Y. The ets transcription factor Fli-1 in development, cancer and disease. *Oncogene*. 2015 Apr 16;34(16):2022–31.
312. Dittmer J. The role of the transcription factor Ets1 in carcinoma. *Semin Cancer Biol*. 2015 Dec;35:20–38.
313. Hock H, Shimamura A. ETV6 in hematopoiesis and leukemia predisposition. *Semin Hematol*. 2017;54(2):98–104.

314. Hollenhorst PC, Jones DA, Graves BJ. Expression profiles frame the promoter specificity dilemma of the ETS family of transcription factors. *Nucleic Acids Res.* 2004;32(18):5693–702.
315. Szymczyna BR, Arrowsmith CH. DNA binding specificity studies of four ETS proteins support an indirect read-out mechanism of protein-DNA recognition. *J Biol Chem.* 2000 Sep 15;275(37):28363–70.
316. Ray-Gallet D, Mao C, Tavitian A, Moreau-Gachelin F. DNA binding specificities of Spi-1/PU.1 and Spi-B transcription factors and identification of a Spi-1/Spi-B binding site in the c-fes/c-fps promoter. *Oncogene.* 1995 Jul 20;11(2):303–13.
317. Shore P, Sharrocks AD. The ETS-domain transcription factors Elk-1 and SAP-1 exhibit differential DNA binding specificities. *Nucleic Acids Res.* 1995 Nov 25;23(22):4698–706.
318. Wei G-H, Badis G, Berger MF, Kivioja T, Palin K, Enge M, et al. Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J.* 2010 Jul 7;29(13):2147–60.
319. Oettgen P, Finger E, Sun Z, Akbarali Y, Thamrongsak U, Boltax J, et al. PDEF, a novel prostate epithelium-specific ets transcription factor, interacts with the androgen receptor and activates prostate-specific antigen gene expression. *J Biol Chem.* 2000 Jan 14;275(2):1216–25.
320. Hollenhorst PC, Chandler KJ, Poulsen RL, Johnson WE, Speck NA, Graves BJ. DNA specificity determinants associate with distinct transcription factor functions. *PLoS Genet.* 2009 Dec;5(12):e1000778.
321. Baillat D, Laitem C, Leprivier G, Margerin C, Aumercier M. Ets-1 binds cooperatively to the palindromic Ets-binding sites in the p53 promoter. *Biochem Biophys Res Commun.* 2009 Jan 9;378(2):213–7.
322. Garvie CW, Pufall MA, Graves BJ, Wolberger C. Structural analysis of the autoinhibition of Ets-1 and its role in protein partnerships. *J Biol Chem.* 2002 Nov 22;277(47):45529–36.
323. Bojović BB, Hassell JA. The PEA3 Ets transcription factor comprises multiple domains that regulate transactivation and DNA binding. *J Biol Chem.* 2001 Feb 9;276(6):4509–21.
324. Hagman J, Grosschedl R. An inhibitory carboxyl-terminal domain in Ets-1 and Ets-2 mediates differential binding of ETS family factors to promoter sequences of the mb-1 gene. *Proc Natl Acad Sci U S A.* 1992 Oct 1;89(19):8889–93.
325. Wasylyk C, Kerckaert JP, Wasylyk B. A novel modulator domain of Ets transcription factors. *Genes Dev.* 1992 Jun;6(6):965–74.
326. Li R, Pei H, Watson DK. Regulation of Ets function by protein - protein interactions. *Oncogene.* 2000 Dec 18;19(55):6514–23.
327. Brass AL, Zhu AQ, Singh H. Assembly requirements of PU.1-Pip (IRF-4) activator complexes: inhibiting function in vivo using fused dimers. *EMBO J.* 1999 Feb 15;18(4):977–91.
328. Kim WY, Sieweke M, Ogawa E, Wee HJ, Englmeier U, Graf T, et al. Mutual activation of Ets-1 and AML1 DNA binding by direct interaction of their autoinhibitory domains. *EMBO J.* 1999 Mar 15;18(6):1609–20.
329. Hassler M, Richmond TJ. The B-box dominates SAP-1-SRF interactions in the structure of the ternary complex. *EMBO J.* 2001 Jun 15;20(12):3018–28.
330. Mo Y, Vaessen B, Johnston K, Marmorstein R. Structure of the elk-1-DNA complex reveals how DNA-distal residues affect ETS domain recognition of DNA. *Nat Struct Biol.* 2000 Apr;7(4):292–7.
331. Yamamoto H, Kihara-Negishi F, Yamada T, Hashimoto Y, Oikawa T. Physical and functional interactions between the transcription factor PU.1 and the coactivator CBP. *Oncogene.* 1999 Feb;18(7):1495–501.
332. Hollenhorst PC, Shah AA, Hopkins C, Graves BJ. Genome-wide analyses reveal

- properties of redundant and specific promoter occupancy within the ETS gene family. *Genes Dev.* 2007 Aug 1;21(15):1882–94.
333. Farnham PJ. Insights from genomic profiling of transcription factors. *Nat Rev Genet.* 2009 Sep;10(9):605–16.
334. Boros J, Donaldson IJ, O'Donnell A, Odrowaz ZA, Zeef L, Lupien M, et al. Elucidation of the ELK1 target gene network reveals a role in the coordinate regulation of core components of the gene regulation machinery. *Genome Res.* 2009 Nov;19(11):1963–73.
335. Heidenreich B, Kumar R. TERT promoter mutations in telomere biology. *Mutat Res.* 2017 Mar;771:15–31.
336. Ben-David Y, Giddens EB, Bernstein A. Identification and mapping of a common proviral integration site Fli-1 in erythroleukemia cells induced by Friend murine leukemia virus. *Proc Natl Acad Sci U S A.* 1990 Feb;87(4):1332–6.
337. Tsuzuki S, Taguchi O, Seto M. Promotion and maintenance of leukemia by ERG. *Blood.* 2011 Apr 7;117(14):3858–68.
338. Martens JHA. Acute myeloid leukemia: a central role for the ETS factor ERG. *Int J Biochem Cell Biol.* 2011 Oct;43(10):1413–6.
339. Marcucci G, Baldus CD, Ruppert AS, Radmacher MD, Mrózek K, Whitman SP, et al. Overexpression of the ETS-related gene, ERG, predicts a worse outcome in acute myeloid leukemia with normal karyotype: a Cancer and Leukemia Group B study. *J Clin Oncol Off J Am Soc Clin Oncol.* 2005 Dec 20;23(36):9234–42.
340. Kim CA, Phillips ML, Kim W, Gingery M, Tran HH, Robinson MA, et al. Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. *EMBO J.* 2001 Aug 1;20(15):4173–82.
341. Moreau-Gachelin F, Tavitian A, Tambourin P. Spi-1 is a putative oncogene in virally induced murine erythroleukaemias. *Nature.* 1988 Jan 21;331(6153):277–80.
342. Nguyen VC, Ray D, Gross MS, de Tand MF, Frézal J, Moreau-Gachelin F. Localization of the human oncogene SPI1 on chromosome 11, region p11.22. *Hum Genet.* 1990 May;84(6):542–6.
343. Poon GMK, Macgregor RB. Base coupling in sequence-specific site recognition by the ETS domain of murine PU.1. *J Mol Biol.* 2003 May 9;328(4):805–19.
344. Joo M, Park GY, Wright JG, Blackwell TS, Atchison ML, Christman JW. Transcriptional regulation of the cyclooxygenase-2 gene in macrophages by PU.1. *J Biol Chem.* 2004 Feb 20;279(8):6658–65.
345. Pongubala J, Van Beveren C, Nagulapalli S, Klemsz M, McKercher, Maki R, et al. Effect of PU.1 phosphorylation on interaction with NF-EM5 and transcriptional activation. *Science.* 1993 Mar 12;259(5101):1622–5.
346. Back J, Allman D, Chan S, Kastner P. Visualizing PU.1 activity during hematopoiesis. *Exp Hematol.* 2005 Apr;33(4):395–402.
347. Nutt SL, Metcalf D, D'Amico A, Polli M, Wu L. Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. *J Exp Med.* 2005 Jan 17;201(2):221–31.
348. DeKoter RP, Singh H. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science.* 2000 May 26;288(5470):1439–41.
349. Mak KS, Funnell APW, Pearson RCM, Crossley M. PU.1 and Haematopoietic Cell Fate: Dosage Matters. *Int J Cell Biol.* 2011;2011:808524.
350. Gangenahalli GU, Gupta P, Saluja D, Verma YK, Kishore V, Chandra R, et al. Stem Cell Fate Specification: Role of Master Regulatory Switch Transcription Factor PU.1 in Differential Hematopoiesis. *Stem Cells Dev.* 2005 Apr;14(2):140–52.
351. Spooner CJ, Cheng JX, Pujadas E, Laslo P, Singh H. A recurrent network involving the transcription factors PU.1 and Gfi1 orchestrates innate and adaptive immune cell fates. *Immunity.* 2009 Oct 16;31(4):576–86.

352. Laslo P, Spooner CJ, Warmflash A, Lancki DW, Lee H-J, Sciammas R, et al. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell*. 2006 Aug 25;126(4):755–66.
353. Reynaud D, Demarco IA, Reddy KL, Schjerven H, Bertolino E, Chen Z, et al. Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nat Immunol*. 2008 Aug;9(8):927–36.
354. Dias S, Månsson R, Gurbuxani S, Sigvardsson M, Kee BL. E2A proteins promote development of lymphoid-primed multipotent progenitors. *Immunity*. 2008 Aug 15;29(2):217–27.
355. Greig KT, de Graaf CA, Murphy JM, Carpinelli MR, Pang SHM, Frampton J, et al. Critical roles for c-Myb in lymphoid priming and early B-cell development. *Blood*. 2010 Apr 8;115(14):2796–805.
356. Stehling-Sun S, Dade J, Nutt SL, DeKoter RP, Camargo FD. Regulation of lymphoid versus myeloid fate “choice” by the transcription factor Mef2c. *Nat Immunol*. 2009 Mar;10(3):289–96.
357. Carotta S, Dakic A, D’Amico A, Pang SHM, Greig KT, Nutt SL, et al. The transcription factor PU.1 controls dendritic cell development and Flt3 cytokine receptor expression in a dose-dependent manner. *Immunity*. 2010 May 28;32(5):628–41.
358. Rosenbauer F, Wagner K, Kutok JL, Iwasaki H, Le Beau MM, Okuno Y, et al. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nat Genet*. 2004 Jun;36(6):624–30.
359. Rosenbauer F, Owens BM, Yu L, Tumang JR, Steidl U, Kutok JL, et al. Lymphoid cell growth and transformation are suppressed by a key regulatory element of the gene encoding PU.1. *Nat Genet*. 2006 Jan;38(1):27–37.
360. Anderson MK, Weiss AH, Hernandez-Hoyos G, Dionne CJ, Rothenberg EV. Constitutive expression of PU.1 in fetal hematopoietic progenitors blocks T cell development at the pro-T cell stage. *Immunity*. 2002 Feb;16(2):285–96.
361. Okuno Y, Huang G, Rosenbauer F, Evans EK, Radomska HS, Iwasaki H, et al. Potential autoregulation of transcription factor PU.1 by an upstream regulatory element. *Mol Cell Biol*. 2005 Apr;25(7):2832–45.
362. Dakic A, Metcalf D, Di Rago L, Mifsud S, Wu L, Nutt SL. PU.1 regulates the commitment of adult hematopoietic progenitors and restricts granulopoiesis. *J Exp Med*. 2005 May 2;201(9):1487–502.
363. Iwasaki H, Somoza C, Shigematsu H, Duprez EA, Iwasaki-Arai J, Mizuno S, et al. Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood*. 2005 Sep 1;106(5):1590–600.
364. Staber PB, Zhang P, Ye M, Welner RS, Levantini E, Di Ruscio A, et al. The Runx-PU.1 pathway preserves normal and AML/ETO9a leukemic stem cells. *Blood*. 2014 Oct 9;124(15):2391–9.
365. Zhang DE, Hetherington CJ, Chen HM, Tenen DG. The macrophage transcription factor PU.1 directs tissue-specific expression of the macrophage colony-stimulating factor receptor. *Mol Cell Biol*. 1994 Jan;14(1):373–81.
366. DeKoter RP, Walsh JC, Singh H. PU.1 regulates both cytokine-dependent proliferation and differentiation of granulocyte/macrophage progenitors. *EMBO J*. 1998 Aug 3;17(15):4456–68.
367. Medina KL, Pongubala JMR, Reddy KL, Lancki DW, Dekoter R, Kieslinger M, et al. Assembling a gene regulatory network for specification of the B cell fate. *Dev Cell*. 2004 Oct;7(4):607–17.
368. Pang SHM, de Graaf CA, Hilton DJ, Huntington ND, Carotta S, Wu L, et al. PU.1 Is Required for the Developmental Progression of Multipotent Progenitors to Common Lymphoid

Progenitors. *Front Immunol.* 2018;9:1264.

369. Polli M, Dakic A, Light A, Wu L, Tarlinton DM, Nutt SL. The development of functional B lymphocytes in conditional PU.1 knock-out mice. *Blood.* 2005 Sep 15;106(6):2083–90.
370. Ye M, Ermakova O, Graf T. PU.1 is not strictly required for B cell development and its absence induces a B-2 to B-1 cell switch. *J Exp Med.* 2005 Nov 21;202(10):1411–22.
371. Willis SN, Tellier J, Liao Y, Trezise S, Light A, O'Donnell K, et al. Environmental sensing by mature B cells is controlled by the transcription factors PU.1 and SpiB. *Nat Commun.* 2017 10;8(1):1426.
372. Sokalski KM, Li SKH, Welch I, Cadieux-Pitre H-AT, Gruca MR, DeKoter RP. Deletion of genes encoding PU.1 and Spi-B in B cells impairs differentiation and induces pre-B cell acute lymphoblastic leukemia. *Blood.* 2011 Sep 8;118(10):2801–8.
373. Garrett-Sinha LA, Su GH, Rao S, Kabak S, Hao Z, Clark MR, et al. PU.1 and Spi-B are required for normal B cell receptor-mediated signal transduction. *Immunity.* 1999 Apr;10(4):399–408.
374. Garrett-Sinha LA, Dahl R, Rao S, Barton KP, Simon MC. PU.1 exhibits partial functional redundancy with Spi-B, but not with Ets-1 or Elf-1. *Blood.* 2001 May 1;97(9):2908–12.
375. Eisenbeis CF, Singh H, Storb U. Pip, a novel IRF family member, is a lymphoid-specific, PU.1-dependent transcriptional activator. *Genes Dev.* 1995 Jun 1;9(11):1377–87.
376. Pang SHM, Minnich M, Gangatirkar P, Zheng Z, Ebert A, Song G, et al. PU.1 cooperates with IRF4 and IRF8 to suppress pre-B-cell leukemia. *Leukemia.* 2016;30(6):1375–87.
377. Vigorito E, Perks KL, Abreu-Goodger C, Bunting S, Xiang Z, Kohlhaas S, et al. microRNA-155 regulates the generation of immunoglobulin class-switched plasma cells. *Immunity.* 2007 Dec;27(6):847–59.
378. Lu D, Nakagawa R, Lazzaro S, Staudacher P, Abreu-Goodger C, Henley T, et al. The miR-155-PU.1 axis acts on Pax5 to enable efficient terminal B cell differentiation. *J Exp Med.* 2014 Oct 20;211(11):2183–98.
379. Pongubala JM, Nagulapalli S, Klemsz MJ, McKercher SR, Maki RA, Atchison ML. PU.1 recruits a second nuclear factor to a site important for immunoglobulin kappa 3' enhancer activity. *Mol Cell Biol.* 1992 Jan;12(1):368–78.
380. Yee AA, Yin P, Siderovski DP, Mak TW, Litchfield DW, Arrowsmith CH. Cooperative interaction between the DNA-binding domains of PU.1 and IRF4. *J Mol Biol.* 1998 Jun 26;279(5):1075–83.
381. McKercher SR, Lombardo CR, Bobkov A, Jia X, Assa-Munt N. Identification of a PU.1-IRF4 protein interaction surface predicted by chemical exchange line broadening. *Proc Natl Acad Sci U S A.* 2003 Jan 21;100(2):511–6.
382. Escalante CR, Brass AL, Pongubala JMR, Shatova E, Shen L, Singh H, et al. Crystal structure of PU.1/IRF-4/DNA ternary complex. *Mol Cell.* 2002 Nov;10(5):1097–105.
383. Iwafuchi-Doi M, Zaret KS. Pioneer transcription factors in cell reprogramming. *Genes Dev.* 2014 Dec 15;28(24):2679–92.
384. Minderjahn J, Schmidt A, Fuchs A, Schill R, Raithel J, Babina M, et al. Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1. *Nat Commun.* 2020 Dec;11(1):402.
385. Rao G, Rekhtman N, Cheng G, Krasikov T, Skoultchi AI. Deregulated expression of the PU.1 transcription factor blocks murine erythroleukemia cell terminal differentiation. *Oncogene.* 1997 Jan 9;14(1):123–31.
386. Hong W, Kim AY, Ky S, Rakowski C, Seo S-B, Chakravarti D, et al. Inhibition of CBP-mediated protein acetylation by the Ets family oncoprotein PU.1. *Mol Cell Biol.* 2002

Jun;22(11):3729–43.

387. Cook WD, McCaw BJ, Herring C, John DL, Foote SJ, Nutt SL, et al. PU.1 is a suppressor of myeloid leukemia, inactivated in mice by gene deletion and mutation of its DNA binding domain. *Blood*. 2004 Dec 1;104(12):3437–44.
388. Walter MJ, Park JS, Ries RE, Lau SKM, McLellan M, Jaeger S, et al. Reduced PU.1 expression causes myeloid progenitor expansion and increased leukemia penetrance in mice expressing PML-RAR. *Proc Natl Acad Sci*. 2005 Aug 30;102(35):12513–8.
389. Steidl U, Rosenbauer F, Verhaak RGW, Gu X, Ebralidze A, Otu HH, et al. Essential role of Jun family transcription factors in PU.1 knockdown-induced leukemic stem cells. *Nat Genet*. 2006 Nov;38(11):1269–77.
390. Mueller BU, Pabst T, Osato M, Asou N, Johansen LM, Minden MD, et al. Heterozygous PU.1 mutations are associated with acute myeloid leukemia. *Blood*. 2002 Aug 1;100(3):998–1007.
391. Lamandin C, Sagot C, Roumier C, Lepelley P, De Botton S, Cosson A, et al. Are PU.1 mutations frequent genetic events in acute myeloid leukemia (AML)? *Blood*. 2002 Dec 15;100(13):4680–1.
392. Döhner K, Tobis K, Bischof T, Hein S, Schlenk RF, Fröhling S, et al. Mutation analysis of the transcription factor PU.1 in younger adults (16 to 60 years) with acute myeloid leukemia: a study of the AML Study Group Ulm (AMLSG ULM). *Blood*. 2003 Nov 15;102(10):3850; author reply 3850-3851.
393. Lavallée V-P, Baccelli I, Kros J, Wilhelm B, Barabé F, Gendron P, et al. The transcriptomic landscape and directed chemical interrogation of MLL-rearranged acute myeloid leukemias. *Nat Genet*. 2015 Sep;47(9):1030–7.
394. Vangala RK, Heiss-Neumann MS, Rangatia JS, Singh SM, Schoch C, Tenen DG, et al. The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. *Blood*. 2003 Jan 1;101(1):270–7.
395. Seki M, Kimura S, Isobe T, Yoshida K, Ueno H, Nakajima-Takagi Y, et al. Recurrent SPI1 (PU.1) fusions in high-risk pediatric T cell acute lymphoblastic leukemia. *Nat Genet*. 2017 Aug;49(8):1274–81.
396. Anderson JR, Armitage JO, Weisenburger DD. Epidemiology of the non-Hodgkin's lymphomas: distributions of the major subtypes differ by geographic locations. Non-Hodgkin's Lymphoma Classification Project. *Ann Oncol Off J Eur Soc Med Oncol*. 1998 Jul;9(7):717–20.
397. Talaulikar D, Shadbolt B, Dahlstrom JE, McDonald A. Routine use of ancillary investigations in staging diffuse large B-cell lymphoma improves the International Prognostic Index (IPI). *J Hematol Oncol* *J Hematol Oncol*. 2009 Nov 22;2:49.
398. Talaulikar D, Dahlstrom JE, Shadbolt B, McNiven M, Broomfield A, Pidcock M. Occult bone marrow involvement in patients with diffuse large B-cell lymphoma: results of a pilot study. *Pathology (Phila)*. 2007 Dec;39(6):580–5.
399. Hu S, Xu-Monette ZY, Balasubramanyam A, Manyam GC, Visco C, Tzankov A, et al. CD30 expression defines a novel subgroup of diffuse large B-cell lymphoma with favorable prognosis and distinct gene expression signature: a report from the International DLBCL Rituximab-CHOP Consortium Program Study. *Blood*. 2013 Apr 4;121(14):2715–24.
400. Hasselblom S, Ridell B, Sigurdardottir M, Hansson U, Nilsson-Ehle H, Andersson P-O. Low rather than high Ki-67 protein expression is an adverse prognostic factor in diffuse large B-cell lymphoma. *Leuk Lymphoma*. 2008 Aug;49(8):1501–9.
401. Armitage JO, Weisenburger DD. New approach to classifying non-Hodgkin's lymphomas: clinical features of the major histologic subtypes. Non-Hodgkin's Lymphoma Classification Project. *J Clin Oncol Off J Am Soc Clin Oncol*. 1998 Aug;16(8):2780–95.
402. Coiffier B, Thieblemont C, Van Den Neste E, Lepage G, Plantier I, Castaigne S, et al.

- Long-term outcome of patients in the LNH-98.5 trial, the first randomized study comparing rituximab-CHOP to standard CHOP chemotherapy in DLBCL patients: a study by the Groupe d'Etudes des Lymphomes de l'Adulte. *Blood*. 2010 Sep 23;116(12):2040–5.
403. Shaffer AL, Young RM, Staudt LM. Pathogenesis of human B cell lymphomas. *Annu Rev Immunol*. 2012;30:565–610.
404. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000 Feb 3;403(6769):503–11.
405. Gutiérrez-García G, Cardesa-Salzmann T, Climent F, González-Barca E, Mercadal S, Mate JL, et al. Gene-expression profiling and not immunophenotypic algorithms predicts prognosis in patients with diffuse large B-cell lymphoma treated with immunochemotherapy. *Blood*. 2011 May 5;117(18):4836–43.
406. Chapuy B, Stewart C, Dunford AJ, Kim J, Kamburov A, Redd RA, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med*. 2018;24(5):679–90.
407. Wright GW, Huang DW, Phelan JD, Coulibaly ZA, Roulland S, Young RM, et al. A Probabilistic Classification Tool for Genetic Subtypes of Diffuse Large B Cell Lymphoma with Therapeutic Implications. *Cancer Cell*. 2020 Apr;37(4):551-568.e14.
408. Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature*. 2011 Jul 27;476(7360):298–303.
409. Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet*. 2011 Jul 31;43(9):830–7.
410. Goodman RH, Smolik S. CBP/p300 in cell growth, transformation, and development. *Genes Dev*. 2000 Jul 1;14(13):1553–77.
411. Bannister AJ, Kouzarides T. The CBP co-activator is a histone acetyltransferase. *Nature*. 1996 Dec 19;384(6610):641–3.
412. Ogryzko VV, Schiltz RL, Russanova V, Howard BH, Nakatani Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell*. 1996 Nov 29;87(5):953–9.
413. Lill NL, Grossman SR, Ginsberg D, DeCaprio J, Livingston DM. Binding and modulation of p53 by p300/CBP coactivators. *Nature*. 1997 Jun 19;387(6635):823–7.
414. Avantaggiati ML, Ogryzko V, Gardner K, Giordano A, Levine AS, Kelly K. Recruitment of p300/CBP in p53-dependent signal pathways. *Cell*. 1997 Jun 27;89(7):1175–84.
415. Bereshchenko OR, Gu W, Dalla-Favera R. Acetylation inactivates the transcriptional repressor BCL6. *Nat Genet*. 2002 Dec;32(4):606–13.
416. Shilatifard A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem*. 2012;81:65–95.
417. Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet*. 2010 Feb;42(2):181–5.
418. Sneeringer CJ, Scott MP, Kuntz KW, Knutson SK, Pollock RM, Richon VM, et al. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proc Natl Acad Sci U S A*. 2010 Dec 7;107(49):20980–5.
419. Davis RE, Brown KD, Siebenlist U, Staudt LM. Constitutive nuclear factor kappaB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells. *J Exp Med*. 2001 Dec 17;194(12):1861–74.
420. Lam KP, Kühn R, Rajewsky K. In vivo ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death. *Cell*. 1997 Sep 19;90(6):1073–

83.

421. Yang Y, Shaffer AL, Emre NCT, Ceribelli M, Zhang M, Wright G, et al. Exploiting synthetic lethality for the therapy of ABC diffuse large B cell lymphoma. *Cancer Cell*. 2012 Jun 12;21(6):723–37.
422. Compagno M, Lim WK, Grunn A, Nandula SV, Brahmachary M, Shen Q, et al. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature*. 2009 Jun 4;459(7247):717–21.
423. Kato M, Sanada M, Kato I, Sato Y, Takita J, Takeuchi K, et al. Frequent inactivation of A20 in B-cell lymphomas. *Nature*. 2009 Jun 4;459(7247):712–6.
424. Migliazza A, Martinotti S, Chen W, Fusco C, Ye BH, Knowles DM, et al. Frequent somatic hypermutation of the 5' noncoding region of the BCL6 gene in B-cell lymphoma. *Proc Natl Acad Sci U S A*. 1995 Dec 19;92(26):12520–4.
425. Pasqualucci L, Migliazza A, Fracchiolla N, William C, Neri A, Baldini L, et al. BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proc Natl Acad Sci U S A*. 1998 Sep 29;95(20):11816–21.
426. Shen HM, Peters A, Baron B, Zhu X, Storb U. Mutation of BCL-6 gene in normal B cells by the process of somatic hypermutation of Ig genes. *Science*. 1998 Jun 12;280(5370):1750–2.
427. Mandelbaum J, Bhagat G, Tang H, Mo T, Brahmachary M, Shen Q, et al. BLIMP1 is a tumor suppressor gene frequently disrupted in activated B cell-like diffuse large B cell lymphoma. *Cancer Cell*. 2010 Dec 14;18(6):568–79.
428. Pasqualucci L, Compagno M, Houldsworth J, Monti S, Grunn A, Nandula SV, et al. Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. *J Exp Med*. 2006 Feb 20;203(2):311–7.
429. Owen RG, Treon SP, Al-Katib A, Fonseca R, Greipp PR, McMaster ML, et al. Clinicopathological definition of Waldenstrom's macroglobulinemia: consensus panel recommendations from the Second International Workshop on Waldenstrom's Macroglobulinemia. *Semin Oncol*. 2003 Apr;30(2):110–5.
430. Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016 19;127(20):2375–90.
431. Remstein ED, Hanson CA, Kyle RA, Hodnefield JM, Kurtin PJ. Despite apparent morphologic and immunophenotypic heterogeneity, Waldenstrom's macroglobulinemia is consistently composed of cells along a morphologic continuum of small lymphocytes, plasmacytoid lymphocytes, and plasma cells. *Semin Oncol*. 2003 Apr;30(2):182–6.
432. Tournilhac O, Santos DD, Xu L, Kutok J, Tai Y-T, Le Gouill S, et al. Mast cells in Waldenstrom's macroglobulinemia support lymphoplasmacytic cell growth through CD154/CD40 signaling. *Ann Oncol Off J Eur Soc Med Oncol*. 2006 Aug;17(8):1275–82.
433. Pedreira CE, Costa ES, Lecrevisse Q, van Dongen JJM, Orfao A, EuroFlow Consortium. Overview of clinical flow cytometry data analysis: recent advances and future challenges. *Trends Biotechnol*. 2013 Jul;31(7):415–25.
434. San Miguel JF, Vidriales MB, Ocio E, Mateo G, Sánchez-Guijo F, Sánchez ML, et al. Immunophenotypic analysis of Waldenstrom's macroglobulinemia. *Semin Oncol*. 2003 Apr;30(2):187–95.
435. Lin P, Mansoor A, Bueso-Ramos C, Hao S, Lai R, Medeiros LJ. Diffuse large B-cell lymphoma occurring in patients with lymphoplasmacytic lymphoma/Waldenström macroglobulinemia. Clinicopathologic features of 12 cases. *Am J Clin Pathol*. 2003 Aug;120(2):246–53.
436. Castillo JJ, Gustine J, Meid K, Dubeau T, Hunter ZR, Treon SP. Histological transformation to diffuse large B-cell lymphoma in patients with Waldenström

- macroglobulinemia. *Am J Hematol*. 2016;91(10):1032–5.
437. Dimopoulos MA, Zervas C, Zomas A, Hamilos G, Gika D, Efstathiou E, et al. Extended rituximab therapy for previously untreated patients with Waldenström’s macroglobulinemia. *Clin Lymphoma*. 2002 Dec;3(3):163–6.
438. Leblond V, Kastritis E, Advani R, Ansell SM, Buske C, Castillo JJ, et al. Treatment recommendations from the Eighth International Workshop on Waldenström’s Macroglobulinemia. *Blood*. 2016 08;128(10):1321–8.
439. Gachard N, Parrens M, Soubeyran I, Petit B, Marfak A, Rizzo D, et al. IGHV gene features and MYD88 L265P mutation separate the three marginal zone lymphoma entities and Waldenström macroglobulinemia/lymphoplasmacytic lymphomas. *Leukemia*. 2013 Jan;27(1):183–9.
440. Martín-Jiménez P, García-Sanz R, Balanzategui A, Alcoceba M, Ocio E, Sanchez ML, et al. Molecular characterization of heavy chain immunoglobulin gene rearrangements in Waldenström’s macroglobulinemia and IgM monoclonal gammopathy of undetermined significance. *Haematologica*. 2007 May;92(5):635–42.
441. Walsh SH, Laurell A, Sundström G, Roos G, Sundström C, Rosenquist R. Lymphoplasmacytic lymphoma/Waldenström’s macroglobulinemia derives from an extensively hypermutated B cell that lacks ongoing somatic hypermutation. *Leuk Res*. 2005 Jul;29(7):729–34.
442. Kriangkum J, Taylor BJ, Treon SP, Mant MJ, Belch AR, Pilarski LM. Clonotypic IgM V/D/J sequence analysis in Waldenstrom macroglobulinemia suggests an unusual B-cell origin and an expansion of polyclonal B cells in peripheral blood. *Blood*. 2004 Oct 1;104(7):2134–42.
443. Martín-Jiménez P, García-Sanz R, Sarasquete ME, Ocio E, Pérez JJ, González M, et al. Functional class switch recombination may occur “in vivo” in Waldenström macroglobulinaemia. *Br J Haematol*. 2007 Jan;136(1):114–6.
444. Sahota SS, Forconi F, Ottensmeier CH, Provan D, Oscier DG, Hamblin TJ, et al. Typical Waldenstrom macroglobulinemia is derived from a B-cell arrested after cessation of somatic mutation but prior to isotype switch events. *Blood*. 2002 Aug 15;100(4):1505–7.
445. Nardini E, Rizzi S, Capello D, Vitolo U, Gaidano G, Menard S, et al. Most immunoglobulin heavy chain switch mu rearrangements in B-cell chronic lymphocytic leukemia are internal deletions. *FEBS Lett*. 2002 May 8;518(1–3):119–23.
446. Kriangkum J, Taylor BJ, Strachan E, Mant MJ, Reiman T, Belch AR, et al. Impaired class switch recombination (CSR) in Waldenstrom macroglobulinemia (WM) despite apparently normal CSR machinery. *Blood*. 2006 Apr 1;107(7):2920–7.
447. Mohamed AN, Bentley G, Bonnett ML, Zonder J, Al-Katib A. Chromosome aberrations in a series of 120 multiple myeloma cases with abnormal karyotypes. *Am J Hematol*. 2007 Dec;82(12):1080–7.
448. Döhner H, Stilgenbauer S, Benner A, Leupolt E, Kröber A, Bullinger L, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. *N Engl J Med*. 2000 Dec 28;343(26):1910–6.
449. Braggio E, Keats JJ, Leleu X, Van Wier S, Jimenez-Zepeda VH, Valdez R, et al. Identification of copy number abnormalities and inactivating mutations in two negative regulators of nuclear factor-kappaB signaling pathways in Waldenstrom’s macroglobulinemia. *Cancer Res*. 2009 Apr 15;69(8):3579–88.
450. Terré C, Nguyen-Khac F, Barin C, Mozziconacci MJ, Eclache V, Léonard C, et al. Trisomy 4, a new chromosomal abnormality in Waldenström’s macroglobulinemia: a study of 39 cases. *Leukemia*. 2006 Sep;20(9):1634–6.
451. Nguyen-Khac F, Lambert J, Chapiro E, Grelier A, Mould S, Barin C, et al. Chromosomal aberrations and their prognostic value in a series of 174 untreated patients with

- Waldenström's macroglobulinemia. *Haematologica*. 2013 Apr;98(4):649–54.
452. Lin S-C, Lo Y-C, Wu H. Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling. *Nature*. 2010 Jun 17;465(7300):885–90.
453. Poulain S, Roumier C, Decambron A, Renneville A, Herbaux C, Bertrand E, et al. MYD88 L265P mutation in Waldenström macroglobulinemia. *Blood*. 2013 May 30;121(22):4504–11.
454. Hunter ZR, Xu L, Yang G, Zhou Y, Liu X, Cao Y, et al. The genomic landscape of Waldenström macroglobulinemia is characterized by highly recurring MYD88 and WHIM-like CXCR4 mutations, and small somatic deletions associated with B-cell lymphomagenesis. *Blood*. 2014 Mar 13;123(11):1637–46.
455. Landgren O, Staudt L. MYD88 L265P somatic mutation in IgM MGUS. *N Engl J Med*. 2012 06;367(23):2255–6; author reply 2256–2257.
456. Wang EL, Qian ZR, Nakasono M, Tanahashi T, Yoshimoto K, Bando Y, et al. High expression of Toll-like receptor 4/myeloid differentiation factor 88 signals correlates with poor prognosis in colorectal cancer. *Br J Cancer*. 2010 Mar 2;102(5):908–15.
457. Zhu Y, Huang J-M, Zhang G-N, Zha X, Deng B-F. Prognostic significance of MyD88 expression by human epithelial ovarian carcinoma cells. *J Transl Med*. 2012 Apr 25;10:77.
458. Liang B, Chen R, Wang T, Cao L, Liu Y, Yin F, et al. Myeloid differentiation factor 88 promotes growth and metastasis of human hepatocellular carcinoma. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2013 Jun 1;19(11):2905–16.
459. Avbelj M, Wolz O-O, Fekonja O, Benčina M, Repič M, Mavri J, et al. Activation of lymphoma-associated MyD88 mutations via allosteric-induced TIR-domain oligomerization. *Blood*. 2014 Dec 18;124(26):3896–904.
460. Ansell SM, Hodge LS, Secreto FJ, Manske M, Braggio E, Price-Troska T, et al. Activation of TAK1 by MYD88 L265P drives malignant B-cell Growth in non-Hodgkin lymphoma. *Blood Cancer J*. 2014 Feb 14;4:e183.
461. Wang JQ, Jeelall YS, Beutler B, Horikawa K, Goodnow CC. Consequences of the recurrent MYD88(L265P) somatic mutation for B cell tolerance. *J Exp Med*. 2014 Mar 10;211(3):413–26.
462. Knittel G, Liedgens P, Korovkina D, Seeger JM, Al-Baldawi Y, Al-Maarri M, et al. B-cell-specific conditional expression of Myd88p.L252P leads to the development of diffuse large B-cell lymphoma in mice. *Blood*. 2016 02;127(22):2732–41.
463. Sewastianik T, Guerrera ML, Adler K, Dennis PS, Wright K, Shanmugam V, et al. Human MYD88L265P is insufficient by itself to drive neoplastic transformation in mature mouse B cells. *Blood Adv*. 2019 Nov 12;3(21):3360–74.
464. Xu L, Hunter ZR, Tsakmaklis N, Cao Y, Yang G, Chen J, et al. Clonal architecture of CXCR4 WHIM-like mutations in Waldenström Macroglobulinaemia. *Br J Haematol*. 2016 Mar;172(5):735–44.
465. Fecteau JF, Côté G, Néron S. A new memory CD27-IgG+ B cell population in peripheral blood expressing VH genes with low frequency of somatic mutation. *J Immunol Baltim Md 1950*. 2006 Sep 15;177(6):3728–36.
466. Wu Y-CB, Kipling D, Dunn-Walters DK. The relationship between CD27 negative and positive B cell populations in human peripheral blood. *Front Immunol*. 2011;2:81.
467. Colonna-Romano G, Bulati M, Aquino A, Pellicanò M, Vitello S, Lio D, et al. A double-negative (IgD-CD27-) B cell population is increased in the peripheral blood of elderly people. *Mech Ageing Dev*. 2009 Oct;130(10):681–90.
468. Jan M, Snyder TM, Corces-Zimmerman MR, Vyas P, Weissman IL, Quake SR, et al. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Sci Transl Med*. 2012 Aug 29;4(149):149ra118.
469. Orlacchio A, Ranieri M, Brave M, Arciuch VA, Forde T, De Martino D, et al. SGK1 Is

- a Critical Component of an AKT-Independent Pathway Essential for PI3K-Mediated Tumor Development and Maintenance. *Cancer Res.* 2017 15;77(24):6914–26.
470. Lu E, Wolfreys FD, Muppidi JR, Xu Y, Cyster JG. S-Geranylgeranyl-L-glutathione is a ligand for human B cell-confinement receptor P2RY8. *Nature.* 2019;567(7747):244–8.
471. Shenoy N, Bhagat T, Nieves E, Stenson M, Lawson J, Choudhary GS, et al. Upregulation of TET activity with ascorbic acid induces epigenetic modulation of lymphoma cells. *Blood Cancer J.* 2017 Jul;7(7):e587–e587.
472. Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, et al. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell.* 2011 Jul 8;146(1):67–79.
473. Guo JU, Su Y, Zhong C, Ming G, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell.* 2011 Apr 29;145(3):423–34.
474. Rosikiewicz W, Chen X, Dominguez PM, Ghamlouch H, Aoufouchi S, Bernard OA, et al. TET2 deficiency reprograms the germinal center B cell epigenome and silences genes linked to lymphomagenesis. *Sci Adv.* 2020 Jun;6(25):eaay5872.
475. Zhou Y, Liu X, Xu L, Hunter ZR, Cao Y, Yang G, et al. Transcriptional repression of plasma cell differentiation is orchestrated by aberrant over-expression of the ETS factor SPIB in Waldenström macroglobulinaemia. *Br J Haematol.* 2014 Sep;166(5):677–89.
476. Delgado MD, Gutiérrez P, Richard C, Cuadrado MA, Moreau-Gachelin F, León J. Spi-1/PU.1 proto-oncogene induces opposite effects on monocytic and erythroid differentiation of K562 cells. *Biochem Biophys Res Commun.* 1998 Nov 18;252(2):383–91.
477. Staber PB, Zhang P, Ye M, Welner RS, Nombela-Arrieta C, Bach C, et al. Sustained PU.1 levels balance cell-cycle regulators to prevent exhaustion of adult hematopoietic stem cells. *Mol Cell.* 2013 Mar 7;49(5):934–46.
478. Lu X, Levine R, Tong W, Wernig G, Pikman Y, Zarnegar S, et al. Expression of a homodimeric type I cytokine receptor is required for JAK2V617F-mediated transformation. *Proc Natl Acad Sci.* 2005 Dec 27;102(52):18962–7.
479. Argyropoulos KV, Vogel R, Ziegler C, Altan-Bonnet G, Velardi E, Calafiore M, et al. Clonal B cells in Waldenström’s macroglobulinemia exhibit functional features of chronic active B-cell receptor signaling. *Leukemia.* 2016;30(5):1116–25.
480. Phelan JD, Young RM, Webster DE, Roulland S, Wright GW, Kasbekar M, et al. A multiprotein supercomplex controlling oncogenic signalling in lymphoma. *Nature.* 2018;560(7718):387–91.
481. Ouk C, Roland L, Saintamand A, Gachard N, Thomas M, Devéza M, et al. B-cell enforced expression of the mouse ortholog of MYD88^{L265P} is responsible for Waldenström-like B-cell lymphoma. *bioRxiv.* 2019 Jan 1;794024.
482. Kyle RA, Therneau TM, Rajkumar SV, Remstein ED, Offord JR, Larson DR, et al. Long-term follow-up of IgM monoclonal gammopathy of undetermined significance. *Blood.* 2003 Nov 15;102(10):3759–64.
483. Hu C-J, Rao S, Ramirez-Bergeron DL, Garrett-Sinha LA, Gerondakis S, Clark MR, et al. PU.1/Spi-B Regulation of c-rel Is Essential for Mature B Cell Survival. *Immunity.* 2001 Oct;15(4):545–55.
484. Bekeredjian-Ding I, Jengo G. Toll-like receptors - sentries in the B-cell response. *Immunology.* 2009 Nov;128(3):311–23.
485. Vaure C, Liu Y. A Comparative Review of Toll-Like Receptor 4 Expression and Functionality in Different Animal Species. *Front Immunol [Internet].* 2014 Jul 10 [cited 2020 Aug 15];5. Available from: <http://journal.frontiersin.org/article/10.3389/fimmu.2014.00316/abstract>
486. Wang L, Zhao Y, Qian J, Sun L, Lu Y, Li H, et al. Toll-like receptor-4 signaling in mantle cell lymphoma: effects on tumor growth and immune evasion. *Cancer.* 2013 Feb

15;119(4):782–91.

487. Xu Y, Zhao Y, Huang H, Chen G, Wu X, Wang Y, et al. Expression and function of toll-like receptors in multiple myeloma patients: toll-like receptor ligands promote multiple myeloma cell growth and survival via activation of nuclear factor- κ B: Expression and Function of TLRs in MM Patients. *Br J Haematol*. 2010 Sep;150(5):543–53.
488. Shrimpton J, Care MA, Carmichael J, Walker K, Evans P, Evans C, et al. TLR-mediated activation of Waldenström macroglobulinemia B cells reveals an uncoupling from plasma cell differentiation. *Blood Adv*. 2020 Jun 23;4(12):2821–36.
489. Rehli M, Poltorak A, Schwarzfischer L, Krause SW, Andreesen R, Beutler B. PU.1 and interferon consensus sequence-binding protein regulate the myeloid expression of the human Toll-like receptor 4 gene. *J Biol Chem*. 2000 Mar 31;275(13):9773–81.
490. Ganley-Leal LM, Liang Y, Jagannathan-Bogdan M, Farraye FA, Nikolajczyk BS. Differential regulation of TLR4 expression in human B cells and monocytes. *Mol Immunol*. 2010 Nov;48(1–3):82–8.
491. Tsukamoto H, Takeuchi S, Kubota K, Kobayashi Y, Kozakai S, Ukai I, et al. Lipopolysaccharide (LPS)-binding protein stimulates CD14-dependent Toll-like receptor 4 internalization and LPS-induced TBK1-IKK ϵ -IRF3 axis activation. *J Biol Chem*. 2018 29;293(26):10186–201.
492. Puente XS, Beà S, Valdés-Mas R, Villamor N, Gutiérrez-Abril J, Martín-Subero JI, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2015 Oct 22;526(7574):519–24.
493. Hoang PH, Dobbins SE, Cornish AJ, Chubb D, Law PJ, Kaiser M, et al. Whole-genome sequencing of multiple myeloma reveals oncogenic pathways are targeted somatically through multiple mechanisms. *Leukemia*. 2018;32(11):2459–70.
494. Bergsagel PL, Kuehl WM. WSU-WM and BCWM.1 should not be assumed to represent Waldenström macroglobulinemia cell lines. *Blood*. 2008 Aug 1;112(3):917–917.
495. Drexler HG, Chen S, Macleod RAF. Would the real Waldenström cell line please stand up? *Leuk Lymphoma*. 2013 Feb;54(2):224–6.
496. Miyaoka Y, Berman JR, Cooper SB, Mayerl SJ, Chan AH, Zhang B, et al. Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Sci Rep*. 2016 Mar 31;6:23549.
497. Ye L, Wang C, Hong L, Sun N, Chen D, Chen S, et al. Programmable DNA repair with CRISPRa/i enhanced homology-directed repair efficiency with a single Cas9. *Cell Discov*. 2018;4:46.
498. Dodsworth BT, Hatje K, Meyer CA, Flynn R, Cowley SA. Rates of homology directed repair of CRISPR-Cas9 induced double strand breaks are lower in naïve compared to primed human pluripotent stem cells. *Stem Cell Res*. 2020;46:101852.
499. Hoppe PS, Schwarzfischer M, Loeffler D, Kokkaliaris KD, Hilsenbeck O, Moritz N, et al. Early myeloid lineage choice is not initiated by random PU.1 to GATA1 protein ratios. *Nature*. 2016 Jul;535(7611):299–302.

Titre : Régulation transcriptionnelle et épigénétique de la différenciation B normale et tumorale : rôle des enzymes TET du facteur de transcription SPI1

Mots clés : Epigénétique, Régulation transcriptionnelle, TET, SPI1, lymphocyte B

Résumé : L'ontogenèse des lymphocytes B (LB) comporte une première phase de différenciation, dans la moelle osseuse en l'absence de toute stimulation antigénique spécifique, aboutissant au LB immature. La seconde phase, d'activation et de maturation finale, est dépendante des antigènes et se déroule dans les organes lymphoïdes secondaires, au sein de structures transitoires appelées centres germinatifs (CG). Elle génère des plasmocytes et des cellules B mémoires spécifiques d'un antigène.

Ce travail de thèse s'intéresse à différents acteurs impliqués dans la régulation épigénétique et transcriptionnelle de la différenciation lymphoïde B terminale : les enzymes TET2 et TET3 et le facteur de transcription (FT) SPI1/PU.1. Des mutations affectant les gènes codant pour ces protéines sont trouvées dans les hémopathies chez l'Homme et nous avons cherché à déterminer leurs conséquences fonctionnelles en utilisant des modélisations *in vivo* et *in vitro*.

J'ai d'une part analysé l'impact de la perte de fonction de TET2 sur la différenciation et la maturation des cellules B. Les résultats montrent un blocage de la différenciation plasmocytaire associé à une hyperplasie des CG et une augmentation du pourcentage et du nombre absolu des LB du CG (BGC). L'analyse par PCR quantitative de l'expression des FT importants pour la différenciation des BGC et des plasmocytes a montré que les cellules déficientes pour TET2 présentent une répression du gène *Prdm1* codant pour BLIMP1, un régulateur essentiel de la différenciation plasmocytaire. Je me suis ensuite intéressée à TET3, autre protéine de la famille TET exprimée dans la lignée B. Les modèles *Tet3*-déficients *in vivo* et *in vitro* n'ont pas montré d'altération marquée de la différenciation B terminale.

J'ai par ailleurs étudié une mutation somatique de *SPI1/PU.1*, identifiée par notre équipe chez des patients atteints de maladie de Waldenström (MW). Dans plus de 95% des cas, la mutation activatrice L265P du gène *MYD88* est également présente. Nous avons montré que la mutation de *SPI1*, bien que n'empêchant pas sa liaison à l'ADN, modifie son affinité de liaison sur les sites normalement reconnus par la forme sauvage. La mutation semble faire adopter à cette protéine ETS de classe III un comportement qui ressemble à celui d'une classe I/IIa. J'ai ensuite cherché à documenter les bases de la coopération oncogénique entre SPI1 et MYD88 de deux façons. La première, en étudiant la prolifération et la différenciation de lymphocytes B naïfs issus d'un modèle murin knock-in pour la mutation *SPI1* développé dans l'équipe, transduits avec un rétrovirus apportant la mutation de *MYD88*. Les résultats montrent une augmentation de la prolifération dans la condition double mutante ainsi qu'une augmentation de la différenciation terminale. La seconde approche consiste à modifier la lignée humaine BCWM1 de MW par CRISPR/Cas9 afin d'y introduire la mutation de *SPI1* en même temps que l'expression de la GFP. Ce modèle sera notamment utilisé pour réaliser des expériences de ChIP-seq afin d'identifier les cibles de la protéine mutante dans un contexte MW-like.

En conclusion, le respect des programmes transcriptionnels est essentiel pour le bon déroulement de la différenciation B terminale et peut être impacté soit directement, par des mutations affectant des FT comme SPI1, soit indirectement lorsque le profil de méthylation de gènes codant pour des FT (*PRDM1*) est altéré suite à des mutations affectant des enzymes comme TET2.

Title : Transcriptional and Epigenetic Regulation of Normal and Malignant B-cell Differentiation: Role of TET Enzymes and Transcription Factor SPI1

Keywords : Epigenetics, Transcriptional regulation, TET, SPI1, B-cell

Abstract : B-cell development involves a first phase of differentiation in the bone marrow, in the absence of any specific antigenic stimulation, leading to immature B-cells. The second phase, staging activation and final maturation, is antigen-dependent and takes place in the secondary lymphoid organs, within transient structures called germinal centers (GC). It generates antigen-specific plasma cells and memory B cells.

This thesis work focuses on different actors involved in the epigenetic and transcriptional regulation of B-cell differentiation: the enzymes TET2 and TET3 and the transcription factor (TF) SPI1/PU.1. Mutations in genes encoding these proteins are found in human neoplasms, we used *in vivo* and *in vitro* models to determine their functional consequences.

I analyzed the impact of TET2 loss of function on the differentiation and maturation of B-cells. The results show an impaired plasma cell differentiation associated with GC hyperplasia and an increase in the percentage and absolute number of GC B-cells (BGC). Quantitative PCR analysis of the expression of key BGC and plasma cell TF showed that *Tet2*-deficient cells exhibit repression of the *Prdm1* gene encoding BLIMP1, a master regulator of plasma cell differentiation. I then turned my attention to TET3, another TET family protein expressed in the B-cell lineage. *In vivo* and *in vitro* *Tet3*-deficient models show that the loss of TET3 does not significantly affect terminal B differentiation.

In addition, I studied a somatic mutation of *SPI1/PU.1*, identified by our team in patients with Waldenström's disease (WM). In more than 95% of cases, the L265P activating mutation of *MYD88* is also present. We have shown that *SPI1* mutation, although not preventing its binding to DNA, alters its binding affinity at sites normally recognized by the wild-type form. The mutation appears to cause this class III ETS protein to behave in a manner similar to a class I/IIa ETS protein. I then sought to document the basis for oncogenic cooperation between *SPI1* and *MYD88*. First, by studying the proliferation and differentiation of naïve B-cells from a locally developed mouse model knock-in for the *SPI1* mutation, transduced with a retrovirus carrying the *MYD88* mutation. The results show an increase in proliferation and in terminal differentiation in the double mutant condition. Second, by modifying the human BCWM1 WM cell line by CRISPR/Cas9 in order to introduce the *SPI1* mutation at the same time as the expression of the GFP. This model will be used in particular to perform ChIP-seq experiments to identify the targets of the mutant protein in a MW-like context.

In conclusion, compliance to transcriptional programs is essential for the smooth progress of B-cell terminal differentiation and can be impacted either directly, by mutations affecting TF such as SPI1, or indirectly when the methylation profile of key TF-encoding genes (*PRDM1*) is altered following mutations in enzymes such as TET2.

