



**HAL**  
open science

# Anomalies moléculaires dans la macroglobulinémie de Waldenström : identification d'une mutation somatique récurrente dans le gène codant pour le facteur de transcription SPI1/PU.1 et description de ses conséquences fonctionnelles

Damien Roos-Weil

## ► To cite this version:

Damien Roos-Weil. Anomalies moléculaires dans la macroglobulinémie de Waldenström : identification d'une mutation somatique récurrente dans le gène codant pour le facteur de transcription SPI1/PU.1 et description de ses conséquences fonctionnelles. Cancer. Université Paris Saclay (COMUE), 2018. Français. NNT : 2018SACLS517 . tel-02420532

**HAL Id: tel-02420532**

**<https://theses.hal.science/tel-02420532>**

Submitted on 20 Dec 2019

**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.

**Anomalies moléculaires dans la macroglobulinémie de  
Waldenström : identification d'une mutation somatique  
récurrente dans le gène codant pour le facteur de transcription  
SPI1/PU.1 et description de ses conséquences fonctionnelles**

Thèse de doctorat de l'Université Paris-Saclay  
préparée à l'Université Paris-Sud

École doctorale n°582

Cancérologie, biologie, médecine et santé (CBMS)

Spécialité de doctorat: Aspects moléculaires et cellulaires de la biologie

Thèse présentée et soutenue à Villejuif, le 19/12/2018, par

**Damien Roos-Weil**

Composition du Jury :

**Alain DELMER**

PU-PH, CHU de Reims

Président

**Mary CALLANAN**

PU-PH, CHU Dijon-Bourgogne

Rapporteur

**Michel COGNE**

PU-PH, Université de Limoges (CNRS UMR 7276)

Rapporteur

**Joëlle WIELS**

DR, Université Paris-Sud (CNRS UMR 8126)

Examinatrice

**Jean-Christophe BORIES**

DR, Université Paris Diderot (INSERM UMR 1126)

Examineur

**Olivier BERNARD**

DR, Université Paris-Sud (INSERM U1170)

Directeur de thèse

## Remerciements

*Je remercie Olivier Bernard de m'avoir accueilli dans son unité et de m'avoir ouvert les portes du monde merveilleux de la génétique des hémopathies. Merci de m'avoir initié à la recherche et au raisonnement scientifique, merci d'avoir aiguisé mon sens critique et de m'avoir fait confiance pour travailler sur tant de projets passionnants. Je peux avouer maintenant que quelques semaines avant mon arrivée au laboratoire une amie m'avait offert (par hasard) un très beau livre « Un si long chemin » de Bernard Ollivier et que j'en avais tiré quelques appréhensions. Mais en réalité rien de tel ! Tout est passé très vite et fut extrêmement enrichissant.*

*Je remercie les Prs Mary Callanan et Michel Cogné d'avoir accepté d'être les rapporteurs de mon travail de thèse. Je remercie les Drs Joëlle Wiels et Jean-Christophe Bories ainsi que le Pr Alain Delmer d'avoir accepté d'examiner ce travail. Merci beaucoup pour votre disponibilité.*

*Je remercie le Pr Florence Nguyen-Khac pour son implication précieuse, son soutien et ses conseils qui ont grandement contribué à la genèse et à l'aboutissement de ce projet. Je suis très content et honoré de pouvoir poursuivre cette collaboration à l'hôpital et dans son équipe de recherche.*

*Je remercie le Pr Véronique Leblond pour la confiance qu'elle m'accorde au quotidien et l'exemple qu'elle représente pour moi dans l'approche humaine, intellectuelle et scientifique de notre métier de médecin. Merci d'avoir éveillé en moi ce goût pour cette maladie étrange, maladie passionnante, « maladie des Princes » qu'est la macroglobulinémie de Waldenström.*

*Je remercie toutes les personnes au sein du laboratoire qui, par leurs échanges et conseils, ont activement participé à ce projet. Merci notamment à Christel Guillouf, Nathalie Droin, Sébastien Malinge, Santos Susin et Thomas Mercher.*

*Je remercie chaleureusement Nabih Azar sans qui rien n'aurait été possible finalement. Merci d'avoir réalisé la quasi-totalité des prélèvements médullaires (de grande qualité) qui ont permis de faire aboutir ce travail. Merci pour ton amitié.*

*Je remercie infiniment Véro Della-Valle pour sa gentillesse, sa disponibilité, ses connaissances sans limites du monde murin, de la biologie moléculaire et cellulaire. Ce fut un immense plaisir de passer ses années à tes côtés. Merci pour ta bonne humeur et ta bienveillance, toujours présentes que ce soient dans les bons moments mais aussi ceux plus difficiles, heureusement plus rares. J'espère que nous pourrons continuer à nous voir et à échanger, ce sera pour moi un plaisir. Enfin, merci de m'avoir appris à tenir une pipette ! Je sais que ce n'était pas gagné...*



*Je remercie Cécile Lopech et Laurianne « Modesto » Scourzic, les reines du « bocal », pour leur soutien sans faille et tout le savoir scientifique qu'elles m'ont transmis. Je garderai longtemps en mémoire tous les moments de rigolade passés ensemble. Oui, il y a bien deux carotides ! J'espère vous revoir bientôt !*

*Je remercie M'Boypa Khadija Diop pour son aide et sa remarquable patience face à nos demandes d'analyses bioinformatiques toujours plus ésotériques... Merci à Diane « Queen of Bandshift » Lara-Guérin, Eric « Iceman » Durot, Cécile Hérate, Camille Decaudin et Marine Armand pour tout leur apport dans le projet SPI1. J'espère qu'elles/il ne l'ont pas ou ne le vivent pas trop mal...*

*Je remercie chaleureusement Hussein pour tout ce qu'il m'a apporté pendant les deux années que nous avons partagées au laboratoire. Je retiendrai nos discussions scientifiques toujours enrichissantes, nos échanges de points de vue sur la politique internationale, la religion et la vie en général. Merci d'avoir essayé de me faire comprendre quelque chose à l'immunologie. Le résultat n'est pas loin d'être affligeant mais la faute en revient à l'élève !*

*Je remercie affectueusement ma famille : mes grands-parents toujours au top malgré les années qui passent (vous voyez, vous y êtes à cette soutenance de thèse !) ; mes parents qui me conseillent et m'épaulent avec réserve et bienveillance depuis tant d'années (vous serez toujours un exemple pour moi) ; mes sœurs Anne, Lucille et Lise, les plus belles, les plus intelligentes et les plus drôles (jugement ne souffrant d'aucune contestation).*

*Comment ne pas à nouveau remercier mes amis Boris, Clément, Florent, Sébastien et Yuta. Merci d'être toujours là, d'avoir partagé le projet d'une vie qu'est le Moulin Gély (oh la rime riche !). Merci de ne pas être des cossards ! Les mots me manquent pour vous exprimer tout ce que je souhaiterais vous exprimer. Vous connaissez mon alexithymie...*

*Merci à tous mes amis, collègues, colocos qui me supportent et je sais que ce n'est pas facile : Florent « Bob », Marie (s), Delphine, Vincent, Julie, Juliette, JB, Rodolphe, Raphaël, Pauline, Marine (s), Marie-Pierre, Lawrence. Le problème des listes, c'est qu'elles ne sont jamais complètes. Pour ceux qui n'y sont pas, ne vous inquiétez pas, vous serez dans les remerciements de ma prochaine thèse !*

*Je remercie mes idoles, les Didiers (Wampas et Super) et les Philippes (Katerine et Risoli). Je sais tout ce que je vous dois. Des pensées spéciales pour mon compagnon de toutes les routes (mon Lounis) JB Chenet, pour la Reine de la niche et sa gin addiction, pour le grain de folie de Salut c'est cool, pour l'énergie d'Idles et la beauté de Salammbô.*

*Si vous avez déjà songé à votre épitaphe et qu'elle se conclura par "Un Cell et au lit",  
Si vous savez bien que Cancer Discovery n'est pas la dernière attraction de Disney World,  
Si on ne vous fera pas avaler que les CAR-T sont la nouvelle animation Pixar,  
Si RhoA n'est pas pour vous que le cri de ralliement guerrier de Rahan (l'incroyable fils des âges farouches),  
Vous êtes trop brillant, cette thèse n'est pas pour vous...*



## Table des matières

1.	Hématopoïèse .....	9
2.	Lymphopoïèse B normale.....	10
a.	Introduction .....	10
b.	Phase antigène-indépendante .....	12
c.	Phase antigène-dépendante .....	14
d.	Réaction post-centro-germinative .....	17
e.	Lymphomagénèse B .....	19
f.	Voies moléculaires d'activation lymphocytaire B.....	22
i.	Signalisation BCR .....	23
ii.	CD40.....	25
iii.	BAFF et APRIL.....	25
iv.	Signalisation TLR .....	26
v.	Voies de signalisation NF-kB .....	27
g.	Maintien de l'identité B .....	29
i.	MEF2B .....	30
ii.	MYC .....	30
iii.	PAX5.....	31
iv.	E2A/TCF3 .....	32
v.	BACH2 .....	33
vi.	BCL6 .....	34
vii.	OBF1 et OCT2 .....	37
viii.	PU.1, SPIB, ETS1 .....	37
h.	Acteurs de la différenciation plasmocytaire.....	38
i.	IRF4 .....	39
ii.	BLIMP1 .....	39
iii.	XBP1 .....	40
i.	IRF4, BLIMP1, XBP1 et transformation tumorale.....	41
i.	Transition plasmablaste-plasmocyte .....	41
j.	Homing et survie des plasmocytes.....	42
i.	Homing.....	42
ii.	Niches .....	42
iii.	Survie .....	43
3.	Macroglobulinémie de Waldenström .....	44
a.	Introduction .....	44
b.	Epidémiologie.....	45
c.	Facteurs de risque et prédisposition .....	46
d.	Aspects cliniques .....	47



e.	Critères diagnostiques .....	48
i.	Caractéristiques morphologiques .....	48
ii.	Caractéristiques immunophénotypiques .....	49
f.	Critères pronostiques .....	50
g.	Thérapeutiques .....	50
i.	Indication .....	50
ii.	Choix du traitement .....	51
iii.	Options thérapeutiques .....	51
4.	Physiopathogénie de la macroglobulinémie de Waldenström .....	53
a.	Contrepartie physiologique .....	53
b.	Anomalies cytogénétiques .....	54
c.	Anomalies moléculaires .....	55
i.	Mutations du gène <i>MYD88</i> .....	56
ii.	Mutations du gène <i>CXCR4</i> .....	58
iii.	Mutations du gène <i>ARID1A</i> .....	61
iv.	Mutations des gènes <i>CD79A</i> et <i>CD79B</i> .....	62
v.	Autres mutations .....	62
vi.	Anomalies du nombre de copies .....	62
vii.	Profil d'expression génique .....	63
5.	Famille des protéines ETS .....	64
a.	Généralités.....	64
b.	Domaine ETS.....	66
c.	Spécificité de liaison à des séquences ADN préférentielles.....	67
d.	Mécanismes de régulation de la liaison à l'ADN .....	68
i.	Auto-inhibition.....	69
ii.	Partenaires protéiques de co-régulation.....	70
iii.	Liens avec les voies de signalisation.....	70
e.	Localisation des liaisons à l'ADN à l'échelle du génome.....	71
f.	Mécanismes de tumorigénèse impliquant les protéines ETS.....	72
i.	Translocations chromosomiques et amplification génique .....	72
1.	Tumeurs solides .....	72
2.	Hémopathies .....	73
ii.	Autres mécanismes de gain de fonction.....	74
g.	SPI1/PU.1 .....	75
i.	Structure.....	76
ii.	Expression et fonctions .....	77
iii.	Rôle et régulation de l'expression de SPI1 dans l'hématopoïèse précoce .....	78
iv.	Partenaires.....	80
v.	Rôle dans la lymphopoïèse B.....	82
1.	Précoce .....	82
2.	Tardive .....	83
vi.	Mécanismes de transformation impliquant <i>SPI1</i> .....	85



OBJECTIFS de notre travail.....	86
RESULTATS .....	88
DISCUSSION.....	130
REFERENCES.....	138



**Liste des abréviations**

ABC : activated B-cell

Ac : anticorps

Ag : antigène

ASC : antibody-secreting cells, cellules sécrétrices d'anticorps

BCR : B-cell receptor

BL : Burkitt lymphoma

CG : centre germinatif

ChIP : chromatin immunoprecipitation

CLL : chronic lymphocytic leukemia, leucémie lymphoïde chronique

CLP : common lymphoid progenitor

CMP : common myeloid progenitor

CSH : cellules souches hématopoïétiques

CSR : class-switch recombination, commutation de classe isotypique

DC : dendritic cells, cellules dendritiques

DLBCL : diffuse large B-cell lymphoma, lymphome B diffus à grandes cellules

FDC : follicular dendritic cells, cellules dendritiques folliculaires

FT : facteur de transcription

GMSI : gammapathie monoclonale de signification indéterminée

GMP : granulocyte-macrophage progenitor

HL : Hodgkin lymphoma

HLH : helix-loop-helix, hélice-boucle-hélice

Ig : immunoglobuline

LB : lymphocytes B

LLP : lymphome lymphoplasmocytaire

LT : lymphocyte T

LZM : lymphome de la zone marginale

MBC : memory B-cells, lymphocytes B mémoire

MCL : mantle cell lymphoma, lymphome à cellules du manteau

MPP : multipotent progenitor





MO : moelle osseuse

MW : maladie de Waldenström

NF- $\kappa$ B : nuclear factor-kappa B

NGS : next generation sequencing

PEG : profil d'expression génique

PI3K : phosphatidylinositol-3 kinase

SHM : somatic hypermutation, hypermutation somatique

SLPC : syndrome lymphoprolifératif chronique

TCR : T-cell receptor

TSS : transcription starting site

UTR : untranslated region

WT : wild type



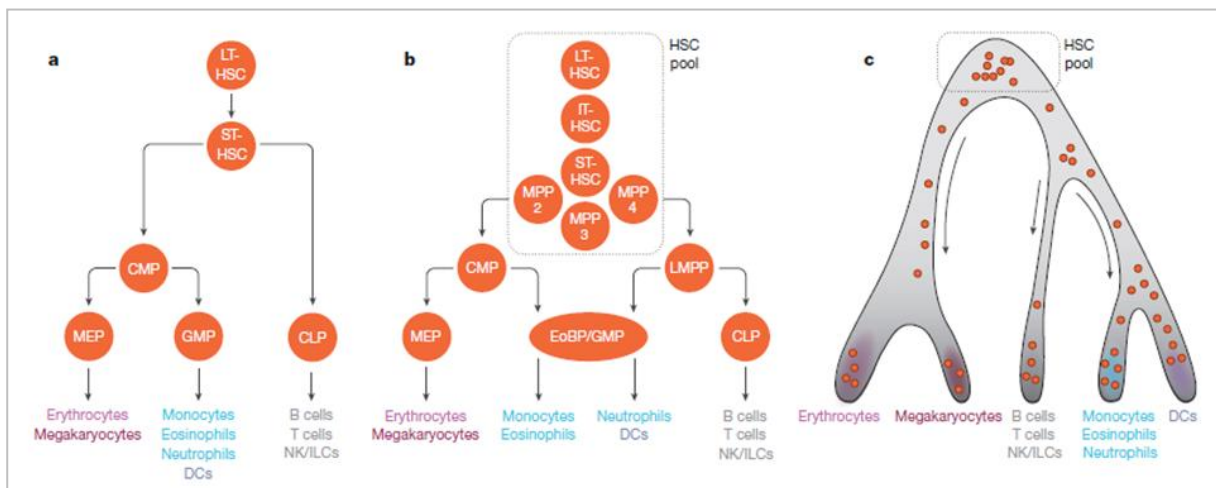
## INTRODUCTION

### 1. Hématopoïèse

L'hématopoïèse est définie comme un processus continu à travers lequel les cellules souches hématopoïétiques (CSH) se développent et se différencient en cellules matures du sang périphérique (1). Chez l'homme, l'hématopoïèse siège dans le sac vitellin, puis dans le foie et la rate avant de gagner progressivement la moelle osseuse (MO). Chez l'adulte, la MO est le site principal de l'hématopoïèse, permettant le développement de la lignée myéloïde et lymphoïde. De nombreux modèles de différenciation hématopoïétiques ont été proposés, basés initialement sur des expériences de transplantation chez la souris puis bénéficiant plus récemment de la caractérisation des différentes populations hématopoïétiques à l'échelle unicellaire chez l'homme ainsi que des techniques de reprogrammation cellulaire (1, 2). Les CSH peuvent s'auto-renouveler ou se différencier de façon contrôlée par des signaux extrinsèques, provenant de leur environnement, qui comprend molécules de la matrice extracellulaire, molécules d'adhérence, facteurs de croissance et cytokines et qui module des voies de signalisation intracellulaire et l'activité de facteurs de transcription instruisant la différenciation (3). Le modèle proposé de l' « arbre de différenciation » des CSH vers les lignées myéloïdes et lymphoïdes, construit sur les expériences de transplantation et ayant prévalu dans les années 2000, repose sur l'existence de trois sous-groupes de CSH : (a) les CSH dites « long terme » (long-term hematopoietic stem cell, LT-HSC) qui peuvent rester quiescentes ou s'auto-renouveler et qui possèdent un potentiel de différenciation large (c.-à-d. pouvant donner naissance à tous les lignages hématopoïétiques), (b) les CSH dites « court terme » (short-term hematopoietic stem cell, ST-HSC) qui dérivent des CSH « long terme » et conservent un potentiel de différenciation multiple mais dont les capacités d'auto-renouvellement sont diminuées, (c) les progéniteurs multipotents (multipotent progenitors, MPP) qui conservent leur capacité de différenciation en n'importe quelle cellule sanguine mais qui ont perdu leur capacité d'auto-renouvellement (4). A partir des MPP, il existe deux voies de développement, soit en précurseur myéloïde commun (common myeloid progenitor, CMP) soit en précurseur lymphoïde commun (common lymphoid progenitor, CLP). Chaque CMP se différencie en progéniteurs encore plus spécialisés : le GMP (granulocyte-macrophage progenitor), qui pourra à son tour se différencier en granulocytes et monocytes, et le MEP (megakaryocyte-erythroid progenitor) qui pourra donner naissance aux mégakaryocytes et érythrocytes. Le CLP a perdu son potentiel myéloïde mais il est capable de se différencier en lymphocytes B, T, NK et cellules dendritiques (dendritic cells, DC). D'autres théories définissent un progéniteur différent, le LMPP (lymphoid primed multipotent progenitor), qui présente



un potentiel érythro-mégacaryocytaire très réduit mais conserve une capacité à se différencier en cellules lymphoïdes et myéloïdes (5). Des modèles de différenciation différents ont émergé plus récemment avec l'identification de nouvelles populations souches et progénitrices lymphoïdes et myéloïdes, remettant en question leurs contributions respectives dans la différenciation hématopoïétique (1, 6). Avec l'avènement des techniques d'études à l'échelle unicellulaire, le modèle s'est encore complexifié avec la subdivision du compartiment progéniteur multipotent en sous-populations distinctes et la caractérisation de l'hétérogénéité fonctionnelle et moléculaire du pool de CSH (1). La vision classique de la hiérarchie hématopoïétique comme structure très compartimentée et stable est remise en question et remplacée par celle d'un ensemble de populations hétérogènes organisées hiérarchiquement, avec une progression graduelle d'une population à l'autre, et s'adaptant aux besoins induits par diverses situations de stress (**Figure 1**).



**Figure 1. Chronologie des différents modèles d'hématopoïèse (1).**

**a)** Modèle basé sur des travaux de recherche effectués dans les années 2000 : les CSH sont représentées comme une population homogène, en aval de laquelle une première bifurcation sépare les lignées myéloïde et lymphoïde par les populations de progéniteurs myéloïdes (CMP) et lymphoïdes (CLP) communs. **b)** Au cours des années 2005-2015, ce modèle intègre de nouvelles découvertes, le pool de CSH est considéré comme plus hétérogène tant en matière d'auto-renouvellement (axe vertical) que de différenciation (axe horizontal). **c)** A partir de 2016, les analyses transcriptomiques à l'échelle unicellulaire indiquent un continuum de différenciation. Chaque point rouge représente une seule cellule et sa localisation le long d'une trajectoire de différenciation.

## 2. Lymphopoïèse B normale

### a. Introduction

Les lymphocytes peuvent être définis comme une population cellulaire participant aux mécanismes de l'immunité, c'est-à-dire aux mécanismes de défense d'un organisme vivant contre des agents

étrangers, notamment infectieux, ou contre des agressions internes, notamment la transformation tumorale. Les lymphocytes T (provenant du thymus) et B (provenant de la moelle osseuse) participent plus spécifiquement à la réponse immunitaire adaptative, c'est-à-dire spécifique d'un antigène (Ag). La lymphopoïèse B englobe un continuum d'étapes dont l'objectif est de générer un répertoire diversifié de lymphocytes B (LB) périphériques, qui donnent naissance à des cellules productrices d'anticorps (Ac) (ASC, antibody-secreting cells) assurant la protection contre des agents pathogènes, tout en demeurant tolérantes aux propres tissus de l'organisme (7). Ce processus se caractérise par le réarrangement séquentiel et productif de l'ADN du locus de la chaîne lourde des immunoglobulines (Ig) ( $Ig\mu$ ) et des loci de la chaîne légère des Ig ( $Ig\kappa$  suivi de  $Ig\lambda$ ), puis par leur expression et assemblage en une structure appelée BCR (B-cell receptor). Les LB peuvent être définis comme une population clonalement diverse de cellules qui expriment à leurs surfaces des BCR reconnaissant des Ag (ou plus précisément leurs épitopes) spécifiques.

Les différentes populations de LB ont été identifiées et définies sur la base de leurs phénotypes, de leur localisation anatomique et de leurs différences fonctionnelles. La majorité des études qui ont caractérisé le développement et la fonction des LB ont été réalisées sur les souris, mais de nombreuses données montrent qu'il existe des différences significatives entre le développement des LB murins et humains, notamment concernant les populations de LB transitionnels et mémoires (8, 9). Chez l'homme et la souris, les différentes étapes de la lymphopoïèse B ont lieu initialement dans la MO, du stade de CSH à celui de LB immatures, puis en périphérie, des LB transitionnels aux LB matures (10). Ce processus de développement suit à la fois un programme intrinsèque à la cellule mais la survie et la différenciation correcte des LB dépendent également de contacts étroits entre les précurseurs des LB et les cellules stromales médullaires, qui leur apportent facteurs de croissance et cytokines. Les cellules stromales, comprenant notamment les cellules mésenchymateuses, les ostéoblastes et les cellules endothéliales, forment des niches particulières où résident les CSH et les précurseurs des LB (11). Les molécules d'adhésion et les intégrines intensifient les interactions entre les cellules stromales, les CSH et les précurseurs des LB, permettant auto-renouvellement et différenciation.

On peut séparer l'ontogenèse des lymphocytes B en deux phases principales, dépendantes ou non de la présence d'Ag. Un Ag est une molécule de toute nature (organique ou non) dont une petite région, appelée épitope, peut être reconnue par un récepteur de l'immunité adaptative (BCR ou TCR, T-cell receptor). Les Ag peuvent également être produits par l'individu lui-même, on parle alors d'auto-Ag ou Ag du soi.



La première phase de différenciation et de maturation des lymphocytes B est principalement indépendante de l'Ag. Elle se déroule dans la MO et aboutit à la génération de LB immatures exprimant une Ig de surface capable de reconnaître un Ag. La seconde phase, d'activation et de différenciation finale, est dépendante des Ag du soi d'abord puis du non-soi, au niveau des organes lymphoïdes secondaires. Elle aboutit à la formation de plasmocytes et de LB mémoires spécifiques d'un Ag.

### **b. Phase antigène-indépendante**

Le LB se différencie dans la MO à partir du CLP dans un microenvironnement spécialisé. Les premières étapes de ce processus sont dépendantes du contact entre les précurseurs des LB et les cellules stromales de la MO, mais ne dépendent pas de l'Ag. Les principaux stades du développement lymphocytaire B précoce (cellules pro-B, pré-B et LB immatures) sont définis et caractérisés par l'expression de marqueurs de surface spécifiques et par une succession de réarrangements impliquant différents segments des gènes des chaînes lourdes (*IGH*; H, heavy) et légères (*IGL*; L, light) d'Ig. Ces étapes de développement qui s'articulent autour de la réorganisation fonctionnelle du récepteur des LB (B-cell receptor, BCR), sont régulées par des étapes de sélection positive et négative et donnent naissance aux LB immatures (12). Les facteurs de transcriptions (FTs) IKZF1 et SPI1/PU.1 vont conditionner la différenciation des CSH en CLP, en régulant l'expression du récepteur à l'IL-7 (IL-7R $\alpha$ ) qui est essentiel à l'initiation de la lymphopoïèse précoce (13, 14). L'expression des facteurs EBF (Early B-cell factor) et LRF (liver regeneration factor) permet la différenciation du CLP en cellule pré-pro-B (15, 16). Cette cellule perd alors toute capacité de différenciation en cellules T ou NK. Les FTs PAX5 (paired box protein 5) et STAT5 (signal transducer and activator of transcription 5) sont nécessaires à la maturation en cellule pro-B (17, 18). Au cours du stade pro-B se mettent en place les réarrangements *IGH*, d'abord *DH-JH* puis *VH-DHJH* (10). L'apparition à la membrane du pré-BCR définit le stade pré-B. Le pré-BCR comprend les deux chaînes lourdes réarrangées identiques et deux pseudo-chaînes légères formées de l'association d'un domaine  $\lambda 5$  (CD179A) et d'un domaine VpréB (CD179B) (19). Cette Ig membranaire s'associe à deux hétérodimères CD79A/CD79B (Ig $\alpha$ /Ig $\beta$ ) impliqués dans la transduction du signal. L'expression d'un pré-BCR initie la progression vers les stades ultérieurs de différenciation et induit un réarrangement des chaînes légères. Les chaînes légères vont être réarrangées (*VL-JL*) et le lymphocyte pré-B perd l'expression du CD34 et exprime le CD20. Ces événements aboutissent au stade de LB immature qui exprime un BCR sous forme d'une Ig de type M (IgM) membranaire complète, 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. Suite à la recombinaison somatique intra-chromosomique aléatoire des gènes codant le domaine variable (*VH* et *VL*), 75% des LB immatures expriment un BCR qui reconnaîtra un auto-Ag (20). Ces cellules seront éliminées par apoptose (BCR auto-réactif de forte affinité) (21) ou deviendront non réactives par un processus d'anergie (BCR auto-réactif de faible affinité) (22). Un autre phénomène d'échappement à l'apoptose est lié à la ré-édition du BCR (ou BCR-editing), du réarrangement de la chaîne L (*VL*) du BCR initialement auto-réactif (23). Les LB immatures de la MO commencent à exprimer des IgD de surface, qui favorisent leurs survies et atténuent les réponses anergiques des LB en réponse à un auto-Ag (24).

Les LB immatures quittent ensuite la MO et émigrent vers les organes lymphoïdes périphériques secondaires pour poursuivre leur maturation en LB naïfs. Les LB entre les stades immature et naïf sont appelées LB transitionnels. Les LB transitionnels émigrent vers les organes lymphoïdes secondaires [rate, ganglion et tissus lymphoïdes associés aux muqueuses (MALT)] via le sang, où ils représentent 5 à 10% de tous les LB (25). Une fois dans le tissu lymphoïde périphérique, ils peuvent se différencier soit en LB folliculaires soit en LB de la zone marginale (MZB) (26). La différenciation en LB folliculaires ou MZB dépend de plusieurs voies de signalisation, incluant celles du BCR, de NOTCH2, du récepteur BAFF (B-cell activating factor) et de la voie canonique de NF- $\kappa$ B, ainsi que de signaux impliqués dans la migration et la domiciliation des MZB.

La majorité des LB immatures deviendra des LB folliculaires IgM+ IgD+. Ces derniers recirculent entre le sang (où ils représentent environ 65% de tous les LB) et les tissus lymphoïdes jusqu'à ce qu'ils rencontrent leur Ag spécifique. Ces cellules sont classiquement impliquées dans les réponses à l'Ag T-dépendantes via la formation du centre germinatif (CG) aboutissant à leur différenciation en LB mémoires (memory B-cells, MBC) ou en plasmablastes/plasmocytes (cellules sécrétrices d'anticorps, ASC). Les autres LB deviennent des MZB, localisés principalement dans le sinus marginal de la rate mais aussi dans les tissus MALT, et porteurs de BCR qui se lient préférentiellement aux Ag véhiculés par le sang, comme les composants de la paroi cellulaire des bactéries. Lorsqu'ils sont ré-exposés à l'Ag, les MZB peuvent se différencier rapidement en plasmocytes sécrétant des IgM et forment ainsi une première ligne de défense contre les pathogènes. Même si ils suivent le même processus de développement que les autres LB naïfs, les MZB ont un phénotype pré-activé et une capacité d'auto-renouvellement, qui les rapprochent des MBC (26). En plus d'être impliqués dans les réponses à l'Ag T-indépendantes, ils participent aux réponses T-dépendantes. Les MZB peuvent ainsi présenter l'Ag et délivrer des signaux de co-stimulation aux



lymphocytes T. Ils sont également capables chez l'homme de circuler entre les organes lymphoïdes et présentent des mutations somatiques *IGHV* (27). Raisons pour lesquelles les MZB sont considérées chez l'homme comme proches des MBC (28, 29).

### c. Phase antigène-dépendante

Après un épisode infectieux, se développent à la fois des plasmocytes à longue durée de vie et des MBC, exprimant des anticorps (Ac) spécifiques d'un Ag et de haute affinité. Ils peuvent persister pendant des décennies et confèrent ainsi une protection contre l'éventuelle récurrence d'agents pathogènes. Ces deux types cellulaires sont formés dans les centres germinatifs (CG) des organes lymphoïdes secondaires au cours de réponses immunitaires T-dépendantes. L'architecture des organes lymphoïdes secondaires, notamment celle des ganglions, est caractérisée par des follicules, composés en majorité de LB naïfs IgM+IgD+, séparés par une région interfolliculaire. Des zones riches en lymphocytes T (LT) bordent ces follicules (30, 31).

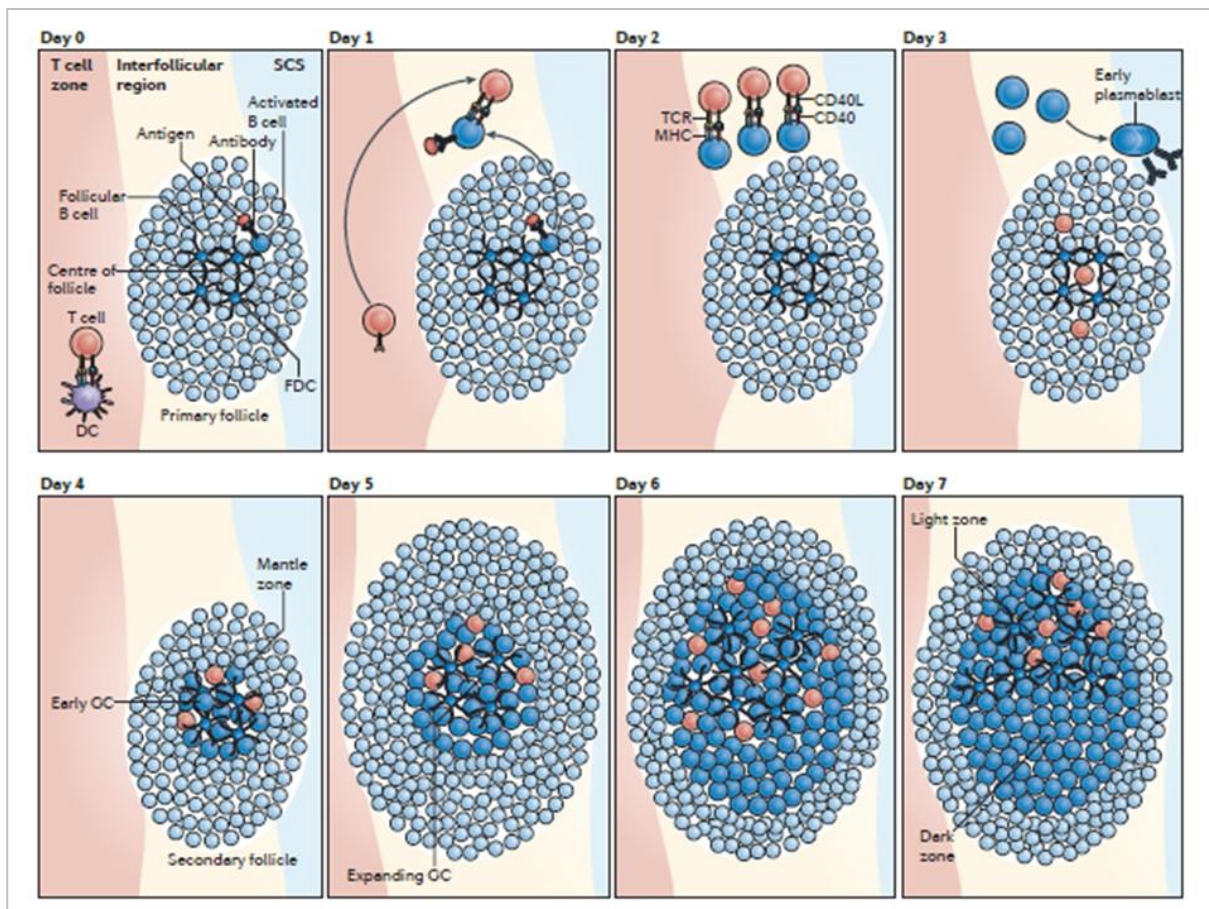
Les CG sont principalement composés de LB en division, de LT folliculaires auxiliaires (T follicular helper, TFH), de cellules dendritiques (dendritic cells, DC) folliculaires (follicular DC, FDC) et de cellules stromales (32).

La première étape de l'initiation du CG est l'activation des LB naïfs par un Ag au sein du follicule (33). La rencontre avec un Ag spécifique est la première étape d'une future interaction entre les LB et les LT afin d'induire une réponse immunitaire optimale. La signalisation du BCR après liaison à l'Ag, induit l'expression de CCR7 sur les LB folliculaires, ce qui permet à ces derniers de répondre à la chimiokine CCL21 produite par les cellules réticulaires fibroblastiques et les cellules endothéliales des veinules de la zone T (**Figure 2**). Ils vont ainsi migrer vers la frontière de cette zone où ils interagiront avec leurs LT auxiliaires correspondants, nommés LT auxiliaires pré-folliculaires (pré-TFH). La rencontre de l'Ag avec les LB fournit les signaux nécessaires pour que les LB régulent positivement leurs propres capacités de présentation d'Ag et de co-stimulation. Ainsi, les LB présentent des complexes peptide-CMH de classe II comme molécules de co-stimulation aux LT, ce qui facilite leur survie et leur production de cytokines. En retour les LT aident les LB via l'interaction CD40-CD40 ligand (CD40L) et par la production de cytokines. Au niveau de la frontière B-T se trouvent également des DC exprimant fortement le CD11c qui contribuent à la survie des LB en présentant l'Ag et en produisant BAFF (B-cell activating factor) et APRIL (a proliferation-inducing





ligand). L'interaction initiale avec les cellules pré-TFH, va conditionner trois destins pour le LB, qui va se différencier soit en plasmocytes extrafolliculaires de courte durée de vie, soit en LB du CG, soit en LB mémoires précoces (34, 35). Les cellules destinées à devenir des LB du CG réintègrent le follicule pour participer à la réaction du CG. Les plasmocytes extrafolliculaires sont de courte durée de vie (3-5 jours) et constituent, au début de la réponse immunitaire, une source importante de production d'Ac de type IgM ou IgG spécifiques de l'Ag et de relative haute affinité (36). L'affinité du BCR pour l'Ag et l'expression différentielle des récepteurs de chimiokines sont les facteurs déterminants du devenir des LB. En effet, les cellules ayant des BCR fortement activés par l'Ag, deviendront des plasmocytes extrafolliculaires, tandis que ceux qui sont moins fortement activés intégreront la réaction du CG (34). L'expression différentielle de récepteurs de cytokines/chimiokines (EBI2, CXCR4, CXCR5) et de FTs spécifiques (BLIMP1, IRF4, BCL6) régie le devenir des LB. Ainsi, une expression élevée de BLIMP1 et d'IRF4 est nécessaire pour le développement des plasmocytes, tandis qu'une forte expression de BCL6 est nécessaire pour la formation du CG, BLIMP1 et BCL6 se réprimant mutuellement (37).





**Figure 2. Initiation de la réaction et de la formation du CG (30).** Au jour 0, les LB et LT sont activés par reconnaissance de leur Ag spécifique dans le follicule primaire et la zone des LT, respectivement. Au jour 1, les LB et LT activés migrent vers la région interfolliculaire et commencent à interagir. Au jour 2, les LB et LT forment des interactions à longue durée de vie, ce qui entraîne l'activation complète des LB. Les LT acquièrent le phénotype caractéristique des LT folliculaires auxiliaires (TFH). Au jour 3, les TFH migrent de la région interfolliculaire vers le follicule. Certains LB activés par l'Ag se différencient en cellules sécrétant des anticorps ou en plasmablastes précoces qui se déplacent vers une région adjacente au sinus sous-capsulaire (SCS). Au jour 4, les LB migrent de la région interfolliculaire vers le centre du follicule - qui est caractérisé par un réseau de cellules dendritiques folliculaires (FDC)- commencent à proliférer et, par conséquent, repoussent les LB folliculaires résidants pour former le CG précoce. Cette structure est également appelée follicule secondaire, ce qui la distingue du follicule primaire "sans CG". Aux jours 5 et 6, le CG se développe rapidement en raison de la prolifération rapide des LB activés. Au jour 7, les zones sombres et claires se forment, ce qui entraîne la mise en place du CG mature. La zone sombre se compose principalement des LB activés, alors que la zone claire contient TFH et FDC. CD40L, ligand de CD40 ; DC, cellules dendritiques ; TCR, T-cell receptor.

Seuls les LB présentant une affinité significative pour l'Ag ont accès à la réaction du CG. Cette entrée dans la réaction du CG est dépendante d'une compétition interclonale entre les LB, contrôlée par les LT (38). L'oligoclonalité caractéristique du CG est ainsi déterminée quelques jours avant que le CG ne commence à se former. La différenciation des LB du CG et des TFH commence à l'extérieur des follicules. Deux études ont établi qu'un jour après l'immunisation par un Ag, les LB et LT spécifiques de l'Ag se sont déplacées et ont commencé à interagir dans la région interfolliculaire du ganglion (39, 40). Les LT pré-TFH, suite à leur interaction avec les DC 20, régulent à la hausse leur expression de BCL6 (B-cell lymphoma 6). En revanche, l'expression de BCL6 dans les LB est détectable à de faibles niveaux seulement 2 jours après l'immunisation. Au jour 2, la majorité des LT ont acquis un phénotype TFH (exprimant CXCR5, PD1 et GL7), et au jour 3, ils ont migré de la région interfolliculaire vers le follicule. Les LB précurseurs du CG se déplacent vers le centre du follicule pour former un CG précoce 1 jour plus tard, au jour 4 (**Figure 2**).

Des CG précoces peuvent être identifiés histologiquement dans les follicules au quatrième jour post-immunisation. À ce moment-là, les LB se développent et deviennent des LB qui se divisent rapidement (centroblastes) et commencent à peupler le réseau de FDC au centre du follicule. Au cours de ce processus, les LB déplacent les cellules IgM+IgD+, entraînant la formation de la zone du manteau autour du CG. 5 à 6 jours après l'immunisation, la taille du CG augmente en raison de la prolifération rapide des centroblastes, qui se poursuit jusqu'au 7e jour, où la réaction du CG est entièrement établie et le CG polarisé en deux microenvironnements distincts connus sous le nom de zone sombre et zone claire (30).



Au sein de la zone sombre (dark zone), qui se situe à proximité de la zone riche en LT, les LB activés (centroblastes) prolifèrent et modifient l'affinité de leurs régions variables par hypermutation somatique (somatic hypermutation, SHM) et leurs régions constantes par commutation isotypique (class switch recombination, CSR). Au sein de la zone claire (light zone), en aval de la zone T, les LB ne prolifèrent plus (centrocytes), sont sélectionnés sur la base de l'affinité antigénique de leurs Ig et sont mélangés avec FDC, TFH et macrophages (30). Les LB des zones sombre et claire peuvent être distingués selon l'expression de différents marqueurs comme CXCR4, CD83 et CD86 (41-43) : phénotype CXCR4 high (hi), CD83 low CD86 low pour les LB de la zone sombre et CXCR4 low CD83 hi CD86 hi pour ceux de la zone claire (42, 43).

Classiquement, la réaction du CG se déroule d'une façon cyclique au cours de laquelle les centroblastes, dans la zone sombre, diminuent l'expression du BCR de surface, subissent les SHM et prolifèrent. Ensuite, ils ré-expriment le BCR muté, et migrent vers la zone claire en tant que centrocytes. A ce stade, les LB avec des BCR mutés de haute affinité ont un avantage concurrentiel en présence de l'Ag. Ils sont sélectionnés positivement par les interactions avec les FDC et les TFH. Les LB positivement sélectionnées quittent le CG pour devenir des LB mémoires, des plasmocytes, ou réintégrer le cycle cellulaire et migrer vers la zone sombre pour un nouveau cycle de prolifération, de SHM et de sélection (44). La majorité des centrocytes qui ont été sélectionnés subissent une CSR qui conduit à la production d'Ac de classe IgG, IgA ou IgE [58]. La CSR est le processus changeant, dans un LB exprimant un Ac spécifique, la composition de la chaîne lourde du gène des Ig. Ce processus joue un rôle essentiel dans la spécificité et l'affinité globale des Ac qui en résultent. La CSR peut se produire d'une manière T-dépendante via les interactions CD40-CD40L ou d'une manière T-indépendante via BAFF et APRIL. Les processus de SHM des régions IGV ainsi que de CSR sont catalysés par la désaminase AID (activation-induced deaminase) exprimée suite à l'activation du BCR dans les centroblastes de la zone sombre (45).

Les LB mémoires quittant le CG recirculent à travers les organes lymphoïdes secondaires où, après réexposition à l'Ag, ils pourront se différencier rapidement en plasmocytes produisant des Ac de haute affinité (46). En revanche, les plasmocytes sortant du CG migrent vers la MO ou la rate, où ils sécréteront de façon continue pendant des décennies des Ac de haute affinité dans la circulation systémique (31).

#### **d. Réaction post-centro-germinative**

Les Ac sont produits par des populations rares de LB très différenciés, appelés plasmablastes et plasmocytes, ou encore cellules sécrétant des Ac (antibody-secreting cells, ASC). Des



modifications importantes de la morphologie, du profil d'expression génique (PEG) et de la durée de vie cellulaire sont associées à la formation des ASC en comparaison à leurs précurseurs dans la différenciation lymphoïde B (47, 48).

La production des ASC en réponse à l'Ag dépendant des LT est un processus en deux étapes, qui fournissent respectivement une protection humorale immédiate et prolongée. Durant la première étape, connue sous le nom de "réponse extrafolliculaire" (c'est-à-dire en dehors d'une réaction du CG), les LB reçoivent un signal dépendant du récepteur de l'Ag, ce qui conduit au développement de LB activés (« lymphoblastes B ») qui se divisent (et peuvent subir une CSR) et se différencient en plasmablastes à courte durée de vie qui sécrètent des Ac (49). La réponse extrafolliculaire n'engendre que peu de SHM et, par conséquent, l'affinité des Ac pour l'Ag tend à être modérée. Néanmoins, cette réponse est la source de la majorité des Ac protecteurs précoces.

Durant la deuxième étape, certains des LB activés pénètrent à nouveau dans le follicule et, sous l'influence des TFH, prolifèrent vigoureusement pour former un CG (50, 51). Le CG produit des plasmocytes à haute affinité et à longue durée de vie qui sont capables de maintenir un niveau élevé de sécrétion d'Ac (52). Le CG produit également des LB mémoires qui ont un phénotype lymphocytaire B, mais qui semblent programmées épigénétiquement pour se différencier rapidement en plasmocyte après une réexposition à l'Ag (53). Cette deuxième phase fournit une protection efficace contre une infection ultérieure, même si elle se produit en parallèle de la phase extrafolliculaire qui offre une protection rapide.

De nombreux aspects de la différenciation en ASC peuvent être reproduits *in vitro*, car les LB naïfs subissent à la fois la CSR et la différenciation terminale en réponse à des stimuli dérivés ou non des LT - par exemple, en réponse à la liaison du CD40 et à certaines cytokines (comme l'interleukine-4 (IL-4), IL-5 ou IL-21)- ou à des signaux comme le lipopolysaccharide (LPS) et l'ADN CpG non méthylé.

L'engagement dans la lignée des ASC est amorcé dans les LB du CG et médié par l'expression d'ensembles de FTs aux actions antagonistes (31). Au niveau transcriptionnel, la transition d'un LB activé en ASC ressemble fortement à un changement de lignée, car elle nécessite la régulation coordonnée de l'expression de centaines de gènes ainsi qu'un remodelage majeur de la chromatine. Ces changements correspondent à la fois à l'extinction d'un groupe de FTs clés qui contrôlent l'identité des LB et à l'expression de régulateurs spécifiques des ASC. Les FTs qui définissent l'identité des LB et des ASC régulent des programmes d'expression génique mutuellement antagonistes, aboutissant à ces deux états cellulaires - les LB et les ASC - qui sont stables dans le temps (54). L'échec de la transition d'un LB vers une ASC, soit par le



maintien des FTs de l'identité des LB, soit par la perte des facteurs cruciaux de différenciation en ASC, est une cause majeure des lymphomes post-CG.

### e. Lymphomagénèse B

Afin de permettre leur croissance et leur survie, les cellules tumorales des hémopathies lymphoïdes B détournent fréquemment les mécanismes qui régissent l'activation et la différenciation des LB. Les LB sont particulièrement sujets aux événements transformants car la machinerie qu'ils utilisent pour la diversification des Ac peut favoriser l'apparition de translocations chromosomiques ou de mutations impliquant des oncogènes. Les hémopathies lymphoïdes B utilisent ou modifient ainsi les voies de signalisation que les LB normaux utilisent pour détecter les Ag, leurs systèmes de régulation transcriptionnelle ou encore les régulateurs épigénétiques contrôlant la structure de la chromatine et l'expression des gènes.

Les différents sous-types d'hémopathies lymphoïdes B matures sont distingués le plus souvent selon des caractéristiques cliniques, biologiques et anatomopathologiques traditionnelles (55). Les caractéristiques moléculaires et de profil d'expression génique (PEG) ont aussi montré leurs importances permettant de différencier des tumeurs histologiquement similaires (56). Les différents sous-types de lymphomes sont caractérisés par des anomalies génétiques récurrentes et une dépendance à certaines voies de signalisation (**Tableau 1**). Chaque sous-type de lymphome présente une ressemblance phénotypique avec des LB à un stade particulier de différenciation, comme en témoignent les analyses de PEG et de la présence ou non de mutations des *IGHV*. Ainsi, le LB mature pré-CG semble être la contrepartie lymphocytaire B normale de la plupart des cas de lymphomes à cellules du manteau (mantle-cell lymphoma, MCL), qui ont des régions *IGHV* non mutées. Le sous-type *IGHV* non muté de leucémie lymphoïde chronique (U-CLL, unmutated chronic lymphocytic leukemia) et de son pendant ganglionnaire, le lymphome lymphocytaire (SLL, small lymphocytic lymphoma), sont aussi considérées d'origine pré-CG, bien que leurs PEG suggèrent tout de même une rencontre avec l'Ag (57). L'autre sous-type de LLC, au profil *IGHV* muté (M-CLL, mutated CLL), pourrait donc provenir d'un LB post-CG (58). Le LB du CG est la contrepartie normale du lymphome folliculaire, du lymphome de Burkitt (Burkitt lymphoma, BL) et du sous-type CG (centro-germinatif ou germinal center, GC) des DLBCL (59, 60). Le sous-type ABC des DLBCL se rapproche de plasmablastes post-CG (56, 59). Les lymphomes des tissus lymphoïdes associés aux muqueuses (MALT) sont d'origine extra-ganglionnaire et phénotypiquement reliés aux LB de la zone marginale post-CG (61). La leucémie à tricholeucocytes a un profil *IGHV* muté et ayant subi une CSR, ainsi qu'un PEG proche de LB



mémoires post-CG (62). La macroglobulinémie de Waldenström (MW) a également un profil *IGHV* muté, suggérant une origine post-CG, mais n'ayant en revanche pas subi de CSR (cf. paragraphe 4). Deux sous-types de lymphomes - le lymphome B primitif du médiastin (PMBCL, primary mediastinal B-cell lymphoma) et le lymphome de Hodgkin classique de sous-type scléro-nodulaire (HL, Hodgkin lymphoma) - pourraient provenir d'une sous-population rare de LB résidant dans le thymus (63) (**Figure 3**).

La lymphomagénèse est un processus multi-étapes impliquant l'accumulation de lésions génétiques qui modifient la structure et/ou l'expression d'oncogènes et de gènes suppresseurs de tumeurs. Comme pour la plupart des tumeurs humaines, plusieurs mécanismes génétiques contribuent à la transformation tumorale, notamment des mutations ponctuelles acquises associées à des gains ou pertes de fonction et des anomalies du nombre de copies des gènes.

Les hémopathies lymphoïdes B ont en plus deux autres types d'altérations génétiques : i) les translocations chromosomiques, dues à des erreurs lors des recombinaisons V(D)J, de la SHM et du CSR et ii) les hypermutations somatiques (SHM) aberrantes (ASHM), évènement satellite de l'activité enzymatique d'AID, utilisée physiologiquement pour la SHM et la CSR. L'importance de ces mécanismes a été démontrée expérimentalement dans des modèles de souris sujettes au développement de lymphomes, pour lesquelles la délétion du gène *Aicda* prévenait l'apparition de translocations *MYC-IGH* et de lymphomes (64, 65).

Nous développerons ci-dessous quelques-uns des évènements oncogéniques impliqués dans la lymphomagénèse B, en parallèle des mécanismes physiologiques d'activation, de différenciation et survie des LB. Le **tableau 1** récapitule les évènements génétiques principaux à l'origine des différentes hémopathies lymphoïdes B.



Tableau 1. Anomalies génétiques principales dans les hémopathies lymphoïdes B

Anomalies génétiques	Type d'anomalies	Phénotype	Hémopathies lymphoïdes B
<b>Signalisation BCR</b>			
<i>CD79A/B</i>	M	S	DLBCL-ABC (20), MW (5-10), FL (2), SMZL (1)
<i>CARD11</i>	M	S	DLBCL-GC (10), FL (5-10)
<i>BCL10</i>	M, Tx	S	MALT (5-10), FL (2)
<i>MALT1</i>	Tx	S	MALT (5-40)
<i>TNFAIP3/A20</i>	M, T, P	S	HL et PMBCL (30-40), DLBCL-ABC (30), MW (5-30), FL (10), SMZL (7)
<i>TCF3</i>	M	S	BL (30), DLBCL-GC (3)
<i>ID3</i>	M, T, P	S	BL (60)
<b>Signalisation TLR</b>			
<i>MYD88</i>	M	S	MW (95), DLBCL-ABC (30), SMZL (5-10), BL (5), CLL (3-5)
<b>Signalisation NF-KB</b>			
<i>BIRC3</i>	M, T	S, P	SMZL (10), CLL (5-10), MCL (5)
<i>NFKBIE</i>	M, T	S	PMBCL et HL (20-25), CLL (5-10)
<i>TNFAIP3/A20</i>	M, T, P	S	HL et PMBCL (30-40), DLBCL-ABC (30), MW (5-30), FL (10), SMZL (7)
<i>TRAF3</i>	M, T	S, P	SMZL (5), MM (4-5), CLL (1-2)
<b>Signalisation NOTCH</b>			
<i>NOTCH1</i>	M, T	S, P	CLL (10-15), MCL (5-10), SMZL (5)
<i>NOTCH2</i>	M, T, P	S, P, Diff	SMZL (20), DLBCL (8)
<i>SPEN</i>	M	S, P	DLBCL (10), SMZL (5), CLL (1)
<i>FBXW7</i>	M	S, P	CLL (2)
<b>Epigénétique</b>			
<i>CREBBP</i>	M, T	P, DNA, Diff	FL (60), DLBCL (25)
<i>EP300</i>	M, T	P, DNA, Diff	FL (10-20), DLBCL (5)
<i>MLL2</i>	M, T	P, DNA, Diff	FL (80), DLBCL (20-25), MCL (10-15), MW (5-20), CLL (1-3)
<i>EZH2</i>	M	P, DNA, Diff	FL (20-30), DLBCL-GC (20), CLL
<i>TET2</i>	M, T	DNA, Diff	DLBCL (5-8), CLL (1)
<i>ARID1A/B</i>	M, T, P	DNA, Diff	DLBCL (8-16), MW (5-10), CLL (1-2)
<b>Facteurs de transcription</b>			
<i>MEF2B</i>	M, T	S, P, DNA, Diff	DLBCL (15), FL (8-15)
<i>BCL6</i>	M, Tx	S, P, DNA, Diff	DLBCL (35)
<i>MYC</i>	M, Tx, G	P	BL (100), DLBCL-GC (20), MM (10-20)
<i>PAX5</i>	M, Tx, G	P, Diff	DLBCL (5), CLL (2-3)*, MM (2-3)*
<i>SP1B</i>	Tx, G	S, Diff	DLBCL-ABC (20)
<i>IRF4</i>	M, G	S, Diff	DLBCL (6), CLL (2-3), MM (2-3)
<i>PRDM1</i>	M, P	Diff	DLBCL-ABC (30), MM (1-2)
<i>XBP1</i>		Diff	MM (1-2)
<b>Régulateurs du cycle cellulaire</b>			
<i>CCND1</i>	M, Tx	P	MCL (90), MM (15-20)
<i>CCDN3</i>	M, Tx	P	MM (1-2)
<i>CDKN2A/B</i>	M, P	P, DNA	MCL (20-25), DLBCL (5-10)
<b>Intégrité du génome</b>			
<i>TP 53</i>	M, T, P	DNA	DLBCL (5-10), MCL (5-10), CLL (5-10), MM (3-12), MW (7-10)
<i>ATM</i>	M, T, P	DNA	MCL (20-25), CLL (15-25), DLBCL (8)
<b>Régulateurs de l'apoptose</b>			
<i>BCL2</i>	M, Tx	S	FL (90), DLBCL-GC (30-35)
<b>Autres voies de signalisation</b>			
<i>CXCR4</i>	M	P, Diff	MW (25-45), SMZL (2-4)
<i>BRAF</i>	M	P	HCL (98), MM (6-12), CLL (2-5)
<i>NRAS</i>	M	P	MM (20), CLL (1-3)
<i>KRAS</i>	M	P	MM (20), CLL (1-3)
<i>PTEN</i>	M, P	S, P	DLBCL (4-10)

Rouge, anomalies gain-de-fonction ; bleu, anomalies perte-de-fonction

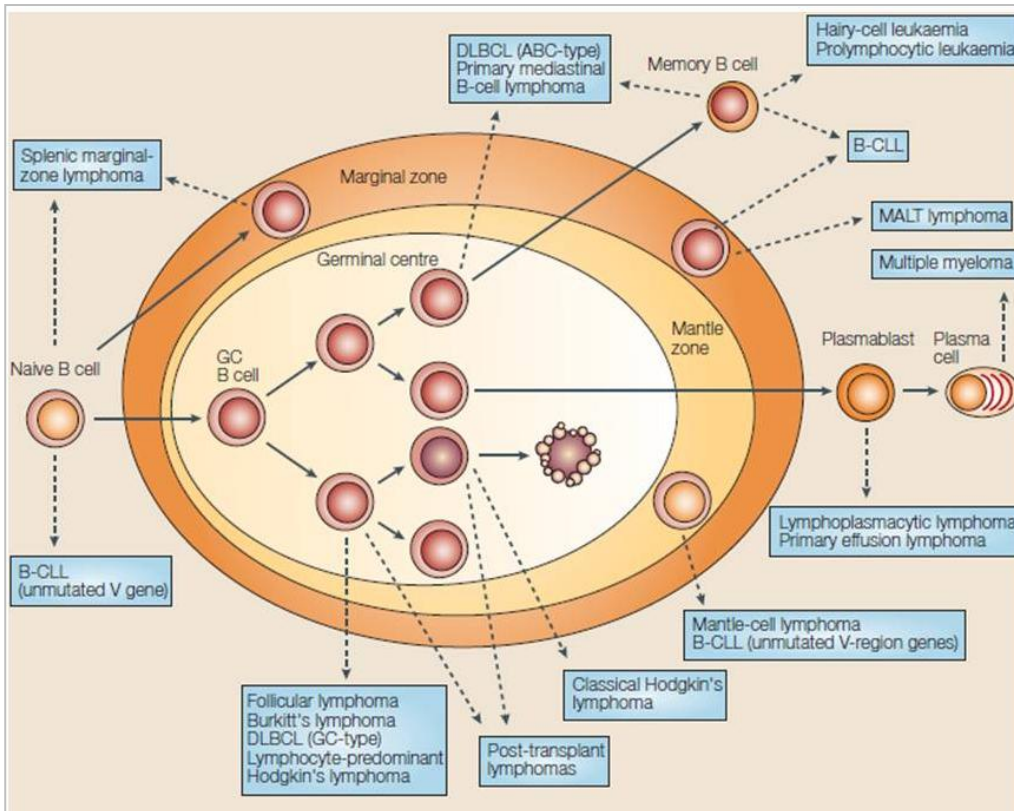
M, missense mutation ; T, truncating mutation (frameshift, nonsense) ; Tx, translocation chromosomique ; P, perte du nombre de copies ; G, gain du nombre de copies

S, survie ; P, prolifération ; DNA, dommages ADN ; Diff, blocage de différenciation

\* mutations dans une région enhancer de *PAX5*

BL, Burkitt lymphoma ; CLL, chronic lymphocytic lymphoma ; DLBCL, diffuse large B-cell lymphoma ; FL, follicular lymphoma ; HCL, hairy-cell leukemia ; HL, Hodgkin lymphoma ; MCL, mantle-cell lymphoma ; MM, multiple myeloma ; MW, maladie de Waldenström ; PMBCL, primary mediastinal B-cell lymphoma ; SMZL, splenic marginal zone lymphoma





**Figure 3. Représentation schématique des lymphomes B et de leur contrepartie physiologique** (66). CLL, chronic lymphocytic leukemia; DLBCL, diffuse large B-cell lymphoma; GC, germinal center; MALT, mucosal associated lymphoid tissue.

#### f. Voies moléculaires d'activation lymphocytaire B

La régulation de l'activation, de la différenciation, de la survie et de la mort cellulaire programmée des LB impliquent un certain nombre de molécules de surface, de facteurs cytokiniques et de voies de signalisation intracellulaires d'aval.

Les cellules B tumorales utilisent les voies de signalisation des LB normaux afin d'assurer leur croissance et leur survie. Ceci est accompli soit par des mutations gain de fonction activant des effecteurs de signalisation, soit par des mutations perte de fonction inactivant des régulateurs négatifs de signalisation, ou encore par l'activation de récepteurs autocrines. De nombreux événements oncogéniques dans les lymphomes convergent vers la voie NF- $\kappa$ B, mais les voies de la phosphatidylinositol-3 kinase (PI3K) et de la Janus kinase (JAK) sont aussi essentielles à la survie de ces cellules.



### i. Signalisation BCR

Le BCR est constituée de chaînes d'Ig lourdes (IgH) et légères (IgL), de liaison à l'Ag, qui sont couplées de façon non covalente aux sous-unités CD79A (Ig $\alpha$ ) et CD79B (Ig $\beta$ ), régulant l'expression de surface, l'internalisation et le recyclage du BCR (67). Lors de la rencontre Ag-BCR, CD79A et CD79B transmettent des signaux à une variété de voies de signalisation en aval. Chacune de ces sous-unités possède 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 doublement 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 activent les voies de signalisation NF- $\kappa$ B, PI3K, NFAT, MAPK et RAS, conduisant à la survie et à la prolifération des LB (68).

La cascade de signalisation aboutissant à l'activation de NF- $\kappa$ B 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  $\gamma$  (PLC $\gamma$ ). PLC $\gamma$  produit ensuite le second messenger DAG (diacylglycérol), qui active la protéine kinase C $\beta$  (PKC $\beta$ ), conduisant à la phosphorylation de CARD11, à l'activation du complexe CBM (CARD11-BCL10-MALT1) et à la signalisation NF- $\kappa$ B (69).

La signalisation BCR est étroitement et spécifiquement contrôlée par des régulateurs négatifs. En plus de participer à l'activation initiale de la signalisation BCR, la kinase LYN initie également une boucle de rétrocontrôle négatif (70). En effet, LYN phosphoryle CD22 et le récepteur Fc $\gamma$ , permettant le recrutement de la phosphatase SHP1 (SH2 domain-containing phosphatase 1), qui agit par déphosphorylation diverses molécules contenant des motifs ITAM, dont CD79A et CD79B.

Malgré l'accumulation de mutations et de translocations intéressant le locus du gène des Ig, les LB tumoraux maintiennent généralement l'expression du BCR à leur surface. Ceci suggère qu'ils sont capables d'utiliser les propriétés du BCR à activer en aval les voies de prolifération et de survie (66). Le BCR est un acteur majeur de la biologie des LB, jouant un rôle primordial dans la différenciation lymphoïde B, la sélection clonale Ag-dépendante et l'immunité humorale.

Les études de séquençage à haut-débit (NGS, next generation sequencing) dans le DLBCL et le BL ont révélé des mutations récurrentes de gènes appartenant à la voie du BCR, aboutissant à deux types de signalisation différents: « chronic active » et « tonic » (71, 72). Les mutations affectant la voie du





BCR sont particulièrement enrichies dans le sous-type ABC des DLBCL (71, 73). Les DLBCL-ABC ont un PEG qui rappelle celui des LB qui ont été stimulés par le BCR (74). Des études de génomique fonctionnelle ont montré que la délétion des sous-unités du BCR ou de ses composants de signalisation d'aval est létale pour les lignées de DLBCL-ABC, suggérant que la survie des DLBCL-ABC dépend des signaux constitutifs induits par le BCR (71, 75). Cette forme constitutive de signalisation a été qualifiée de « chronic active » en raison d'un certain nombre d'éléments fonctionnels qui la distinguent d'une signalisation « tonic » (71, 76). D'abord, la signalisation « chronic active » rappelle celles des LB stimulés par l'Ag, tandis que la signalisation « tonic » est typique d'un LB au repos, est dépendante du BCR mais ne nécessite pas la liaison de l'Ag (77, 78). De plus, la signalisation « chronic active » est dépendante de CARD11, contrairement à la « tonic ». Enfin, la signalisation « chronic active » agit via NF- $\kappa$ B, qui est le pilier de la survie des DLBCL-ABC, alors que la signalisation « tonic » favorise la survie principalement par la voie PI3K (71, 76).

Les mécanismes génétiques contribuant à la signalisation BCR « chronic active » dans les DLBCL-ABC ont été en partie révélés par des analyses de NGS. Dans 21% des cas, il s'agit de mutations gain de fonction dans les motifs ITAM de CD79B (ou, plus rarement CD79A) (71), qui maintiennent la signalisation BCR en atténuant le rétrocontrôle négatif impliquant l'activation par phosphorylation de la kinase LYN. Dans 10% des cas, les mutations impliquent le gène codant pour CARD11 (75), un composant du complexe CBM (CARD11-BCL10-MALT1), appelé également "signalosome", qui doit être assemblé pour la bonne transduction de la signalisation BCR (79). Ces deux types d'anomalies génétiques augmentent l'amplitude de la signalisation par le BCR, mais n'initient pas cette cascade ; ils ne peuvent donc pas expliquer seuls la dépendance des DLBCL-ABC à la signalisation BCR « chronic active ». En effet, des analyses récentes ont révélé le rôle des auto-Ag dans la survie des cellules de DLBCL-ABC, en accord avec l'expression d'un répertoire restreint des régions variables des gènes d'Ig et la démonstration que la survie des lignées cellulaires DLBCL-ABC dépend de la capacité de leur BCR à interagir avec ces auto-Ag (80). La dépendance des cellules de DLBCL-ABC à la voie de signalisation du BCR est soulignée par des études cliniques récentes rapportant des résultats positifs (30% de réponses) avec l'utilisation d'agents inhibiteurs de BTK, sélectivement chez les patients présentant ce sous-type moléculaire (81). Il est à noter que le BCR a de multiples effecteurs d'aval en plus de NF- $\kappa$ B, notamment PI3K, MAPK/ERK et NFAT, qui peuvent tous jouer un rôle important dans la transformation des cellules de DLBCL-ABC.

La signalisation « tonic » à partir du BCR est essentielle à la survie des LB au repos (76). Ce mode de signalisation est apparemment indépendant de l'Ag et favorise la survie en activant la voie PI3K.



Dans les LB, le maintien d'une signalisation « tonic » nécessite le FT TCF3, qui lie le motif ADN consensus E2-box (GCAGGGTGGG) et fonctionne soit comme activateur, soit comme répresseur transcriptionnel (82). En effet, d'une part, TCF3 régule positivement l'expression des IgH et IgL, composants principaux du BCR, mais d'autre part, réprime le gène *PTPN6*, codant pour l'inhibiteur SHP1 (82). Les études NGS combinant les séquençages du génome entier, de l'exome entier et du transcriptome ont révélé l'importance de cette voie de signalisation dans la pathogenèse du BL (72, 81, 83). La plupart des BL (70%) arborent des mutations gain de fonction touchant le gène *TCF3* ou des mutations inactivatrices d'*ID3*, régulateur négatif de TCF3. Ces mutations sont rares dans les autres hémopathies lymphoïdes B. *TCF3* est la cible de substitutions faux-sens dans le domaine d'hétérodimérisation HLH (helix-loop-helix) de la protéine. Les mutations d'*ID3* sont soit des mutations frameshift, soit des mutations non-sens, soit des substitutions faux-sens ciblant le domaine HLH, qui est nécessaire pour l'interaction avec TCF3 et l'activité inhibitrice d'*ID3*. L'activité accrue de TCF3 résultant de ces mutations aboutit à l'augmentation d'expression des gènes du BCR et à une baisse d'expression de l'inhibiteur SHP1. Ces modifications se traduisent par une augmentation de l'activité PI3K (72).

## ii. CD40

L'activation T-dépendante des LB est contrôlée par l'interaction CD40-CD40L. CD40 est un membre de la famille du TNF (TNFRSF5) constitutivement exprimé sur les LB. L'engagement de CD40 par CD40L, exprimé sur les LT auxiliaires activés, favorise le recrutement des molécules TRAF (TNF receptor-associated factor), conduisant à l'activation de plusieurs voies de signalisation, telles que les voies NF- $\kappa$ B (classique et alternative), MAPK, PI3K, et PLC $\gamma$  (34). La voie CD40 induit notamment AID, enzyme clé pour les mécanismes de SHM et CSR. La délétion des gènes codant pour CD40 ou CD40L chez la souris provoque des défauts profonds de CSR et dans la 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 de longue durée de vie (34).

## iii. BAFF et APRIL

BAFF (également connu sous BlyS et TNFSF13B) et APRIL (également connu sous le nom TNFSF13) jouent un rôle central dans la fonction des LB. Les études sur les souris déficientes indiquent que BAFF est un facteur essentiel pour la maturation et la survie des LB et qu'APRIL est un facteur important pour la CSR en IgA (84). BAFF et APRIL partagent deux récepteurs: TACI (transmembrane activator and



CAML interactor) et BCMA (B-cell maturation antigen). TACI lie BAFF et APRIL avec une affinité modérée, alors que BCMA a une forte affinité pour APRIL et une faible affinité pour BAFF. BAFF se lie également avec une haute affinité à son récepteur BAFF-R. BAFF-R est un puissant régulateur de la survie des LB matures et contrôle également la production d'Ig au cours des réponses T-indépendantes. Les mutations délétères de BAFF-R empêchent le développement correct des LB, de la réponse T-indépendante et de la production d'Ig (85). La liaison de BAFF à son récepteur active deux voies de signalisation favorisant la survie des LB, les voies NF- $\kappa$ B et PI3K/AKT (86). TACI interagit également avec MYD88 (myeloid differentiation primary response gene-88), conduisant à l'activation de la voie NF- $\kappa$ B classique, ainsi qu'à la CSR et à la différenciation en plasmocytes (87). Des mutations de TACI sont ont été identifiées dans les déficits immunitaires communs variables. La signalisation via BCMA active fortement la voie NF- $\kappa$ B classique et favorise la différenciation lymphoïde B terminale ainsi que la survie des ASC.

Dans les lymphomes B non hodgkiniens, les niveaux sériques de BAFF sont élevés et corrèlent avec l'agressivité et le pronostic de la maladie (88) ; les mutations des récepteurs de BAFF et APRIL y sont en revanche rares (89).

#### iv. Signalisation TLR

Dans les LB, les TLR sont au cœur de la réponse à l'Ag, indépendante du BCR, en détectant une variété de motifs moléculaires associés à des pathogènes (90). Dix TLR humains ont été identifiés à ce jour. TLR1/2/6/7/9 et 10 sont exprimés sur les LB, TLR1 et 9 sur les plasmocytes (91). Au moment de la fixation du ligand, les TLR de surface s'agrègent et initient la signalisation intracellulaire en activant divers adaptateurs intracytoplasmiques, y compris MYD88. Après stimulation des TLR, MYD88 est recruté comme homodimère et forme un complexe avec IRAK4, conduisant à l'activation d'IRAK1 et IRAK2. TRAF6 est ensuite activé par IRAK1, et catalyse la polyubiquitination de la kinase MAP3K7, qui à son tour phosphoryle IKK $\beta$  et déclenche l'activation de NF- $\kappa$ B (90). Les composants de parois bactériennes se fixant principalement aux TLR de surface n'activent pas fortement les LB au contraire des TLR endosomaux comme le TLR7 et 9 (91).

Les études de séquençage haut-débit ont identifié des mutations de *MYD88* dans un grand nombre d'hémopathies lymphoïdes B (92-95). MYD88 est une protéine adaptatrice qui couple les récepteurs à domaines TIR, comme les TLR (particulièrement le TLR9) ou le récepteur à l'IL-1 (IL-1R), à diverses voies de signalisation dont celles de NF- $\kappa$ B, MAPK, interféron type I ou encore JAK/STAT3 (96). MYD88



se compose d'un domaine N-terminal (death domain, permettant la liaison aux protéines IRAK), d'une région de liaison et d'un domaine C-terminal (domaine TIR, servant de médiateur avec les domaines TIR des récepteurs lors de l'activation du signal) (97). Presque toutes les mutations de *MYD88* affectent le domaine TIR. Bien qu'il existe de nombreuses mutations différentes, la plus répandue est la substitution d'une leucine pour une proline en position 265 (L265P). Elle se produit dans environ 95% des MW, 50% des GMSI IgM, 30% des DLBCL-ABC, 10% des SMZL et 3% des LLC. La mutation L265P, ainsi que d'autres mutations de *MYD88* (par exemple M232T), se situe sur un feuillet  $\beta$  d'une boucle  $\beta$ - $\beta$  du domaine TIR. Ces mutations vont modifier la structure tertiaire de MYD88 et ainsi permettre une interaction spontanée (indépendante de l'activation par les récepteurs TIR) avec IRAK4 et IRAK1. Cela conduit notamment à l'activation de la voie NF- $\kappa$ B et favorise la survie cellulaire (92, 94).

#### v. Voies de signalisation NF- $\kappa$ B

NF- $\kappa$ B a d'abord été défini dans les LB comme un FT régulant l'expression de la chaîne légère kappa (98). On sait maintenant que cette voie est utilisée par de nombreuses cellules normales et cancéreuses pour favoriser leurs survies. L'activation constitutive de la voie NF- $\kappa$ B est ainsi le mécanisme le plus couramment utilisé par les hémopathies lymphoïdes B pour éviter la mort cellulaire (99).

Dans les LB, l'activation NF- $\kappa$ B se produit de façon transitoire en aval de nombreux récepteurs, tels que le BCR, le CD40, le récepteur BAFF et divers TLR. NF- $\kappa$ B fait référence à une famille de FTs qui sont formés par homo- et hétéro-dimérisation des sous-unités p65 (codé par *RELA*), RelB, c-Rel, p50 (codé par *NFKB1*), et p52 (codé par *NFKB2*). Ces facteurs se déplacent entre le cytoplasme et le noyau, déplacements régulés par divers événements de signalisation intracellulaire. La voie "classique" de l'activation de NF- $\kappa$ B utilise un complexe I $\kappa$ B kinase (IKK) composé de sous-unités  $\alpha$ ,  $\beta$  et  $\gamma$  (NEMO). L'activation de ce complexe IKK par des signaux en amont est complexe, nécessitant la phosphorylation d'IKK $\beta$  et l'ubiquitination d'IKK $\gamma$ . Une fois actif, IKK $\beta$  phosphoryle I $\kappa$ B $\alpha$ , 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 $\kappa$ B $\alpha$  phosphorylé est ubiquitiné par l'ubiquitine ligase FBXW1A (F-box/WD repeat-containing protein 1A ; connu aussi sous le nom de  $\beta$ TrCP) puis dégradé par le protéasome, permettant aux dimères NF- $\kappa$ B de s'accumuler dans le noyau où ils activent la transcription de gènes cibles. La voie « alternative » de NF- $\kappa$ B est initiée par les récepteurs CD40 et BAFF. Lors de la liaison au récepteur, le complexe régulateur négatif TRAF3/MAP3K14-TRAF2/BIRC3 est désolidarisé, ce qui permet la libération cytoplasmique et la stabilisation de MAP3K14, kinase centrale de l'activation de cette signalisation alternative. Cette



voie dépend d'un complexe IKK composé de deux sous-unités IKK $\alpha$ . Cette forme d'IKK est classiquement activée par la kinase MAP3K14 (Mitogen-Activated Protein Kinase 14 ; aussi connu sous le nom de NIK) (100). Ce complexe IKK $\alpha$  phosphoryle p100, produit de traduction 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- $\kappa$ B activent la transcription d'un module anti-apoptotique, comprenant les membres de la famille BCL2, comme BCL2L1 (BCL2 like 1 ; connu aussi sous le nom de BCL-XL) et BCL2A1 (BCL2 Related Protein A1) ainsi que CFLAR (CASP8 and FADD like apoptosis regulator ; connu aussi sous le nom de c-FLIP), BIRC2 (baculoviral IAP repeat containing 2 ; connu aussi sous le nom de c-IAP1), BIRC3 (ou c-IAP2) et d'autres protéines (101, 102).

NF- $\kappa$ B a un rôle essentiel dans la génération, le maintien et la fonction des LB aux différentes étapes de leur différenciation (103). Au sein du CG, la signature transcriptionnelle de NF- $\kappa$ B est absente des LB de la zone sombre mais présente dans une sous-population de LB de la zone claire (104). Différentes sous-unités de NF- $\kappa$ B, participant à des complexes protéiques distincts, sont nécessaires à différentes étapes 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 (105). Dans les LB de la zone claire, NF- $\kappa$ B transactive *IRF4*, qui régule à la baisse l'expression de *BCL6* et favorise ainsi la différenciation post-CG (106). L'importance de l'axe NF- $\kappa$ B-IRF4 dans la physiologie du CG a été bien documentée par les modèles murins où l'activation constitutive de NF- $\kappa$ B conduisait à l'accumulation de plasmocytes alors que la délétion d'IRF4 diminuait leur formation (107-109).

Les hémopathies lymphoïdes B associées à une activation de NF- $\kappa$ B comprennent les DLBCL-ABC, le HL, le PMBCL, le lymphome gastrique MALT et le myélome multiple. Un mécanisme récurrent de lymphomagénèse est l'inactivation de gènes codant pour des régulateurs négatifs des voies NF- $\kappa$ B. Ainsi, des mutations et des délétions d'*IKBKA*, aboutissant à des formes tronquées de la sous-unité I $\kappa$ B $\alpha$ , se produisent dans environ un quart des cas d'HL. Ces anomalies entraînent la migration constitutive des dimères NF- $\kappa$ B vers le noyau (110). *TNFAIP3*, qui code pour l'inhibiteur A20, est inactivé par mutation, délétion ou par un mécanisme d'extinction épigénétique dans de nombreux sous-types de lymphomes, y compris les HL, PMBL, DLBCL-ABC et MALT (111-113), ainsi que le myélome (114). A20 est une enzyme désubiquitinante, qui inactive IKK $\beta$  en supprimant les chaînes d'ubiquitine liées à la lysine K63 d'IKK $\gamma$ . *TNFAIP3* est un gène cible de NF- $\kappa$ B et fonctionne donc comme un rétrocontrôle négatif de l'activation de NF- $\kappa$ B. La réexpression d'une protéine A20 fonctionnelle dans des lignées cellulaires



de DLBCL conduit à la relocalisation cytoplasmique de NF- $\kappa$ B et à l'apoptose de ces cellules (73). Il faut noter que l'inactivation d'A20 coexiste souvent avec d'autres lésions génétiques qui activent NF- $\kappa$ B via la signalisation du BCR ou de MYD88 (92).

Des délétions récurrentes de 4 paires de bases du gène *NFKBIE*, codant pour I $\kappa$ B $\epsilon$ , un régulateur négatif de NF- $\kappa$ B dans les LB normaux, ont été décrites dans 5-10% des LLC (115, 116) et 23% des PMBL (117) et associées à un pronostic péjoratif.

Le gène *BIRC3* est fréquemment altéré par des mutations, des délétions ou une combinaison des deux dans environ 10% des lymphomes spléniques de la zone marginale (SMZL) (118) et 5% des LLC (119). Les mutations inactivatrices de *BIRC3* sont principalement des mutations décalant le cadre de lecture (frameshift) ou aboutissant à un codon stop (non-sens), responsables de la perte du domaine C-terminal RING, dont l'activité d'E3 ubiquitine ligase est nécessaire pour diriger MAP3K14 vers le protéasome. Outre *BIRC3*, TRAF3, un autre composant du complexe régulateur négatif TRAF3/MAP3K14-TRAF2/BIRC3, est muté dans environ 5% des SMZL et des 2-3% des LLC (118, 120). Ce sont aussi des mutations inactivatrices entraînant la perte du domaine C-terminal MATH de la protéine qui fournit le site d'ancrage pour MAP3K14 et est nécessaire pour le recrutement puis la dégradation de MAP3K14 par *BIRC3*. Une des conséquences fonctionnelles des mutations de *BIRC3* et *TRAF3* est la stabilisation de MAP3K14 dans le cytoplasme et l'activation constitutive de la voie alternative de NF- $\kappa$ B. A noter que le même type de mécanisme oncogénique est retrouvé dans le cas des t(11;18), observé dans 25% des lymphomes MALT gastriques et aboutissant à la fusion du domaine N-terminal de *BIRC3* et la partie C-terminale de *MALT1* (121). Le domaine RING de *BIRC3* est donc manquant dans ce gène de fusion.

#### **g. Maintien de l'identité B**

Les FTs, exprimés dans les LB matures, favorisent des programmes d'expression de gènes participant à la réponse à l'Ag et aux LT et empêchent la différenciation prématurée en ASC (**Figure 4**). Les facteurs les mieux caractérisés et leurs implications éventuelles dans les mécanismes de transformation sont discutés ci-dessous.



### i. MEF2B

Le FT MEF2B (myocyte-specific enhancer factor 2B) est exprimé dans les LB du CG et à un stade précoce de la formation du CG (122). Son implication dans la physiologie du CG avait été suggérée jusqu'ici par le fait qu'il est directement transactivé par BCL6 (B-cell lymphoma 6) (122) et par l'identification dans les lymphomes de mutations dérégulant son activité (122, 123). Ces mutations augmentent l'activité transcriptionnelle de MEF2B (et notamment l'expression de BCL6) soit en empêchant l'interaction physique avec le co-répresseur CABIN1 (mutations faux-sens dans le domaine N-terminal MEF/MAD) soit par la perte de ses motifs régulateurs négatifs sujets à la phosphorylation et sumoylation dans le domaine C-terminal de la protéine (mutations tronquantes) (122). Son rôle a été confirmé tout récemment (124) dans des modèles murins : la délétion de *Mef2b* réduisait la formation du CG alors que l'expression du mutant D83V de *Mef2B* entraînait une hypertrophie du CG et le développement d'un lymphome.

### ii. MYC

*MYC* a été le premier gène à être impliqué dans la lymphomagenèse de tumeurs dérivant du CG (125, 126) mais son rôle dans la physiologie du CG n'a commencé à être compris que récemment. *MYC* a un profil d'expression bimodal dans le CG : son expression est induite précocement lors de l'initiation du CG puis elle est éteinte dans les LB de la zone sombre, par l'action du répresseur transcriptionnel BCL6 (127, 128), et est enfin ré-induite de façon transitoire dans une sous-population de LB de la zone claire qui sont destinés à entrer à nouveau dans la zone sombre (104, 127-129). Bien que l'importance biologique de l'extinction de l'expression de *MYC* dans les LB de la zone sombre reste inconnue, l'induction initiale de son expression implique des programmes transcriptionnels qui sont associés à ces nombreuses fonctions, comme la réplication de l'ADN, l'entrée dans le cycle cellulaire et la maintenance des télomères (130, 131).

La dérégulation de *MYC* est impliquée dans la pathogenèse d'un certain nombre de lymphomes. Elle peut être un événement primaire comme dans le BL ou secondaire comme dans d'autres lymphomes agressifs (DLBCL, lymphome plasmablastique, lymphome à cellules du manteau). Les principaux mécanismes à l'origine de la dérégulation de l'expression de *MYC* sont les translocations chromosomiques et les amplifications géniques. Les translocations chromosomiques associées aux hémopathies lymphoïdes B ne génèrent généralement pas de protéine de fusion. Ils juxtaposent plutôt des séquences régulatrices hétérologues (promoteurs ou enhancers), dérivées de chromosomes partenaires, impliquant souvent les loci des Ig, à proximité de proto-oncogènes, conduisant à l'expression dérégulée de leurs protéines normales (132). Les translocations impliquant *MYC* peuvent être divisées en deux types distincts relatifs à





leurs mécanismes de génération : i) celles médiées par des erreurs dans le processus de CSR dépendant d'AID (translocations *IGH-MYC* du BL sporadique) ; et ii) celles satellites du mécanisme d'hypermutation somatique (SHM) médié par AID, qui génèrent également des cassures d'ADN (translocations *IGH-MYC* du BL endémique). Ces translocations sont à l'origine d'une "dérégulation hétérotopique", c'est-à-dire aboutissant à l'expression d'une protéine normalement non exprimée (128, 133). Des marques d'hypermutation somatique aberrante (ASHM), mécanisme de dommages à l'ADN menant à l'accumulation de multiples mutations autour des régions 5' UTR de gènes qui ne sont normalement pas mutées dans les LB normaux du CG, ont été observées dans le gène *MYC* (134) et pourraient aussi participer à sa dérégulation. Plus récemment, des variations du nombre de copies d'éléments régulateurs en cis de *MYC*, modifiant son expression, ont été décrites dans le myélome (135).

### iii. PAX5

Très peu d'aspects du programme transcriptionnel des LB peuvent fonctionner indépendamment de PAX5 (également connu sous le nom de BSAP, B-cell specific activator protein). PAX5, exprimé tout au long de la lymphopoïèse B, est primordial à l'identité lymphoïde B. En effet, l'inactivation de *PAX5* dans les LB matures entraîne la perte de l'identité lymphoïde B et le retour à un stade progéniteur (136). Chez les souris déficientes pour *Pax5* (*Pax5*<sup>-/-</sup>), le développement des progéniteurs lymphoïdes est bloqué au stade pro-B (137). Ces souris sont dépourvues de lymphocytes B, bien que leurs précurseurs conservent un potentiel de développement et soient capables de se différencier vers l'ensemble des autres lignées hématopoïétiques (138). Ce n'est que lorsque PAX5 est exprimé que l'on observe une restriction forte vers la lignée B. Différentes études montrent que pour établir et maintenir l'identité des LB, PAX5 active non seulement des gènes importants pour leurs fonctions mais surtout, réprime des gènes nécessaires au développement des autres lignées, comme *Notch1* pour la différenciation T (138, 139). PAX5 se lie à plusieurs milliers de sites ADN dans les LB folliculaires (140). PAX5 régule ainsi l'expression de composants du BCR (tels que l'IgH et CD79A), d'autres récepteurs (tels que CD19 et CD21), ainsi que des FTs comme IRF4 (interferon-regulatory factor 4), IRF8, BACH2 (BTB et homologue CNC 2), Aiolos (également connu sous le nom d'IKZF3) et SPIB (141, 142). Il a été démontré que l'inhibition de PAX5 est l'une des premières étapes de la différenciation plasmocytaire, aboutissant à la formation des pré-plasmablastes (143). PAX5 réprime l'expression de la chaîne J (144), de *XBP1* (145) et de *PRDM1* (BLIMP1), gènes clés du développement plasmocytaire (146, 147). PAX5 réprime également l'expression d'un ensemble de gènes qui ne sont normalement pas





exprimés par les LB, comprenant FMS-related tyrosine kinase 3 (*FLT3*), CC-chemokine receptor 2 (*CCR2*) et *CD28*. Ces gènes sont à nouveau exprimés suite à la régulation négative de *PAX5* dans les ASC (143, 147).

Bien que l'on suppose que l'expression de *PAX5* doit être éteinte pour permettre la différenciation en ASC, cela n'a pas été formellement démontré. L'expression forcée de *PAX5* dans des LB inhibe partiellement la différenciation en ASC (148, 149). Il a été suggéré que *PAX5* favoriserait la lymphomagenèse B par une activation de la signalisation BCR (150). L'expression de *PAX5* peut être maintenue dans certains lymphomes suite notamment à des translocations impliquant son locus génique et celui des *IGH* (151). Bien que des modèles murins intégrant ces translocations aient été décrits, leur implications directes dans les anomalies de différenciation en ASC n'est pas encore complètement établie (152). Une autre possibilité est que ce soit la fonction de *PAX5*, et non son expression, qui soit modifiée afin d'initier la différenciation en ASC. En effet, il a été démontré que la réexpression des gènes réprimés par *PAX5* est un événement très précoce dans la différenciation en ASC (143). Cette réexpression se produit en présence d'une concentration normale de la protéine *PAX5*, ce qui suggère qu'il existe une altération de la fonction de *PAX5*, potentiellement par une modification post-traductionnelle ou un changement dans les interactions avec ses co-facteurs.

Récemment, des mutations dans des régions enhanceurs de *PAX5*, réduisant son expression, ont été décrites dans la LLC (153) et le myélome (135). Des mutations ponctuelles et des amplifications de *PAX5* ont également été identifiées dans 5% des DLBCL (154). Le lien avec les mécanismes de transformation tumorale sont en revanche non connus.

#### iv. E2A/TCF3

Le gène *TCF3* code pour deux isoformes protéiques d'E2A, E12 et E47, appartenant à la famille des FTs HLH. Dans les tissus non lymphoïdes, E2A est majoritairement exprimé sous la forme d'un hétérodimère E47/E12, alors que dans les LB il s'agit essentiellement d'un homodimère E47/E47. Son expression est indispensable à la survie des cellules aux stades précoces du développement lymphoïde B (82) et est essentielle à la régulation de la transcription et à l'expression d'EBF1 (early B-cell factor 1).

Les souris dans lesquelles le gène *Tcf3* a été invalidé (souris *Tcf3* <sup>-/-</sup>) se caractérisent par un arrêt du développement lymphoïde B à son stade le plus précoce, par l'absence totale de réarrangements des *IGH* et par un défaut d'expression de CD19, RAG1, CD79a, λ5 et *PAX5* (155).



Dans les LB matures, E2A/TCF3 est préférentiellement exprimé dans les LB de la zone sombre (41), dans lesquels il favorise la progression du cycle cellulaire en induisant la transcription de *CCND3* (qui code pour la cycline D3) et *E2F2*, et en inhibant *RB1* (72). E2A/TCF3 régule également l'expression des composants du BCR en transactivant les gènes *IGH* et *IGL* et en réprimant *PTPN6* (tyrosine protein phosphatase non-receptor 6), qui code pour SHP1, un régulateur négatif de la signalisation en aval du BCR. L'activité d'E2A/TCF3 est modulée négativement par hétéro-dimérisation avec ID3 (inhibitor of DNA binding 3) (82). L'axe E2A-ID3 semble être important dans les LB du CG car E2A est nécessaire à la formation du CG (156) et car il existe des anomalies génétiques inactivatrices fréquentes ciblant les gènes *TCF3* et *ID3* dans le lymphome de Burkitt, un sous-type de lymphome dérivant du CG (cf. paragraphe 2.f.i) (72, 83).

#### v. BACH2

BACH2 (BTB (broad complex, tramtrak et bric-à-brac) et CNC homology 2) appartient à la famille des FTs de type leucine zipper. Il fonctionne comme un répresseur transcriptionnel par hétérodimérisation avec les petites protéines Maf (MafK, MafG et MafF) puis en se liant à MARE (Maf recognition elements) (157, 158). BACH2 est exprimé tout au long de la différenciation lymphoïde B mais, conformément à sa régulation positive par PAX5 (129), n'est pas exprimé dans les plasmocytes. Les souris déficientes en Bach2 n'ont pas de CG et n'expriment pas Aid (activation-induced cytidine deaminase, enzyme clé des mécanismes de SHM et CSR) (159). Une cible majeure de BACH2 dans les LB est *PRDM1* (160, 161). En l'absence de BACH2, la formation de LB matures est relativement normale, mais si ces LB sont activés, BLIMP1 s'exprime prématurément, ce qui entraîne une répression d'AID et une exacerbation de la différenciation plasmocytaire (162). La répression de BACH2 a également été décrite comme un élément clé de la capacité accrue des LB mémoires à engendrer une différenciation plasmocytaire après une réexposition à l'Ag (53). Les données sur la liaison à l'ADN de BACH2 dans les LB sont rares. Une étude récente de ChIP-seq (chromatin immunoprecipitation sequencing) réalisée dans des LB spléniques murins (163) a révélé que BACH2 est principalement lié à l'ADN avec son hétérodimère partenaire MAFK. Le complexe BACH2-MAFK se lie à des ensembles de gènes connus pour être impliqués dans la réponse du CG, dont 60% sont également des cibles de BCL6. Environ 30% des pics de BACH2 chevauchent ceux de BCL6, comprenant notamment des séquences régulatrices en cis du gène *PRDM1*. BCL6 module également la stabilité de la protéine BACH2 et leurs niveaux d'expression protéique sont positivement corrélés dans les LB du CG.



Plusieurs études suggèrent que *BACH2* puissent fonctionner comme un gène suppresseur de tumeur. Une perte d'hétérozygotie de *BACH2* a été décrite à une fréquence de 20% dans les lymphomes B (164). L'expression ectopique de *BACH2* dans des lignées cellulaires de BL réduit la prolifération cellulaire et augmente les effets cytotoxiques des espèces réactives de l'oxygène (ROS) produites par les chimiothérapies (165). Les patients atteints de DLBCL et MW présentant une expression de *BACH2* élevée ont un meilleur pronostic (166, 167).

## vi. BCL6

BCL6 (B-cell CLL/lymphoma 6) est un répresseur transcriptionnel appartenant à la famille des FTs à doigt de zinc (Krüppel-type zinc finger). Il est indispensable à la formation du CG (168). Il est exprimé par les LB du CG, essentiellement par les centroblastes et par la majorité des centrocytes (169). Il possède la particularité de subir des mutations somatiques au cours de la maturation lymphoïde B (170). BCL6 favorise la prolifération des LB du CG et, par l'inhibition de la réponse aux dommages de l'ADN, participe à leurs tolérances de taux élevés de SHM. BCL6 se lie aux régions régulatrices de plusieurs milliers de gènes dans les LB du CG, y compris ceux qui contrôlent l'expression de régulateurs du cycle cellulaire comme p21 et p53, ainsi que des oncogènes, notamment *MYC* et *BCL2* (163, 171, 172). L'activité de répresseur transcriptionnel de BCL6 dépend de sa capacité à se lier à des motifs ADN spécifiques et à recruter des complexes d'histone déacétylases de classe I et II, directement ou par interaction (via son domaine BTB) avec de multiples co-répresseurs (comme NCOR1 (nuclear receptor co-repressor 1) et NCOR2 (également connu sous le nom de SMRT)) (171, 173-175).

BCL6 assure l'inhibition de la différenciation en plasmocytes des LB du CG en réprimant l'expression de BLIMP1, permettant ainsi à la réaction du CG de se poursuivre pendant une durée suffisante avant que la différenciation en plasmocytes ne se produise (47). Les souris transgéniques exprimant de manière constitutive BCL6 ne génèrent pas de plasmocytes (176). Dans une réponse immunitaire T-dépendante, les souris *Bcl6*<sup>-/-</sup> produisent plus de plasmocytes que les souris témoins. De plus, les LB isolés de souris *Bcl6*<sup>-/-</sup> traitée par LPS sécrètent des Ig à un taux plus élevé que les cellules contrôles (177). L'IL21 (178, 179) ainsi que les FTs IRF8 et MEF2B (180, 181) induisent l'expression de BCL6, tandis qu'IRF4 et BLIMP1 la répriment (106, 182). Après son induction, le niveau d'expression de BCL6 est finement réglé par un mécanisme autorégulateur au cours duquel BCL6 se lie à son propre promoteur afin de diminuer sa propre



transcription (183). L'importance de ce mécanisme est souligné par la présence de lésions génétiques abolissant ce circuit autorégulateur dans les DLBCL de type CG.

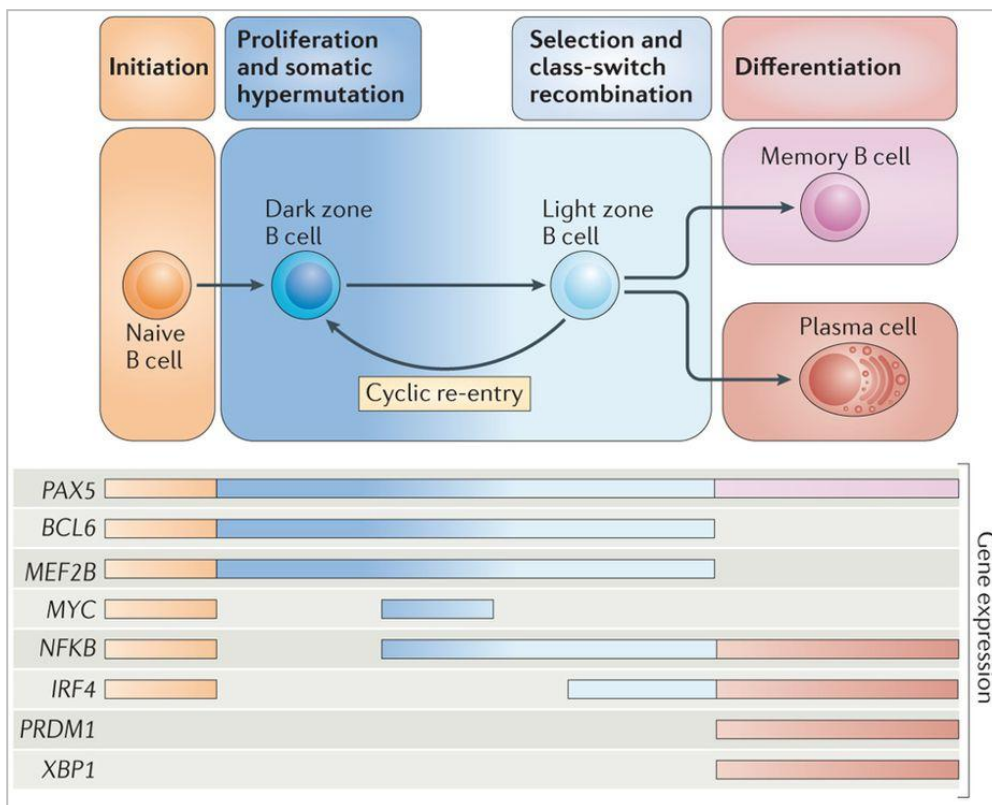
Un partenaire majeur de BCL6 dans la réaction du CG et dans les mécanismes de lymphomagenèse B est l'histone méthyltransférase EZH2 (enhancer of zeste homologue 2). EZH2 est un composant du complexe PRC2 (polycomb repressive complex 2) et catalyse la méthylation du résidu lysine 27 de l'histone 3 (H3K27) (184-186), marque de répression transcriptionnelle. L'expression d'EZH2 dans les LB matures est limitée au CG (187, 188) et est nécessaire à son développement complet (189, 190). L'activité épigénétique d'EZH2 favorise la prolifération des LB du CG et inhibe la différenciation plasmocytaire (191). Une proportion de promoteurs ciblés par EZH2 est également modulée par BCL6, ce qui suggère que ces deux facteurs peuvent coopérer dans la répression transcriptionnelle d'un sous-ensemble de gènes cibles (191). Il a été confirmé plus récemment que la fonction d'EZH2 dans la formation du CG et la lymphomagenèse B nécessitait la liaison de BCL6 à certains sites ADN spécifiques et la présence d'un complexe PRC1-BCOR-CBX8 (192).

La dérégulation de BCL6 joue un rôle crucial dans la lymphomagenèse en renforçant la prolifération des LB du CG, en supprimant les réponses aux dommages à l'ADN et en bloquant la différenciation lymphoïde B terminale (193). Une variété de mécanismes concourt à la dérégulation de l'expression et de l'activité de BCL6 dans les lymphomes. Le locus *BCL6* est ciblé par des translocations chromosomiques qui placent la séquence codante, complète et non mutée, de la protéine BCL6 en aval de régions régulatrices hétérologues de chromosomes partenaires (194). Ces régions comprennent principalement les loci *IGH* et *IGL*, mais aussi les promoteurs d'une variété de gènes qui se caractérisent par une large gamme d'expression tout au long de la différenciation lymphoïde B, y compris dans les stades post-CG (194, 195). Ces mécanismes de substitution de promoteurs empêchent notamment la diminution d'expression de BCL6 normalement associée à la différenciation post-CG. Ils induisent ainsi l'expression constitutive de BCL6 et le maintien pathologique du phénotype CG, y compris la tolérance aux dommages à l'ADN et le blocage de la différenciation terminale. De plus, la liaison de BCL6 lui-même ou d'IRF4 au promoteur de *BCL6* peut être altérée par des mutations qui contribuent à la dérégulation de l'expression de *BCL6* en interférant avec son circuit autorégulateur (183), ou avec la répression de l'expression de *BCL6* induite par le CD40 et IRF4 (106). La dérégulation de l'expression et/ou de l'activité de BCL6 peut également être engendrée par une variété de mécanismes indirects : i) des mutations gain de fonction de son régulateur positif MEF2B (122), ii) l'inactivation de FBXO11, qui est un adaptateur spécifique pour l'ubiquitination et la dégradation de



BCL6 (196), et iii) des anomalies perte de fonction des acétyltransférases CREBBP et EP300, qui sont impliquées dans l'inactivation de *BCL6* par acétylation (197, 198).

Les anomalies génétiques impliquant une dérégulation de BCL6 sont des événements fréquents dans les DLBCL et les FL, survenant dans environ 30 et 15% des cas respectivement. Les propriétés oncogéniques de ces dérégulations de BCL6 ont été confirmées dans des modèles murins (176). De plus, des translocations *IGH-BCL6* ont également été identifiées dans des DLBCL de type CG (193).



**Figure 4. Réseaux transcriptionnels à l'origine de la réaction du CG (199).** De multiples FTs interviennent dans la régulation de l'initiation, de l'expansion et de la sortie du CG. PAX5 est exprimé tout au long de la vie des LB matures - à l'exception des plasmocytes - alors que l'expression de BCL6 et de MEF2B s'arrête à la sortie du CG. MYC est nécessaire lors de l'initiation du CG et de la ré-entrée dans la zone sombre à partir de la zone claire. NF-κB et IRF4 sont exprimés lors de l'initiation du CG, sont absents dans les LB de la zone sombre et réapparaissent ensuite dans les LB de la zone claire. PRDM1 et XBP1 sont induits dans les cellules qui sont engagées dans la différenciation des plasmocytes. Ces FTs sont souvent impliqués dans la modulation transcriptionnelle des uns et des autres. NF-κB active IRF4, qui contribue à l'induction de BCL6 pendant l'initiation du CG, mais favorise la régulation à la baisse de BCL6 et l'expression de PRDM1 à la sortie du CG. BCL6 régule négativement l'expression de MYC et PRDM1.

### vii. OBF1 et OCT2

OCT2 (octamer-binding factor 2) (codé par le gène *POU2F2*) et son co-activateur OBF1 (codé par *POU2AF1*) ont été découverts du fait de leurs activités de liaison au niveau des promoteurs des gènes des Ig ; cependant, l'expression des *IGH* et *IGL* est normale en leur absence (200). En revanche, ces FTs sont nécessaires au développement des MZB (201-203). Les LB déficients en OBF1 ne peuvent pas former de CG et, lorsqu'ils sont exposés à des stimuli T-dépendants, initient mais ne peuvent pas accomplir complètement leur différenciation plasmocytaire (204, 205). Cela reflète en partie une réponse défectueuse à l'IL4. OCT2 participe à la différenciation plasmocytaire par le contrôle direct de l'expression de la chaîne alpha du récepteur à l'IL5 (206). Enfin, OCT2 et OBF1 permettent aux LB activés de sécréter des niveaux optimaux d'IL6, inducteur de la différenciation des cellules TFH (207). OCT2 et OBF1 sont exprimés dans un certain nombre de lymphomes mais leur éventuelle implication dans la transformation tumorale n'est pas connue.

### viii. PU.1, SPIB, ETS1

Les FTs de la famille ETS, PU.1 (également connu sous le nom de SPI1) et SPIB, peuvent se fixer sur l'ADN en formant des complexes avec IRF4 et IRF8 (208). PU.1 (209) et IRF8 (210) sont exprimés de manière constitutive tout au long de la différenciation lymphoïde B, alors que l'expression de SPIB est hétérogène au cours de la différenciation, maximale dans les LB du CG (211), où elle est régulée par OBF1 (212). L'expression des trois facteurs, PU.1, IRF8 et SPIB, est ensuite régulée à la baisse dans les plasmocytes, la répression de SPIB étant médiée par BLIMP1 (213).

L'inactivation spécifique de PU.1 (214, 215) ou d'IRF8 (216) dans les LB n'entraîne que des altérations minimales du développement et de la fonction des LB, alors que la perte de SPIB empêche le maintien du CG (217). Ces phénotypes relativement bénins sont surprenants car PU.1 et IRF8 lient les régions régulatrices de milliers de gènes dans les LB (218, 219). Des données récentes suggèrent que ces FTs jouent un rôle important dans la différenciation tardive des LB, rôle qui est dose-dépendant et qui peut être masqué par des mécanismes de redondance. Les LB dépourvus de PU.1 et IRF8 montrent une hyperréactivité marquée aux stimuli T-dépendants et indépendants et génèrent un grand nombre de plasmocytes et de cellules qui ont subi une CSR, en particulier des cellules sécrétrices d'IgE (181). Au niveau cellulaire, le complexe PU.1-



IRF8 contrôle la probabilité de différenciation en plasmocytes en agissant notamment sur la division cellulaire des LB activés (220). Ce complexe maintient aussi l'expression de BCL6, PAX5 et MEF2C (myocyte-specific enhancer factor 2), tout en réprimant simultanément BLIMP1 (181).

PU.1 fonctionne de manière dose-dépendante pendant l'hématopoïèse précoce et l'expression forcée de PU.1 ou SPIB à des stades tardifs perturbe la CSR et la différenciation plasmocytaire (181, 221, 222). Il a été montré que la concentration de PU.1 est notamment contrôlée par le microARN miR-155, qui se lie à la région 3' UTR de PU.1 (223). Ce contrôle semble être physiologiquement important puisque la délétion spécifique de la séquence 3'UTR de *SPI1* reconnue par miR-155 multiplie par deux la concentration de PU.1. Il en résulte une augmentation de l'expression de PAX5 et une diminution des réponses immunitaires T-dépendantes *in vivo*, ainsi qu'une altération de la différenciation plasmocytaire *in vitro* (223).

Plus récemment, il a été montré que les souris KO pour *PU.1* et *SpiB* dans les LB matures ne génèrent pas de CG et d'anticorps à haute affinité après immunisation (224). Ces LB double KO présentent un défaut de survie après engagement des récepteurs CD40 ou TLR (toll-like receptor), malgré une différenciation plasmocytaire paradoxalement augmentée. Dans ce modèle, PU.1 et SpiB régulent l'expression de nombreux composants de la voie de signalisation du BCR et les récepteurs des ligands CD40L, BAFF et TLR. Ces résultats suggèrent que PU.1 et SpiB sont primordiaux pour permettre aux LB de répondre de manière appropriée aux signaux environnementaux.

Un troisième membre de la famille ETS, ETS1, joue également un rôle important dans la différenciation terminale des LB. Les LB déficients en ETS1 ont une capacité considérablement accrue à se différencier en plasmocytes en présence de CpG (225), qui n'est normalement pas un inducteur efficace de la différenciation plasmocytaire. Comme pour PU.1, ETS1 a pour fonction de promouvoir l'expression de PAX5 et d'inhiber simultanément BLIMP1, bien qu'il soit proposé dans ce cas que l'inhibition soit post-traductionnelle (226). Différents membres des familles ETS et IRF sont cruciaux pour la production normale des plasmocytes.

Le rôle des facteurs ETS dans la transformation tumorale est discuté dans la section 5.

#### **h. Acteurs de la différenciation plasmocytaire**

La formation de plasmocytes nécessite que le programme transcriptionnel des LB soit complètement éteint, permettant l'induction d'un transcriptome de type plasmocytaire. Ce dernier est orienté à la fois vers la production d'une grande quantité d'Ig ainsi que vers le homing de ces cellules vers la MO et leur





survie dans leur niche médullaire (31). Cette transition est réalisée par l'action de trois FTs principaux - IRF4, BLIMP1 et XBP1 (X-box-binding protein 1) (**Figure 5**)- dont les fonctions sont décrites ci-dessous.

### i. IRF4

IRF4 est essentiel à la fois au stade de LB matures, pour la formation du CG et la CSR, mais aussi pour la différenciation en ASC (108, 109, 227, 228). IRF4 a des niveaux d'expression différents au cours du développement lymphocytaire B, avec un pic d'expression dans les plasmocytes. Il a un profil d'expression bimodal dans le CG ; son expression est induite dans les cellules initiatrices du CG, dans lesquelles il contribue à l'expression de BCL6, puis elle est réprimée dans la zone sombre et ensuite ré-induite dans une proportion de cellules de la zone claire destinées à se différencier en ASC (228, 229). IRF4 se lie faiblement à l'ADN lorsqu'il est seul, mais sa liaison est plus forte en présence de SPI1/PU.1. Ses autres partenaires sont BATF (basic leucine zipper transcriptional factor ATF-like), composant du complexe AP-1 (activator protein 1) (230-232) et NFAT (nuclear factor of activated T cells) (233). Le complexe IRF4-BATF joue probablement un rôle important dans les LB activés car les souris dépourvues de l'un ou l'autre facteur présentent des défauts similaires en matière de formation du CG et de CSR (108, 109, 234). Les complexes IRF4-ETS sont impliqués dans la régulation des loci *IGH* et *IGL* (235), et se lient également à l'enhancer de *PAX5* (236). IRF4 fonctionne de manière dose-dépendante ; de faibles niveaux d'expression d'IRF4 favorisant la formation du CG et la CSR, par l'activation de *Aicda*, *Pou2af1* et *Bcl6* (109, 227), alors que des quantités élevées d'IRF4 répriment *Bcl6* et activent à la fois *Blimp1* et *Zbtb20* (BCL-6-related transcription factor zinc finger and BTB domain-containing protein 20) (109), facilitant la différenciation en ASC.

### ii. BLIMP1

BLIMP1 (B lymphocyte induced maturation protein 1) est un répresseur transcriptionnel codé par le gène *PRDM1* (PR domain containing 1, with ZNF domain) et membre de la famille des FTs Krüppel-type zinc finger. Il a été cloné et initialement décrit comme une protéine induite au cours de la différenciation des plasmocytes, capable d'induire, dans des cellules de la lignée de lymphome BCL1-3B3, un phénotype plasmocytaire caractérisé par l'expression de la chaîne J et de CD138, par la synthèse d'immunoglobulines IgM et par la régulation positive des gènes impliqués dans la sécrétion des Ig (237). Dans la lymphopoïèse B, BLIMP1 est exclusivement exprimé dans les ASC (plasmablastes et plasmocytes) ; les plasmablastes se distinguant des plasmocytes à longue durée de vie par leur expression inférieure





de BLIMP1, chez la souris et chez l'homme (238-240). Bien que l'expression de BLIMP1 soit connue pour nécessiter l'action d'IRF4 et de FOS, facteur du complexe AP-1 (109, 241, 242), les facteurs qui contrôlent le niveau des concentrations de BLIMP1 dans les ASC restent mal connus.

BLIMP1 a de nombreuses fonctions dans les ASC. Il réprime les principaux régulateurs du programme transcriptionnel des LB, notamment les gènes qui codent pour SPIB, BCL6, ID3, MYC et PAX5 (213, 243). Bien qu'il soit essentiel à la formation des ASC (244), BLIMP1 n'est pas indispensable à l'initiation de leur programme transcriptionnel, une population de pré-plasmablastes pouvant se développer en son absence (143). En l'absence de BLIMP1 fonctionnel, l'expression de cibles telles que PAX5 et BCL6 sont régulées à la baisse. Il n'existe également qu'une compréhension limitée de la fonction physiologique de BLIMP1 dans les plasmocytes matures. Des études suggèrent qu'il joue un rôle dans la présentation de l'Ag, dans les réponses au stress cellulaire et la production d'anticorps (245-247).

BLIMP1 est exprimé dans toutes les tumeurs plasmocytaires, notamment le myélome multiple, et il semble jouer un rôle essentiel dans la progression de la maladie dans des modèles murins (248). Cependant, des mutations récurrentes de *PRDM1*, principalement inactivatrices, sont décrites dans les DLBCL de type ABC (activated B-cell) (249) et dans le myélome (250), suggérant que ce FT puisse avoir aussi une fonction suppressive de tumeur. En faveur de cette hypothèse, la délétion conditionnelle de *Prdm1* chez la souris, spécifiquement au stade de CG, conduit au développement de DLBCL de type ABC. Cette activité opérerait en synergie avec l'activation constitutive de la voie canonique NF- $\kappa$ B pour accélérer la lymphomagenèse (251, 252).

### iii. XBP1

XBP1 (X-box-binding protein-1) appartient à la famille des FTs de type leucine zipper CREB/ATF.

Dans les LB, PAX5 réprime *XBP1*, la régulation à la baisse de l'expression de PAX5 pendant la différenciation plasmocytaire contribuant à l'activation de *XBP1* (145). Des études récentes utilisant l'inactivation spécifique de *Xbp1* dans des LB murins ont toutefois montré que les cellules plasmocytaires se forment relativement normalement en l'absence de *Xbp1* (253, 254). Au lieu de cela, XBP1 fonctionne principalement en promouvant la traduction de l'ARNm des gènes *IGH*, la sécrétion d'Ig et le remodelage associé au réticulum endoplasmique (RE) qui est caractéristique des plasmocytes (253, 254). Le FT XBP1 est induit par le stress du RE dans de nombreux types cellulaires, y compris les plasmocytes, et est un acteur de la réponse UPR (unfolded protein response) (255). La réponse UPR est induite suite à un



stress du RE résultant de l'accumulation de protéines mal formées. En raison de leurs taux majeurs de synthèse et de sécrétion d'Ig, les ASC sont particulièrement sensibles au stress du RE.

#### i. IRF4, BLIMP1, XBP1 et transformation tumorale

Deux mécanismes mutuellement exclusifs, convergeant vers la régulation négative de *PRDM1*, ont été décrits dans les DLBCL-ABC, pouvant expliquer l'incapacité des LB tumoraux à se différencier en ASC (249, 251, 256). D'une part, une inactivation bi-allélique de *PRDM1* est observée dans environ 30% des cas de DLBCL-ABC. D'autre part, la dérégulation de *BCL6* par translocations chromosomiques, plus fréquentes dans les DLBCL-ABC que -CG (257), conduit également à une répression constitutive de *PRDM1* (177). Compte tenu de la variété des lésions génétiques pouvant modifier l'expression et l'activité de *BCL6* (cf. paragraphe 2.g.vi), la répression de *PRDM1* pourrait être encore plus fréquente. Enfin, environ 25% des cas de DLBCL-ABC ont des anomalies gain de fonction, principalement sous la forme de translocations, de SPIB (258, 259), qui est un FT de la famille ETS qui peut former un complexe avec IRF4 et contribue à l'inactivation de *PRDM1* en réprimant directement sa transcription (221, 260, 261). Les souris inactivées pour *Prdm1* dans les LB du CG développent des DLBCL-ABC caractérisés par une activation constitutive de NF-κB, suggérant que ces deux événements sont nécessaires pour la transformation tumorale de ce type de lymphome (262).

Les études de NGS (250, 263) ont révélé que *XBP1*, *IRF4* et *BLIMP1* étaient tous mutés dans le myélome, ce qui est relativement surprenant puisque ces trois FTs sont considérés comme indispensables à la survie des cellules myélomateuses. Bien que la nature des mutations d'*IRF4* ne soient pas clairement établies, il a été démontré qu'*IRF4* était essentiel à la survie des cellules myélomateuses par un mécanisme appelé « non-oncogene addiction » (264). Les mutations de *BLIMP1* et *XBP1* sont elles clairement délétères, suggérant que ces FTs ne sont pas absolument nécessaires à la survie des cellules myélomateuses.

#### i. Transition plasmablaste-plasmocyte

Beaucoup de recherches se sont concentrées sur la description des facteurs qui contrôlent l'engagement initial vers la différenciation terminale ; cependant, moins de données sont disponibles concernant la transition plasmablaste (cellule en division et à courte durée de vie)-plasmocyte (cellule ne cyclant



plus et à longue durée de vie). Chez la souris et l'homme, les plasmablastes et les plasmocytes peuvent 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 (238, 240). Les plasmocytes éteignent ainsi l'expression de MYC, qui est une cible de la répression transcriptionnelle médiée par BLIMP1, et une partie du programme transcriptionnel induisant le cycle cellulaire (243, 265).

#### **j. Homing et survie des plasmocytes**

Même si les plasmocytes existent dans de multiples organes lymphoïdes de l'organisme et peuvent être observés dans des organes non lymphoïdes au cours de situations pathologiques, la majorité de ces cellules se situe dans la MO chez les sujets sains. Leur propriété de longévité ne semble pas être purement intrinsèque puisque le transfert de ces cellules en dehors de leur microenvironnement médullaire entraînent une mort rapide (266).

##### **i. Homing**

Le processus d'adressage et de rétention des ASC dans la MO reste mal compris. L'activation de S1PR1 (sphingosine-1-phosphate receptor 1) est nécessaire pour la circulation des ASC des organes lymphoïdes secondaires vers le sang périphérique, puis CXCL12 (chemokine CXC-chemokine ligand 12) et son récepteur CXCR4 (CXC-chemokine receptor 4) sont importants pour le recrutement des ASC dans la MO et leur maintien sur ce site (267, 268). Les chimiokines inflammatoires, telles que CXCL9, CXCL10 et CXCL11 – dont le récepteur est CXCR3 - favorisent l'adressage des plasmablastes vers les sites inflammatoires, y compris d'organes non lymphoïdes (269). Enfin, le maintien des ASC dans la MO et la maturation en plasmocytes implique la présence de VLA4 (very late antigen 4), CD44, CD28 et CD93 sur les plasmocytes, ainsi que l'expression des FTs KLF2 (Kruppel-like factor), ZBTB20 et IKZF3 (270-272).

##### **ii. Niches**

Des études longitudinales ont démontré que le nombre de "niches" médullaires n'est pas infini (52, 273). Ces niches se composent au moins d'une cellule stromale CXCL12+VCAM1+ et d'une cellule hématopoïétique qui sécrète un facteur de survie des LB, APRIL (a proliferation-inducing ligand). Des données récentes suggèrent que la cellule éosinophile puisse être cette cellule hématopoïétique, bien qu'il soit possible que d'autres types cellulaires puissent remplir ce rôle (274, 275). Les arguments pour un rôle des éosinophiles sont d'une part l'observation, chez les souris dépourvues d'éosinophiles, d'anomalies de maturation

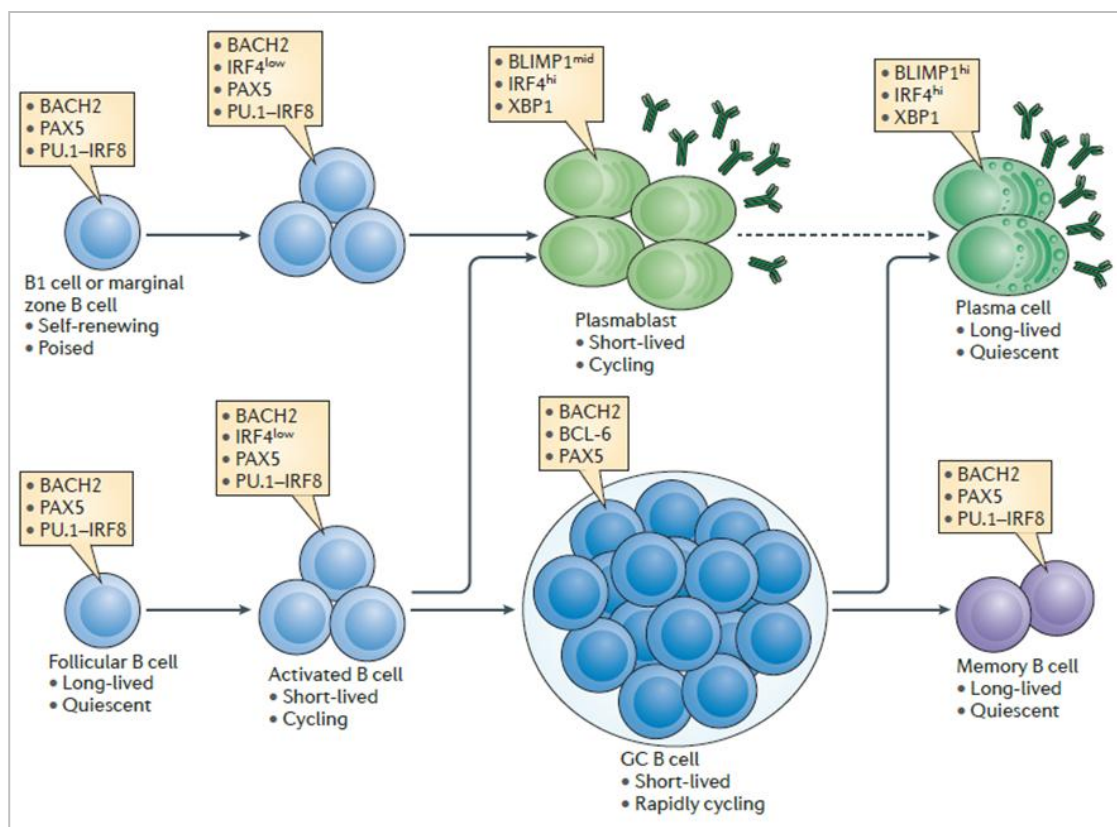


des plasmocytes médullaires et d'une diminution de leur nombre; et d'autre part, la sécrétion par les éosinophiles d'APRIL et d'IL-6 (274).

### iii. Survie

Plusieurs facteurs ont été décrits comme garant de la survie des plasmocytes dans la MO, notamment l'IL6, le TNF (tumor necrosis factor) et APRIL (276-278). APRIL et son récepteur, BCMA (B-cell maturation antigen ; également connu sous le nom de TNFRSF17), sont les composants de la niche les plus importants sur le plan fonctionnel (279, 280). La façon dont l'expression de BCMA est régulée dans les plasmocytes n'est pas claire, bien que le processus soit connu pour être indépendant de BLIMP1 (280). Une conséquence clé de la signalisation par l'axe APRIL-BCMA est l'expression de la protéine anti-apoptotique MCL1 (myeloid cell leukaemia 1), qui est essentielle à la survie de tous les ASC (280). La très courte demi-vie de MCL1 suggère un modèle dans lequel l'exposition continue à des signaux provenant de la niche, y compris APRIL, est nécessaire pour maintenir les niveaux d'expression de MCL1, rendant ainsi la survie du plasmocyte très dépendante de sa localisation. Il est intéressant de noter que le rôle anti-apoptotique d'APRIL et de l'IL6 nécessite l'activité de l'oxyde nitrique synthase inductible, ce qui suggère que le NO est un intermédiaire de signalisation dans la survie de l'ASC (281). Les plasmocytes expriment également le récepteur inhibiteur FcγRIIb, qui induit l'apoptose des plasmocytes, soulevant l'hypothèse d'une boucle de régulation négative dans laquelle les Ac sériques affectent l'homéostasie des plasmocytes (282).





**Figure 5. Différents stades de la différenciation lymphoïde B tardive** (54). La majorité des LB matures se trouvent dans les follicules des organes lymphoïdes et sont appelées LB folliculaires. Il existe d'autres populations spécialisées de LB : les cellules de la zone marginale, qui se trouvent dans la région située entre la pulpe rouge et la pulpe blanche de la rate, et les cellules B1, décrites chez la souris, qui se trouvent dans les cavités péritonéale et pleurale. Toutes les populations de LB matures expriment les FTs PAX5, PU.1, IRF8 et BACH2, tandis que de faibles (low) niveaux d'IRF4 sont induits par la signalisation BCR. Les LB activés par l'Ag sont capables d'une prolifération rapide, de CSR et d'une différenciation en plasmablastes de courte durée de vie, exprimant des taux élevés (high, hi) de XBP1, IRF4 et intermédiaires (mid) de BLIMP1, et en ASC. Les LB folliculaires peuvent également réguler à la hausse BCL6 et réprimer l'expression d'IRF4 pendant la réaction du CG. Les LB dotés de BCR à haute affinité quittent le CG et se différencient soit en LB mémoires, soit en plasmocytes à longue durée de vie, qui expriment des niveaux élevés de BLIMP1, IRF4 et XBP1, et produisent des quantités importantes d'Ac.

### 3. Macroglobulinémie de Waldenström

#### a. Introduction

La macroglobulinémie de Waldenström (MW) est un syndrome lymphoprolifératif B chronique (SLPC) rare, représentant 1 à 2% de toutes les hémopathies malignes. La MW fut décrite en 1944 par Jan Gosta Waldenström, qui rapportait l'observation de deux patients avec épistaxis, anémie, adénopathies, fibrinopénie et hypergammaglobulinémie. La MW est définie comme un sous-type de lymphome

lymphoplasmocytaire (LLP), infiltrant la MO et associé à la sécrétion d'une Ig monoclonale de classe M (IgM) détectable dans le sérum (55). Les LLP sont des hémopathies lymphoïdes B matures indolentes, ne répondant pas aux critères des autres néoplasies lymphoïdes B à petites cellules. Ils sont caractérisés par la présence d'un mélange de cellules allant du LB au plasmocyte, en passant par une importante fraction de cellules intermédiaires n'ayant pas achevé leur différenciation terminale, dénommées lymphoplasmocytes. L'association avec une IgM sérique monoclonale définissant la MW est retrouvée dans environ 90% des cas, mais il existe d'autres types de LLP produisant une Ig monoclonale d'une autre classe (le plus souvent A ou G), voire non sécrétants. De même, la présence d'une IgM monoclonale, quelle que soit sa concentration, ne peut pas être considérée comme un critère diagnostique suffisant pour une MW, car elle peut être observée dans d'autres situations comme les gammopathies monoclonales de signification indéterminée (GMSI) IgM ou d'autres SLPC B (lymphome de la zone marginale (LZM), LLC, exceptionnellement myélome à IgM).

La médiane d'âge est de 72 ans. Une estimation exacte de la prévalence de la MW est difficile, car les critères diagnostiques sont longtemps restés mal définis, et également en raison de l'existence de formes asymptomatiques. La présentation clinique est protéiforme en raison des propriétés physicochimiques ou de l'activité anticorps de l'IgM monoclonale. Cette hémopathie est actuellement incurable en dehors d'une stratégie d'allogreffe de moelle et le but du traitement est d'améliorer la durée et la qualité de vie. La médiane de survie est de 5 à 8 ans selon les études, mais 20% des patients survivent plus de dix ans, et 10 à 20% meurent d'une autre cause.

### **b. Epidémiologie**

La MW est une hémopathie maligne rare, ne représentant que 1 à 2% de toutes les hémopathies, avec une incidence de 3 à 4 cas par million de personnes et par an (283). En raison de sa rareté, l'acquisition de données épidémiologiques fut et reste difficile. Une grande partie des informations disponibles provenait initialement de petites séries cliniques, mais ces données ont été récemment remplacées par des études de population à large échelle, notamment aux Etats-Unis. La MW est une maladie du sujet âgé, avec moins de 1% des patients diagnostiqués avant l'âge de 40 ans (284). L'âge médian au diagnostic est d'environ 72 ans. L'incidence augmente fortement avec l'âge jusqu'à 60-69 ans puis plus lentement au-delà de 70 ans (285). Les sujets noirs ont l'âge médian le plus jeune au moment du diagnostic (63 ans) alors que celui des sujets blancs est le plus élevé (73 ans) (286). La MW est jusqu'à deux fois plus fréquente chez les hommes. La prédominance masculine est constante dans toutes les catégories d'âge (285).



L'incidence est très variable d'un continent à l'autre, la MW étant jusqu'à 10 fois moins fréquente dans certains pays d'Asie (287). Cependant, cette incidence est plus faible que celle des immigrants asiatiques aux États-Unis (288), confirmant l'importance des facteurs liés à l'environnement et au mode de vie dans la pathogenèse de la MW.

### c. Facteurs de risque et prédisposition

L'étiologie de la MW est en grande partie inconnue. De nombreuses études ont évalué le rôle des facteurs environnementaux et biologiques sur le risque de développement de la MW. Le facteur de risque le plus connu est la présence d'une GMSI IgM, des antécédents personnels de GMSI IgM étant associés à un risque accru (46 fois supérieure) de développer une MW (289). D'après les données de suivi à long terme de la Mayo Clinic (Rochester, E-U), un antécédent personnel de GMSI IgM est associé à un risque annuel de 1 à 1,5% de développer une MW et à un risque supplémentaire important de développer une LLC ou un autre SLPC B associé à une IgM (290). Les GMSI IgM peuvent ainsi être considérées comme un état précoce de la MW, plutôt que comme une cause de la maladie, comme cela est observé pour les GMSI IgG et IgA dans le myélome multiple (MM) (291).

La stimulation antigénique chronique est susceptible de jouer un rôle dans le développement de la MW. Les cellules B tumorales de MW portent des mutations somatiques *IGHV*, suggérant que ces cellules proviennent de LB qui ont expérimenté une stimulation antigénique répétée et une sélection dans les CG des follicules lymphoïdes (292-294). Un certain nombre d'études épidémiologiques (295-298) ont ainsi évalué le rôle de divers types de conditions de stimulation antigénique chronique par rapport au risque de développer une MW. Un risque accru de MW était associé à des antécédents personnels de maladies auto-immunes (sclérodermie, lupus érythémateux, syndrome de Sjögren, anémie hémolytique auto-immune, polyarthrite rhumatoïde et maladie de Horton) et infectieuses (infection par le VHC, pneumonie, septicémie, pyélonéphrite, sinusite, herpès zona et grippe). Le lien observé entre la MW et les infections/maladies auto-immunes suggère l'implication éventuelle d'une perturbation du système immunitaire comme événement initial du développement de la MW [46]. La stimulation chronique des LB dans ce contexte pourrait favoriser la survenue d'événements mutationnels à l'origine de la transformation tumorale, conduisant au développement de la MW.

Enfin, l'existence de formes familiales est connue de longue date : il existe pour 20% des patients d'autres cas de MW ou de SLPC B dans la famille. De plus, la prévalence d'une IgM monoclonale est plus





élevée dans les familles de patients ayant une MW : elle est estimée à 0,25-0,64% dans la population générale et à 3,2-6,3% dans les familles de patients ayant une MW (299). Dans une vaste étude cas-témoin portant sur 2144 patients ayant un LLP/WM (1539 WM [72 %] et 605 LLP [28%]) et plus de 8000 parents au premier degré et plus de 30 000 témoins (299), les parents au premier degré des patients atteints de LLP/MW présentaient un risque 20 fois plus élevé de développer une LLP/MW et, en outre, un risque accru de développer un autre SLPC B, une LLC ou une GMSI.

Les études de liens génétiques dans les cas de MW familiales sont rares. Des liens forts impliquant les chromosomes 1q, 3q et 4q ont été identifiés en compilant des cas familiaux de MW et de GMSI IgM (300). Une analyse de séquençage de l'exome entier a également été réalisée sur 4 membres d'une famille atteints de MW et 246 cas contrôles sporadiques ou familiaux. Deux variants constitutionnels *LAPTM5*<sup>c403t</sup> et *HCLS1*<sup>g496a</sup> ont été identifiés comme présents de façon récurrente dans les cas familiaux (301). Une seule étude toute récente d'association pangénomique (GWAS, genome-wide association sequencing) a été entreprise à ce jour dans la MW sur 530 cas et 4360 contrôles. Elle a identifié deux polymorphismes de haut-risque en 6p25.3 (près d'*EXOC2* et *IRF4*) et 14q32.13 (près de *TCL1*) (302)

Une autre approche de la prédisposition à la MW et aux autres hémopathies lymphoïdes B a été l'étude des cibles protéiques des Ig monoclonales (dénommées « paratargs »), sécrétées par le clone tumoral. Il a ainsi été démontré que 11% des IgM monoclonales de cas de MW/GMSI IgM réagissaient spécifiquement contre la forme phosphorylée de la paratarg-7 (pP-7) (303). Les analyses familiales de parents de patients atteints de MW/GMSI IgM réagissant contre la pP-7 ont montré que l'état hyperphosphorylé de la P-7 est hérité de façon autosomale dominante. Seulement 2% des sérums de sujets sains réagissent contre la pP-7. Les porteurs de pP-7 ont un risque 6 fois plus élevé de développer une MW/GMSI IgM.

#### **d. Aspects cliniques**

Les manifestations cliniques sont liées soit à l'infiltration tumorale par les cellules lymphoplasmocytaires, soit aux propriétés physicochimiques et antigéniques de l'IgM (304, 305). Elles sont résumées dans le tableau 2.





**Tableau 2.** Manifestations clinico-biologiques dans la MW

Signes généraux	Sueurs, fièvre, arthralgies Altération de l'état général
Manifestations liées à l'IgM	Syndrome d'hyperviscosité Troubles de l'hémostase (Willebrand acquis) Neuropathie périphérique Cryoglobulinémie Maladie des agglutinines froides Amylose AL
Manifestations liées à l'infiltration tumorale	Adénopathies périphériques, hépato-splénomégalie Cytopénies

### e. Critères diagnostiques

La MW est caractérisée par une infiltration faite de cellules lymphoplasmocytaires et la production d'une IgM monoclonale. Les critères diagnostiques sont les suivants : i) IgM monoclonale sérique, quelle que soit sa concentration ; ii) infiltration médullaire par des petits lymphocytes avec différenciation plasmocytaire ; iii) infiltration histologique médullaire souvent diffuse ; iv) phénotype des cellules tumorales : IgM+, CD5-/+ , CD10-, CD19+, CD20+, CD22+, CD23-, CD25+, CD27+, FMC7+, CD103-.

Le diagnostic est effectué sur l'analyse médullaire, habituellement sur la biopsie ostéo-médullaire. Un immunophénotypage lymphocytaire permet de caractériser cette population lymphoplasmocytaire et permet d'éliminer d'autres hémopathies lymphoïdes B associées à une IgM monoclonale. Une analyse cytogénétique est également effectuée, retrouvant une anomalie dans 45 à 50% des cas (cf. paragraphe 4.b). L'analyse moléculaire comprend la recherche de la mutation du gène *MYD88*. En effet, celle-ci est présente dans 90 à 95% des patients ayant une MW (94, 306).

### i. Caractéristiques morphologiques

La MW est caractérisée par une infiltration médullaire faite de cellules lymphoïdes et lymphoplasmocytaires et de plasmocytes. L'atteinte médullaire est caractérisée par un infiltrat nodulaire, diffus et/ou interstitiel, constitué de petits lymphocytes prédominants, mêlés en proportions variables à des lymphoplasmocytes et plasmocytes. L'infiltration médullaire est variable et représente 20 à 100% des cellules médullaires,



au sein desquelles il est fréquent d'observer la présence de mastocytes, en proportion augmentée chez environ un quart des patients (307). Dans la plupart des cas, les ganglions ont partiellement conservé une architecture normale, avec des sinus dilatés contenant du matériel PAS+ (*periodic acid Schiff* positif) et parfois quelques petits centres germinatifs résiduels. On trouve une population lymphoïde relativement monotone de petits lymphocytes, lymphoplasmocytes et plasmocytes et très peu de grandes cellules. Dans d'autres cas, l'architecture ganglionnaire est totalement détruite, avec souvent une plus grande proportion de cellules plasmocytaires et/ou de grandes cellules. Dans la rate, on peut observer un infiltrat lymphoplasmocytaire formant des nodules dans la pulpe rouge ou se développant de façon plus diffuse.

## ii. Caractéristiques immunophénotypiques

Des études immunophénotypiques sont nécessaires pour un diagnostic précis de la MW, et ceci peut être réalisé soit par une analyse en cytométrie de flux, soit par une analyse immunohistochimique, bien que la première méthode permette une évaluation plus approfondie des déterminants antigéniques (55). Dans la MW, il est généralement possible de démontrer le caractère monotypique à la fois des LB et des plasmocytes, mais un phénotypage étendu n'est habituellement effectué que sur la composante lymphocytaire B.

Les LB clonaux de la MW sont généralement caractérisés par l'expression constante de marqueurs pancellulaires B (CD19, CD20, CD22) ainsi que par l'expression monotypique d'IgM de surface (slg), restreinte à la chaîne légère  $\kappa$  ou  $\lambda$ . La prédominance de la chaîne légère  $\kappa$  sur  $\lambda$  est plus élevée dans la MW ( $\kappa:\lambda$ , ratio de 5:1) comparée aux autres SLPC B (308). Par rapport aux LB normaux matures, l'expression du CD22 dans la MW est constamment plus faible, tandis que celle du CD20 et du slg est habituellement équivalente. Les LB clonaux sont positifs pour FMC7 et CD25 dans 70% et 66% des cas, respectivement. FMC7 est généralement hétérogène, tandis que le CD25, lorsqu'il est présent, est exprimé de façon homogène par pratiquement tous les LB tumoraux. Bien qu'il ne soit pas entièrement spécifique (p. ex., LZM ou LLC), le profil d'expression phénotypique aberrant CD22 faible CD25+ est évocateur de la MW. L'expression du CD103 et du CD10, informatifs respectivement pour le diagnostic différentiel avec la leucémie à tricholeucocytes et le lymphome folliculaire, est constamment négative dans la MW. CD11b est absent et CD11c est rarement positif et, s'il est exprimé, c'est à une faible intensité (308, 309). Le CD5 et le CD23, deux antigènes classiquement exprimés dans les cas de LLC, ne sont présents que chez moins d'un cinquième des patients MW et lorsqu'ils sont présents, ils sont d'expression hétérogène. Le CD38 est présent dans la moitié des cas de MW et son expression est plus faible que celles observées



sur les progéniteurs lymphoïdes B et les plasmocytes. Les LB de la MW sont aussi typiquement positifs pour CD24, HLA-DR et BCL2, alors que l'antigène CD27 associé à un phénotype mémoire est exprimé de manière hétérogène (de négatif à positif) chez plus de la moitié des patients (308, 309). BCL2 et PAX5 présentent généralement une immunoréactivité nucléaire, ce qui nécessite de les évaluer par immunohistochimie sur des coupes tissulaires (310). Le CD79b et le CD81 sont constamment positifs bien que des niveaux d'expression plus faibles puissent être observés dans certains cas ; le CD200 est positif chez 60 % des patients, et le CD305 (LAIR1) est généralement négatif (311). Ce dernier marqueur est important, car les LB normaux sont généralement positifs pour le CD305. En conséquence, le profil d'expression immunophénotypique le plus courant du clone de la MW pourrait être résumé comme suit : CD5-, CD10-, CD11c-, CD19+, CD20+, CD22 faible, CD23-, CD25+, CD27 hétérogène (het), CD38het, CD79b+, CD81+, CD79b-, CD103-, CD200het, CD305-. La nature monotypique des LB doit être démontrée par l'expression restreinte de la chaîne légère intracytoplasmique ou de surface ( $\kappa$  ou  $\lambda$ ). Environ 90% des cas expriment CD22 faiblement, CD25, CD27 et l'IgM au contraire du CD5, CD10, CD11c, CD23 et CD103 qui sont négatifs. Ces éléments fournissent des critères diagnostiques solides et permettent en particulier de faire la distinction avec un LZM, dont l'immunophénotype est le plus souvent CD22+CD25-CD103+.

#### f. Critères pronostiques

Le score IPSS (*International Prognostic Staging System*) a été proposé par Morel et al. (312) pour les patients avec une MW symptomatique. L'analyse de 587 cas de MW a permis d'identifier cinq facteurs ayant un impact sur la survie des patients : âge > 65 ans, hémoglobine  $\leq$  115 g/L, plaquettes  $\leq$  100 G/L, béta-2-microglobuline > 3 mg/L, IgM monoclonale > 70 g/L. Ces facteurs permettent de déterminer 3 groupes pronostiques différents en matière de médiane de survie globale. Récemment, le génotype *MYD88* non muté / *CXCR4* non muté a été associé à une survie globale inférieure et à une résistance à certaines thérapeutiques (313-315).

#### g. Thérapeutiques

##### i. Indication

Un traitement est initié uniquement chez les patients présentant une maladie symptomatique, que les symptômes soient liés aux propriétés antigéniques ou physico-chimiques de l'IgM ou au syndrome tumoral. En effet, il n'a jamais été démontré de bénéfice au traitement d'une MW asymptomatique. Celle-ci peut rester stable de nombreuses années. La surveillance est donc préconisée chez ces patients. Au



diagnostic, 30 à 50% des patients sont asymptomatiques et ne nécessitent pas de traitement. Le risque de progression à 5 ans est de 59%. La présence d'une anémie est l'indication de traitement la plus fréquemment retrouvée.

## ii. Choix du traitement

La difficulté dans le choix entre les différentes options thérapeutiques réside dans le manque d'essais prospectifs randomisés. Les traitements ont longtemps reposé sur l'utilisation des agents alkylants (chloraminophène, cyclophosphamide) et des analogues des purines (fludarabine). Bien que ces agents soient toujours utilisés, d'importants progrès ont été réalisés avec l'introduction des anticorps monoclonaux anti-CD20 (rituximab), l'utilisation de la bendamustine ou des inhibiteurs du protéasome (bortezomib) et, plus récemment, l'apparition d'inhibiteurs du BCR (ibrutinib, idelalisib).

Parmi toutes les options thérapeutiques, le choix se fait principalement en fonction de la nécessité ou non d'obtenir une réponse rapide et en fonction du profil de toxicité des différents agents thérapeutiques. Malgré ces importants progrès, peu de patients atteignent une rémission complète. Au-delà de l'augmentation de la survie, l'amélioration de la qualité de vie est donc un objectif principal.

## iii. Options thérapeutiques

Les cellules de MW exprimant l'antigène de surface CD20, le rituximab est un élément clé du traitement. Son utilisation est recommandée dans le cadre d'un traitement y associant d'autres drogues, car les taux de réponse en monothérapie sont modestes [taux de réponse globale autour de 50% et de réponse majeure autour de 25%] (316, 317). Cependant, le rituximab en monothérapie peut être utilisé chez les patients fragiles présentant des comorbidités élevées ou lorsque la symptomatologie est due à l'activité auto-immune de l'IgM (neuropathie, cryoglobulinémie) (318).

L'association rituximab, cyclophosphamide et dexaméthasone (RCD) est actuellement considérée comme le traitement standard de première ligne (319). C'est un traitement bien toléré, avec un taux de réponse majeure (incluant réponses partielle et complète) de 74% et de réponse globale de 83% (incluant réponses mineure et majeure). Les médianes de délai avant le traitement suivant, de survie sans progression (SSP) et de survie globale (SG) étaient respectivement de 51, 35 et 95 mois (320).

L'association bendamustine et rituximab (BR) semble également être une option thérapeutique intéressante, surtout pour les patients avec une forte masse tumorale ou une amylose associée. Dans une analyse de sous-groupe de patients MW d'une étude randomisée comparant BR et une polychimiothérapie de



type R-CHOP (rituximab, cyclophosphamide, adriamycine, vincristine, prednisone), l'association BR permettait d'obtenir une SSP plus longue et avait un meilleur profil de tolérance. La SSP médiane n'était pas atteinte avec BR alors qu'elle était de 40 mois pour le R-CHOP (321).

L'association bortézomib (inhibiteur réversible du protéasome), dexaméthasone et rituximab (BDR) a été évalué dans une étude récente de phase 2 incluant des patients nouvellement diagnostiqués avec des facteurs pronostiques défavorables. Le taux de réponse majeure était de 68 % et la SSP médiane de 43 mois. Le délai avant le traitement suivant était de 73 mois (322). Une étude de phase 3 est actuellement en cours, comparant le RCD avec ou sans bortézomib (NCT01788020).

Les analogues nucléosidiques des bases puriques, comme la fludarabine et la cladribine, peuvent aussi être utilisés de façon efficace dans le cadre de combinaisons. L'association fludarabine et rituximab a montré un taux de réponse globale de 95% et permettait un contrôle rapide de la maladie (323). Le chlorambucil en monothérapie peut être une option pour les patients très fragiles, pour lesquels un traitement d'association est considéré comme potentiellement trop toxique (taux de réponse globale de 38 %, SSP de 27 mois) (324). C'est un traitement bien toléré mais associé à un risque de développement de tumeurs secondaires, en particulier de leucémie et de syndrome myélodysplasique. La fludarabine en monothérapie s'est révélée plus efficace et plus sûre que le chlorambucil (324).

L'ibrutinib, inhibiteur de la tyrosine kinase de Bruton (BTK), a été évalué en monothérapie à la posologie de 420 mg/j, par voie orale. Les taux de réponse globale dépassent les 90% (315). Les meilleurs taux de réponse sont observés chez les patients présentant la mutation L265P du gène *MYD88* en l'absence de mutation du gène *CXCR4*. L'ibrutinib dispose actuellement d'une Autorisation de Mise sur le Marché, en première ligne de traitement, aux Etats-Unis et en Europe, chez des patients ne pouvant recevoir de chimiothérapie ou en situation de rechute.

L'autogreffe de cellules souches hématopoïétiques est une option thérapeutique envisageable chez des patients jeunes, en situation de rechute précoce, avec des facteurs pronostiques défavorables. Compte-tenu d'une morbi-mortalité non négligeable, la place de l'allogreffe ne fait pas l'objet d'un consensus à l'heure actuelle.

D'autres agents sont en cours d'essai ou de développement, comme le venetoclax, un inhibiteur de BCL2, ainsi que des thérapies ciblant CXCR4 ou MYD88.



#### 4. Physiopathogénie de la macroglobulinémie de Waldenström

##### a. Contrepartie physiologique

L'identification de la cellule d'origine (cellule dans laquelle va survenir le premier événement oncogénique) et de la contrepartie physiologique (compartiment cellulaire dans lequel va survenir la transformation) des cellules tumorales est importante afin de déterminer les mécanismes physiopathologiques aboutissant au développement de l'hémopathie et afin d'envisager d'éventuelles thérapeutiques ciblées. L'étude des caractéristiques morphologiques, immunophénotypiques, du statut mutationnel *IGHV*, des PEG et épigénétiques peuvent aider à appréhender cette question.

Dans le schéma de classification de l'origine des différentes hémopathies lymphoïdes B, proposé par Küppers et al. (325), les lymphomes/SLPC B sont divisés en trois groupes principaux selon leur contrepartie physiologique : (1) ceux d'origine pré-CG, dérivant de LB n'ayant pas établi de contact avec l'Ag et, par conséquent, sans SHM des gènes des Ig et sans CSR ; (2) ceux d'origine CG, dérivant de LB ayant rencontré un Ag présenté par les cellules dendritiques et ayant expérimenté un processus de SHM des gènes *IGV* ; cette étape étant suivie d'une CSR et, parfois, d'une recombinaison secondaire des gènes *VH* (les cellules de ces hémopathies du CG n'ont pas complètement terminé la réaction du CG et sont caractérisées par la persistance du processus mutationnel (« ongoing mutation »)) ; et enfin (3) ceux d'origine post-CG, dérivant de LB ayant subi toute la réaction du CG, se traduisant par un taux élevé de SHM (sans « ongoing mutation ») et une CSR.

L'immunophénotype des différentes populations de la lymphopoïèse B du stade le plus précoce au plus mature a été déterminé chez > 600 sujets sains et les résultats ont été comparés à ceux de > 250 individus présentant une GMSI IgM et/ou un SLPC B (308, 309, 311, 326-328). L'absence d'expression ou la faible expression par les cellules de MW du CD10, CD22 ou CXCR4 n'étaient pas en faveur d'une origine CG. Leur phénotype et leur statut *IGHV* muté évoquent plutôt qu'elles sont issues d'un LB post-CG, soit d'un LB en voie de différenciation vers le plasmocyte soit d'un LB un peu plus immature. Les cellules tumorales de MW ont une différenciation intraclonale avec des petits lymphocytes exprimant des Ig de surface, des lymphoplasmocytes et des plasmocytes avec des Ig intracytoplasmiques. Deux études rapportent la séquence complète du gène *IGH* réarrangé de 22 cas de MW (329, 330). Comme dans la plupart des hémopathies lymphoïdes B, il existe un biais d'utilisation du répertoire des *IGH*, avec ici une surreprésentation de la famille *IGHV3* et du gène *IGHJ4*. Dans 91% des cas, le gène est muté, présentant moins de 98% d'homologie (médiane 93%, 88-97) avec la séquence germinale



correspondante. Une étude menée sur 31 cas a confirmé ces données et montré l'existence d'une utilisation préférentielle du gène *IGHV3-23* (35%) (331). Les mutations *IGHV* ont une distribution et des caractéristiques qui ne témoigneraient pas toujours d'un contact antigénique. Il est intéressant de constater qu'à l'issue de leur différenciation, vraisemblablement dans les CG où elles ont subi le processus de SHM, les cellules tumorales restent bloquées avant la commutation isotypique. En effet, à la différence des cas de myélome à IgM, on ne retrouve pas d'événement transcriptionnel de CSR (présence de transcrits mu ou mu + delta et absence de transcrits alpha, gamma ou epsilon). De plus, l'étude du transcriptome des cellules tumorales montre que ces cellules ont un transcriptome plus proche des LB matures que des plasmocytes (cf. paragraphe 4.c.vii).

Ces différents éléments ajoutés à l'expression du CD27 (308, 309) par les cellules tumorales, l'absence de CSR et la monotypie des compartiments lymphoïdes et plasmocytaires (332), impliquant un certain degré de capacité de différenciation, ont fait soulever l'hypothèse que la contrepartie physiologique des cellules MW soit un LB mémoire. Bien que la plupart des LB mémoires soient générés dans le CG, une fraction mineure peut être produite par une stimulation antigénique T-indépendante à l'extérieur du CG. L'absence de marqueurs univoques capables d'identifier les LB ayant circulé dans le CG et le fait que la nature et l'origine de la population CD27+IgM+IgD+ ne soient pas encore claires laisse la question de la contrepartie physiologique du MW encore ouverte.

### **b. Anomalies cytogénétiques**

Les techniques de cytogénétique conventionnelle et d'hybridation *in situ* fluorescente (*fluorescent in situ hybridization* [FISH]) ont permis d'identifier un certain nombre d'anomalies chromosomiques récurrentes dans la MW. Ces deux techniques objectivent au moins une anomalie chez 50% des patients (333). L'utilisation de puces génomiques à haute résolution de type *single nucleotide polymorphism* (SNP)-array ou *comparative genomic hybridization* [CGH]-array permet de détecter des anomalies génomiques chez 61 à 83% des patients MW, avec une médiane de 2-3 anomalies par patient selon les séries (334, 335). Chez les patients avec une MW symptomatique, le nombre médian d'anomalies chromosomiques est supérieure à 3 (335).

La délétion d'une partie du bras long du chromosome 6 (6q) est l'anomalie la plus fréquemment retrouvée : elle est présente dans 22 à 54% des cas selon les catégories de patients étudiés (MW indolents ou symptomatiques, au diagnostic ou à la rechute) et les techniques utilisées (333-336). Quatre





régions minimales de délétion distinctes ont été identifiées en 6q, deux d'entre elles incluant des gènes suppresseurs de tumeurs, *PRDM1* (BLIMP1) en 6q21 et *TNFAIP3* en 6q23, mais leur rôle direct dans la pathogenèse de la MW n'est pas encore établi. Bien que la délétion 6q soit associée à des caractéristiques de pronostic défavorable (anémie, hypoalbuminémie, taux élevé de bêta-2-microglobuline), cette anomalie ne semble pas avoir d'impact négatif sur les devenir des patients (333, 337, 338).

La trisomie 4 est décrite dans 8 à 20% des cas de MW, correspondant soit à un gain complet soit partiel du chromosome 4 (333, 339). Aucun gène candidat n'a été identifié à ce jour sur ce chromosome (339). La trisomie 18 est également fréquente (15%), et il a été montré que trisomies 4 et 18 étaient significativement associées dans la MW (333). D'autres anomalies récurrentes sont décrites, telles que la délétion 13q14, la délétion 17p13 (impliquant le gène *TP53*), la délétion 11q22 (impliquant le gène *ATM*) et la trisomie 12, qui sont observées dans moins de 15% des cas de WM. Les translocations impliquant les gènes *IGH* en 14q32 sont très rares (2%) (333). Certaines études ont décrit un gain sur le bras court du chromosome 6 (6p), toujours secondaire à une délétion de 6q, chez 16% des patients MW (334). Le travail de Nguyen-Khac et al.(333), réalisé sur une grande série de patients non préalablement traités, a montré qu'aucune anomalie chromosomique n'influçait la survie globale. En revanche, les patients avec une délétion de *TP53* avaient une survie sans progression et sans maladie significativement plus courte. La trisomie 12, bien que rare, était aussi associée à une survie sans progression plus courte.

Il n'existe à ce jour pas d'anomalie chromosomique spécifique de la MW mais la fréquence de ces anomalies est différente entre la MW et les autres SLPC B (LZM, LLC, MM), pouvant donc être une aide au diagnostic. La délétion 6q et la trisomie 4 sont ainsi plus fréquentes dans la MW, la délétion 13q plus fréquente dans les LLC et le MM, et les translocations impliquant le locus *IGH* plus fréquentes dans le MM.

### c. Anomalies moléculaires

L'avènement des techniques de séquençage haut débit (next generation sequencing, NGS) a permis l'analyse du génome tumoral de la MW à grande échelle et l'identification d'anomalies moléculaires ayant des implications pour le diagnostic et le traitement de cette maladie. Jusqu'à présent, 30 échantillons de MW ont été analysés par WGS (94), 4 par WES (Jimenez et al., IWWM9 2016) et 305 par NGS ciblé avec des panels comprenant entre 5 et 12 gènes (340, 341) (Jimenez et al., IWWM9 2016).





### i. Mutations du gène *MYD88*

Le séquençage du génome entier (*whole genome sequencing*, WGS) de cellules CD19+ de la MO de 30 patients atteints de MW a conduit à la découverte d'une mutation somatique hétérozygote (c.978T>C) du gène *MYD88* (Myeloid Differentiation Primary Response 88), gène situé sur bras court du chromosome 3 (en 3p22). Cette mutation, présente chez plus de 90% des patients MW (94, 306), aboutit au remplacement d'une leucine par une proline en position 265, p.Leu265Pro (L265P).

MYD88 agit comme une protéine adaptatrice en aval des récepteurs de la famille des TLR et de l'IL1R, en se liant directement au récepteur ou, comme pour TLR4, par l'intermédiaire de la molécule adaptatrice TIRAP (TIR Domain Containing Adaptor Protein). Suite à la liaison d'un ligand à ces récepteurs, MYD88 subit un changement conformationnel favorisant son homodimérisation puis le recrutement de protéines de signalisation comme IRAK4 et IRAK1 (342, 343), concourant à la formation d'un complexe multi-protéique appelé Myddosome, qui déclenche une signalisation d'aval aboutissant à l'activation de NF-κB (344). La mutation L265P se produit au niveau d'un résidu hautement conservé d'un feuillet β du noyau hydrophobe du domaine TIR de la protéine. Des analyses biochimiques prédisent que cette mutation altérerait la structure tertiaire du domaine TIR. D'autres mutations de *MYD88*, beaucoup moins fréquentes, identiques à celles observées dans les DLBCL, ont été décrites dans la MW. Elles se situent également pour la très grande majorité sur les séquences codant pour le domaine TIR de la protéine. Les mutations somatiques de *MYD88* induisent de façon constitutive son homodimérisation et sa signalisation d'aval, indépendamment de l'activation des récepteurs (92, 94, 345). Cela conduit au recrutement et à l'activation d'IRAK1/4 puis de TRAF6 (TNF Receptor Associated Factor 6) et enfin à l'activation de NF-κB (**Figure 6**) (345, 346). La protéine mutante MYD88 L265P peut également déclencher une signalisation NF-κB, indépendante des protéines IRAK, par interaction directe avec BTK (345). L'importance de la signalisation NF-κB dans la croissance et la survie des cellules de MW a été bien documentée *in vitro*, l'inhibition de la dimérisation de MYD88 empêchant la localisation nucléaire de NF-κB et favorisant l'apoptose (94, 347). Les protéines IRAK4, IRAK1 et BTK sont toutes directement associées au sein du Myddosome et leurs niveaux de phosphorylation peuvent être réduits par l'inhibition de MYD88. L'inhibition de la kinase BTK par interférence ou knockdown est suffisante pour réduire la signalisation NF-κB et favoriser l'apoptose.

La mutation de *MYD88* peut aussi favoriser la transcription de la kinase HCK, un membre de la famille des



kinases SRC, qui peut aussi être activée par l'IL6, dont la transcription peut elle-même être activée par MYD88 muté (348). HCK contribue à la croissance et à la survie des cellules de MW via BTK et l'activation des voies de signalisation PI3K/AKT et MAPK/ERK (348).

La présence de la mutation *MYD88* L265P chez plus de 90% des patients MW a été confirmée par plusieurs groupes ayant utilisé différentes techniques (séquençage Sanger, PCR allèle-spécifique (AS-PCR), NGS) (94, 95, 349-351). Les mutations de *MYD88* sont typiquement hétérozygotes bien que des amplifications du nombre de copies et l'acquisition d'événements de disomies uniparentales (*acquired uniparental disomies*, aUPD) affectant le locus de *MYD88* aient été décrites (306, 351). Les mutations activatrices de *MYD88* ne sont pas spécifiques de la MW et ont d'abord été décrites dans les DLBCL-ABC où elles ont été observées chez 29% des patients (92). D'autre part, elles sont absentes ou exprimées à faible fréquence dans d'autres SLPC B qui partagent des caractéristiques morphologiques et clinicopathologiques similaires à la MW, notamment le myélome IgM (0%), le LZM (6 à 10%) et la LLC (3 à 8%) (95). Les mutations décrites ne se limitent pas à la mutation L265P. Dans la MW, les autres mutations de *MYD88* comprennent p.Ser219Cys, p.Met232Thr, et p.Ser243Asn, également observées dans les DLBCL. La prévalence de la mutation L265P est en revanche plus importante, supérieure à 95%, dans la MW comparée aux autres hémopathies.

Les mutations *MYD88* L265P sont le plus souvent clonales et peuvent être détectées dans environ 50 à 80% des cas de GMSI IgM en utilisant une technique d'AS-PCR (95, 349, 350). Ceci suggère que la mutation de *MYD88* L265P joue un rôle précoce dans l'oncogenèse de la MW. Il a ainsi été montré dans des modèles murins que l'expression spécifique de la forme mutante de *MYD88* au cours de la lymphopoïèse B pouvait entraîner avec une longue latence la survenue de lymphoproliférations B et plus rarement de DLBCL-ABC (352).

La présence ou l'absence de mutations de *MYD88* semble également distinguer deux populations de patients MW. Les patients pour lesquels aucune mutation de *MYD88* n'est identifiée (*MYD88* sauvage ou wild-type (WT)) présentent une maladie avec des caractéristiques histologiques similaires à celle des patients présentant des mutations de *MYD88*, mais une infiltration médullaire plus faible et des taux d'IgM sériques moins élevés (313). Même si cela nécessite confirmation sur de plus grandes séries de patients, la survie globale des patients *MYD88* WT semble plus courte que celle des patients *MYD88* mutés (313).



## ii. Mutations du gène *CXCR4*

Les mutations du gène *CXCR4* (C-X-C Motif Chemokine Receptor 4), situé sur bras long du chromosome 2 (en 2q21), sont dans la MW les deuxièmes en termes de fréquence après celles de *MYD88*. Elles sont identifiées chez 30 à 40% des patients selon les techniques utilisées (Sanger, NGS ou AS-PCR) (306, 313).

La protéine *CXCR4* est un récepteur couplé à la protéine G (GPCR) dont le ligand est le facteur-1 dérivé des cellules stromales (SDF-1 $\alpha$  ou CXCL12). Ses rôles physiologiques sont multiples : permettre la rétention des cellules souches hématopoïétiques CD34+ au sein de leurs niches dans la MO, stimuler la lymphopoïèse, réguler la circulation des lymphocytes dans le sang périphérique ou encore favoriser la survie cellulaire. *CXCR4* se compose de sept hélices transmembranaires et d'une région carboxy-terminale (C-terminale) intra-cellulaire (queue, *tail*). Les boucles des hélices forment à la fois une poche de liaison extracellulaire pour le ligand CXCL12 mais aussi un domaine de signalisation intracellulaire qui agit par l'intermédiaire des protéines G hétérotrimériques (**Figure 6**). La région C-terminale intracellulaire est importante pour la régulation et l'internalisation du récepteur mais ne participe pas à la signalisation via la protéine G (353). La liaison de CXCL12 à *CXCR4* entraîne l'activation de la sous-unité  $\alpha_i$  de la protéine G (G $\alpha_i$ ) en catalysant l'échange d'un GDP pour un GTP, ce qui libère l'hétérotrimère de *CXCR4* et dissocie l'hétérodimère G $\beta\gamma$  de G $\alpha_i$  (354). Les sous-unités G $\beta\gamma$  activent les voies de signalisation de la phospholipase C (PLC) et de la PI3K, jusqu'à ce qu'elles soient inhibées par une sous-unité G $\alpha_i$  lié à un GDP. La sous-unité G $\alpha_i$  de son côté active la signalisation par PI3K, jusqu'à ce que le GTP soit hydrolysé en GDP par un régulateur de la famille des membres de la signalisation de la protéine G. Cette signalisation basée sur la protéine G s'achève lorsque la région C-terminale de *CXCR4* est phosphorylée et que le récepteur est internalisé. Une fois le récepteur activé, plusieurs kinases, dont la kinase du récepteur de la protéine G (GRK) et les membres de la famille de la protéine kinase C (PKC) phosphorylent les sérines de la région C-terminale. Cela provoque le recrutement de la  $\beta$ -arrestin, inhibant alors toute signalisation supplémentaire, et entraînant l'internalisation de *CXCR4*.

Les mutations de *CXCR4* sont des mutations engendrant soit un codon stop (mutation non-sens) soit un décalage du cadre de lecture (mutation frameshift) et se situent exclusivement dans le domaine C-terminal (306, 313, 355) (**Figure 7**). Plus de 30 mutations différentes ont été décrites à ce jour. Elles sont le plus souvent hétérozygotes et se produisent presque exclusivement chez les patients porteurs de mutations de *MYD88*. Contrairement aux mutations de *MYD88*, celles de *CXCR4* sont fréquemment sous-clonales (356). Différentes mutations de *CXCR4* peuvent aussi survenir chez le même patient,



éventuellement dans des clones différents (357). La nature sous-clonale des mutations de *CXCR4* suggère que ces mutations sont acquises après celles de *MYD88*. Les mutations de *CXCR4* sont quasiment spécifiques de la MW, car, à l'exception de quelques cas de LZM, elles n'ont pas été jusqu'à présent décrites dans d'autres hémopathies (356).

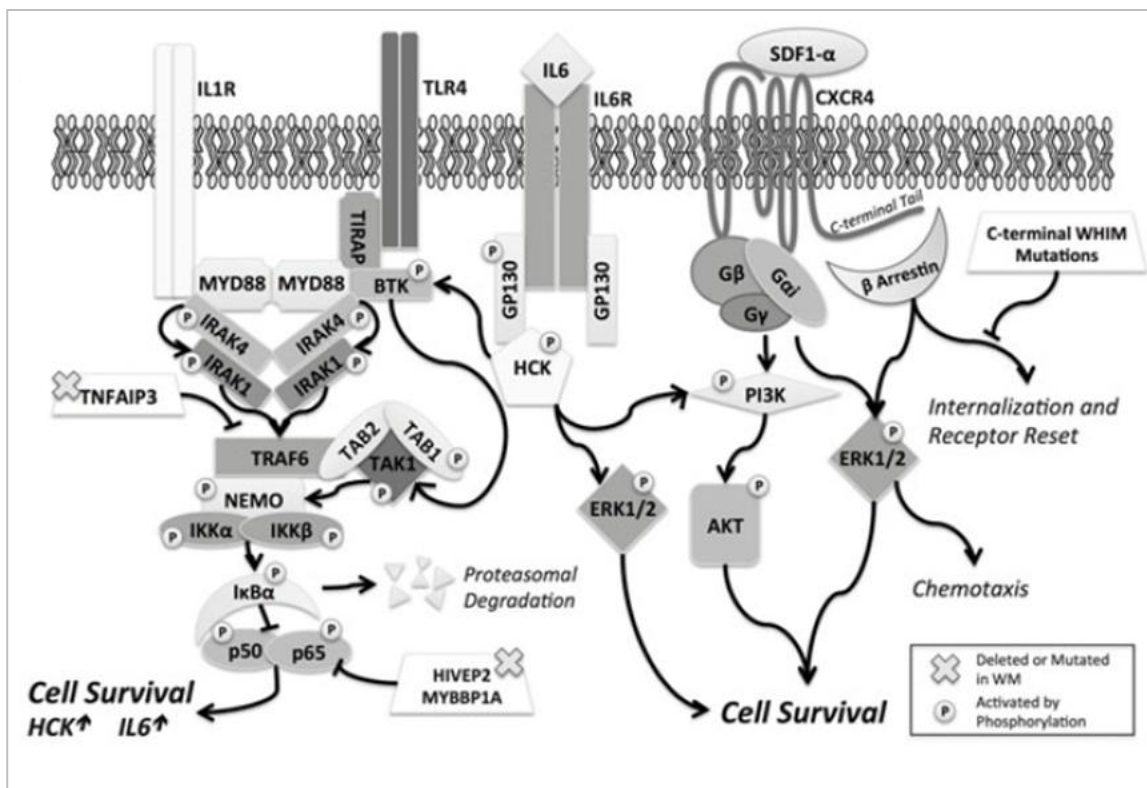
Les patients porteurs de mutations de *CXCR4* présentent moins fréquemment un syndrome tumoral ganglionnaire et ceux avec des mutations non-sens ont une infiltration médullaire plus importante, des taux d'IgM sériques plus élevés et/ou un risque d'hyperviscosité symptomatique supérieur (313, 355, 358). Malgré ces différences de présentation clinique, les mutations de *CXCR4* ne semblent pas avoir d'impact négatif sur la survie globale des patients (313, 355). Une étude récente suggère qu'elles pourraient avoir un impact négatif sur la réponse à l'ibrutinib (315).

Ces mutations sont identiques aux variations germinales observées dans le syndrome WHIM, déficit immunitaire dont l'acronyme signifie "verruques-hypogammaglobulinémie-infections-myélokathexis" (359). Dans les deux maladies, ces mutations se produisent après la dernière des sept hélices transmembranaires, aboutissant à la perte de la région C-terminale mais laissant intactes les régions responsables de la liaison du SDF-1 $\alpha$  et de la signalisation d'aval via les protéines G (360). La perte des sérines régulatrices de la région C-terminale limite l'internalisation du récepteur (353, 361, 362) mais le reste du récepteur restant intact, l'activation des voies de signalisation PI3K/AKT et MAPK/ERK est conservée et prolongée. Après stimulation avec du SDF-1 $\alpha$ , les lignées cellulaires issues de patients WM, transduites afin d'exprimer des mutants de *CXCR4*, présentaient une diminution de l'internalisation du récepteur et une augmentation de la signalisation des voies AKT et MAPK, comparativement aux lignées transduites avec un vecteur vide ou exprimant une forme sauvage de *CXCR4* (357, 363). Les cellules WM transduites avec un mutant de *CXCR4* montraient également une résistance accrue à de multiples thérapeutiques, y compris la bendamustine, la fludarabine, le bortézomib, l'idélalisib et l'ibrutinib, en présence de CXCL12 (357, 363). Cette résistance pouvait être levée par l'utilisation d'un agent bloquant la liaison de CXCL12 à *CXCR4*.

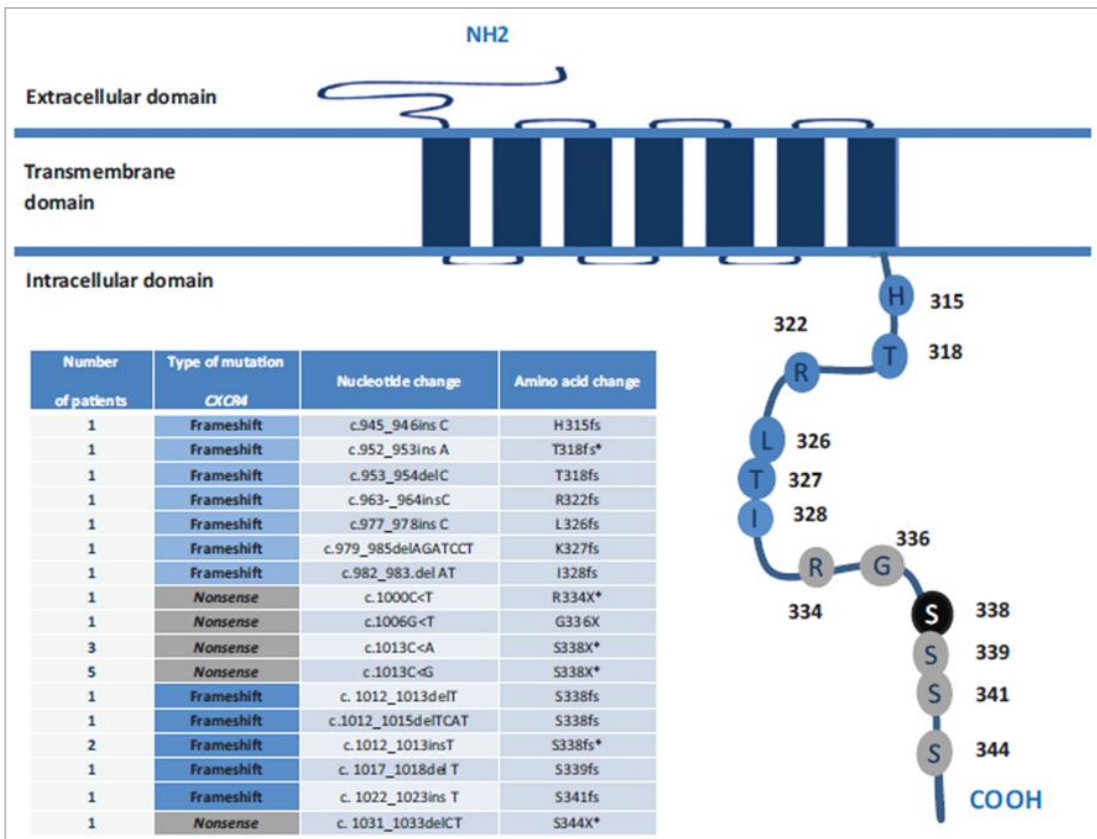
La compréhension actuelle du rôle de Cxcl12 et Cxcr4 dans la biologie des lymphocytes provient principalement d'études effectuées sur des souris déficientes, de façon constitutive ou conditionnelle, pour Cxcr4 (364-367). Le couple Cxcl12/Cxcr4 régule le trafic des LT et orchestre le homing, la maturation et la différenciation des LB dans les organes lymphoïdes secondaires ainsi que ceux des plasmocytes vers des niches spécifiques dans les organes lymphoïdes secondaires et la MO (42, 364, 366).



Des modèles de souris knockin (KI) portant une mutation hétérozygote de Cxcr4 (368) ont également été développés. La désensibilisation anormale de Cxcr4 aboutissait à des anomalies du trafic des LB matures entre la circulation sanguine et les organes lymphoïdes secondaires. En effet, les souris KI présentaient une hypoplasie folliculaire splénique sans altération de la compartimentation lymphocytaire, alors que leur architecture ganglionnaire était perturbée avec un déploiement de la zone des LT au sein des follicules primaires des LB. Malgré cela, les taux sériques d'IgM et d'IgG étaient augmentés chez les souris mutantes par rapport aux WT, suggérant des anomalies dans la qualité et le contrôle de la réponse immunitaire humorale dans ce modèle. De façon contre-intuitive, le gain de fonction de Cxcr4 inhibait le maintien des titres d'Ac après immunisation dans ce modèle KI (369). Alors que la forme mutée de Cxcr4 favorisait intrinsèquement la réponse du CG et la différenciation plasmocytaire, les plasmocytes spécifiques d'un Ag n'étaient que très rarement détectés dans la MO, un défaut reflété par l'accumulation précoce de plasmablastes immatures occupant potentiellement les niches de survie des plasmocytes à longue durée de vie. Un réglage fin de la désensibilisation de Cxcr4 semble indispensable pour une différenciation et un maintien efficaces des plasmocytes, et l'absence d'un tel processus de régulation pourrait expliquer l'immunité humorale déficiente observée chez les patients atteints de WHIM (369).



**Figure 6. Représentation schématique des signalisations d'aval de MYD88 (partie gauche du schéma) et CXCR4 (partie droite) (370)**



**Figure 7. Représentation schématique des mutations identifiées dans le domaine C-terminal de CXCR4 chez les patients atteints de MW (355).** Les mutations décrites dans (306) sont suivies d'un astérisque, les autres proviennent de (355). Les mutations frameshift et non-sens sont respectivement représentées en bleu et gris. L'acide aminé S338, siège de mutations hotspot, est représenté en noir.

### iii. Mutations du gène *ARID1A*

Des mutations somatiques du gène *ARID1A* (AT-rich interactive domain 1A) ont été identifiées chez 10 à 17% des patients atteints de MW (94, 340). Elles comprennent principalement des mutations frameshift et aboutissent à une protéine tronquée. Les patients avec une mutation d'*ARID1A* présentaient, comparativement aux patients qui n'en avaient pas, une infiltration médullaire plus importante ainsi qu'un taux d'hémoglobine et de plaquettes plus faible. *ARID1A* et son homologue *ARID1B*, dont les délétions sont fréquentes dans la MW (306), font partie de la famille de protéines SWI/SNF. Les membres de la famille SWI/SNF participent au remodelage de la chromatine et donc à la régulation de l'expression des gènes. Bien que les mécanismes soient encore mal compris dans le contexte des hémopathies malignes, *ARID1A* peut moduler l'expression de *TP53* et agir comme un suppresseur de tumeur dans diverses tumeurs solides



(371), dans lesquels les mutations d'ARID1A ont fait l'objet d'études fonctionnelles plus approfondies (372-374).

#### iv. Mutations des gènes *CD79A* et *CD79B*

Les protéines CD79A et CD79B font partie du complexe BCR. Elles forment un hétérodimère qui s'associe à la chaîne lourde de l'immunoglobuline, élément requis pour l'expression du BCR à la surface du LB et sa signalisation. Des mutations activatrices dans les séquences codant pour le motif ITAM de *CD79A* et *CD79B* ont été identifiées dans les DLBCL-ABC. Elles activent la signalisation du BCR, induisant la survie et la prolifération cellulaire, par une cascade qui inclut SYK, PLC $\gamma$ 2 et BTK (92). Le même type de mutations de *CD79A/B* a été observé dans la MW chez 8 à 12% des patients (340, 341). Dans une étude, les mutations de *CD79A/B* étaient observées de façon presque mutuellement exclusive aux mutations de *CXCR4*, suggérant éventuellement deux voies de transformation tumorale après un premier événement qui serait constitué par les mutations de *MYD88* (355). Dans une petite série de patients atteints de MW, la présence de mutations de *CD79B* était associée à un risque plus important de transformation en DLBCL (Jimenez et al., IWWM9 2016).

#### v. Autres mutations

D'autres mutations récurrentes ont été identifiées chez les patients MW, intéressant des gènes connus le plus souvent pour être mutés dans d'autres hémopathies malignes, comme *MYBBP1A*, *TP53*, *KMT2D* (anciennement *MLL2*), *HIST1H1E*, *TRAF3* ou encore *NOTCH2* (94, 340). Une étude récente de séquençage ciblé, réalisée chez plus de 100 patients MW, a ainsi observé jusqu'à 24% de mutations de *KMT2D* (340). Ces mutations étaient le plus souvent sous-clonales. Elles n'étaient pas associées à un pronostic particulier. *KMT2D* est une histone méthyltransférase qui méthyle notamment la lysine 4 de l'histone 3 (H3K4me). Des mutations de *KMT2D* sont décrites dans le lymphome folliculaire (89%) et les DLBCL (32%). Dans cette même étude, des mutations de *PRDM1* ont également décrites, sans qu'il soit précisé s'il s'agissait de mutations inactivatrices.

#### vi. Anomalies du nombre de copies

Les anomalies du nombre de copies sont fréquentes chez les patients atteints de MW, peuvent intéresser diverses régions chromosomiques dont la région 6q (cf. paragraphe 4.b) et impliquent des gènes ayant d'importantes fonctions de régulation au sein du LB (306, 335). Dans la région 6q, on observe la perte de gènes modulant l'activité NF- $\kappa$ B (*TNFAIP3*, *HIVEP2*, *IBTK*), la famille des protéines BCL2





(*BCLAF1*), l'apoptose (*FOXO3*), BTK (*IBTK*), la différenciation plasmocytaire (*PRDM1*) mais aussi *ARID1B*. Les gènes fréquemment délétés et qui ne sont pas situés dans la région 6q comprennent *ETV6*, un répresseur transcriptionnel ; *BTG1*, dont la délétion est fréquente dans les DLBCL et qui est associée à une cortico-résistance dans la leucémie aiguë lymphoblastique; et *LYN*, une kinase qui régule la signalisation du BCR. Les délétions de *PRDM2* et *TOP1*, qui participent à la signalisation liée à P53, sont également fréquentes chez les patients MW (306).

### vii. Profil d'expression génique

Les études de profil d'expression génique (PEG) sont peu nombreuses dans la MW. Les deux premières études les plus significatives, utilisant une approche de PEG par puces ARN, ont comparé les PEG de cellules de MW, de LLC, de MM, de LB et plasmocytes normaux (375, 376). L'analyse non supervisée montrait notamment que le PEG des MW était plus proche de la LLC que du MM, mais également plus proche de celui des LB que des plasmocytes (376). L'un des résultats les plus significatifs de ces travaux était le niveau d'expression élevé de l'IL6 dans la MW par rapport au MM, la LLC et les LB normaux. L'IL6 est une cytokine majeure de l'inflammation. Elle stimule notamment l'activité lymphocytaire B et la production d'Ac, en activant la voie MAPK. Les analyses du transcriptome (RNA-seq) de LB médullaires de patients atteints de MW, comparé à celui de LB de donneurs sains, ont montré une dérégulation de gènes reliés à des voies de signalisation primordiales pour la survie et la prolifération cellulaire : IL6 (*ABCB1*, *GRB2*), NF- $\kappa$ B (*TRAF3*, *IGF1R*, *MOPR2*, *TBK1*), JAK/STAT (*GRB2*, *PTPN1*), PI3K/AKT (*GAB2*, *GRB2*, *CDKN1B*, *IL4R*, *ITPR2*, *ITPR1*) (377). Il existait également une expression accrue des gènes de recombinaison VDJ comme *DNTT*, *RAG1* et *RAG2*, ainsi que des gènes de la voie CXCR4 comme *CXCL12*, *VCAM1* et *CXCR4* lui-même (377), suggérant dans ce dernier cas un rôle de la signalisation CXCR4 dans la MW indépendamment du statut mutationnel de *CXCR4*. Une dérégulation des membres de la famille BCL2, notamment une augmentation d'expression de BCL2 et BCL2L1, a également été observée. Parmi les patients MW, trois groupes pouvaient être distingués selon le statut mutationnel de *MYD88* et *CXCR4* (*MYD88* WT/ *CXCR4* WT, *MYD88* muté/ *CXCR4* WT, *MYD88* et *CXCR4* mutés). Parmi les patients *MYD88* muté, on notait chez les patients *CXCR4* muté, comparativement aux patients *CXCR4* WT, une diminution d'expression de gènes suppresseurs de tumeurs comme *WNK2*, *CDKN1C*, *PRDM5* et *CABLES1*. Ces résultats nécessiteraient d'être confirmés sur une plus grande série de patients et intégrés à une analyse plus large comprenant notamment l'étude de populations lymphocytaires B à différents stades de maturation



ainsi que différents SLPC B afin de mieux appréhender les éventuelles spécificités transcriptomiques de la MW et le rôle du profil des mutations les plus récurrentes.

## 5. Famille des protéines ETS

### a. Généralités

Les protéines à domaine ETS sont un groupe de FTs qui dirigent l'expression de gènes impliqués dans divers mécanismes cellulaires normaux et pathologiques, en se liant à des promoteurs et activateurs spécifiques et en facilitant l'assemblage d'autres composants de la machinerie transcriptionnelle. Le domaine ETS, très bien conservé et liant l'ADN, définit la famille et est responsable de la reconnaissance spécifique d'une séquence au motif commun, 5'-GGA(A/T)-3'.

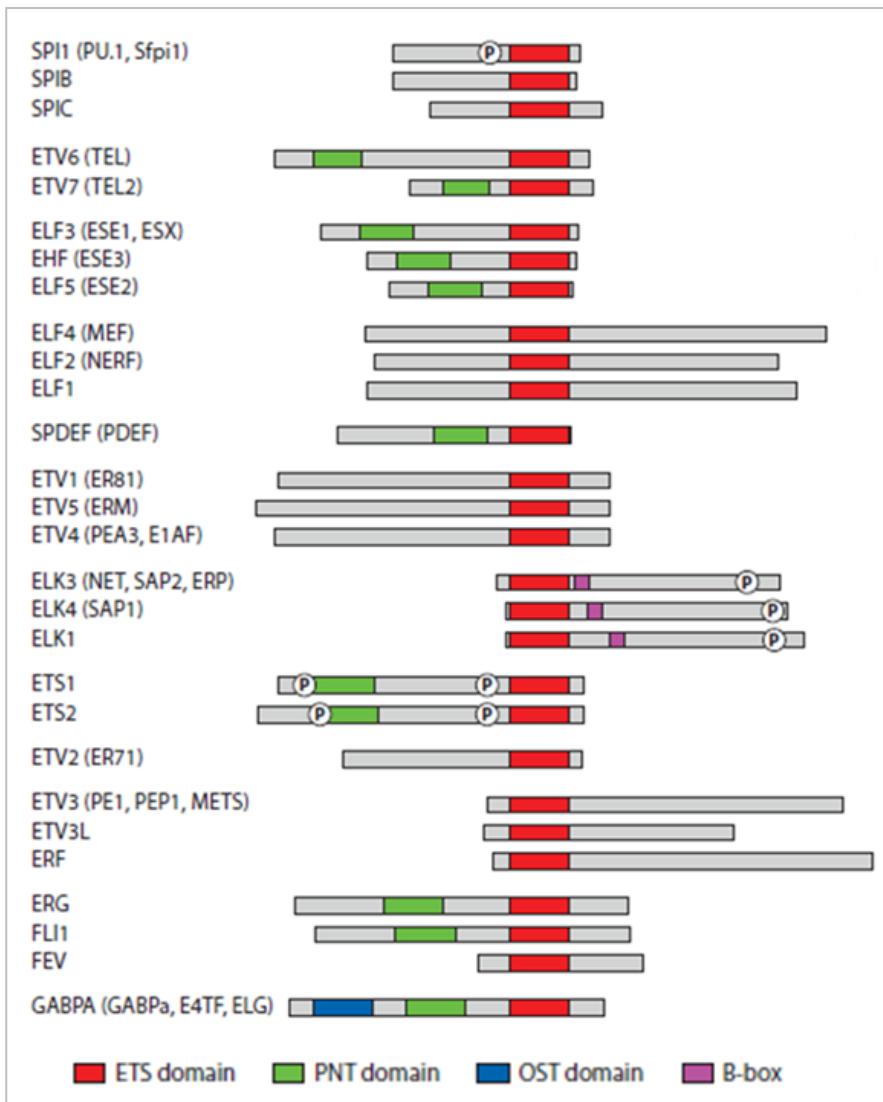
Le 1<sup>er</sup> membre de cette famille, ETS1, a été identifié chez le poulet comme l'une des deux séquences nucléotidiques (*ets-1* et *myb*) transduite par le virus de l'érythroblastose aviaire E26 (378, 379). Ces 2 séquences étaient fusionnées aux séquences rétrovirales *gag*, formant l'oncogène E26. La séquence ETS (E twenty six, E26) était indispensable à la transformation tumorale chez les poulets infectés. Sur la base de leur homologie de séquence avec le domaine de liaison à l'ADN d'ETS1, de nombreuses autres protéines à domaine ETS ont été identifiées dans divers organismes (28 chez l'homme, 27 chez la souris), même si elles semblent restreintes aux métazoaires (380). Les membres de la famille ETS sont regroupés en différentes sous-classes sur la base de leurs similarités de séquence du domaine ETS et de la présence d'autres domaines conservés (**Figure 8**) (381). Parmi ces régions conservées, le domaine PNT (pointed domain), initialement identifié chez la *Drosophila melanogaster* (382), est présent chez 11 membres de la famille ETS. Il participe à l'homo-oligomérisation (383, 384), l'hétérodimérisation (385) et la répression transcriptionnelle (386). Le domaine B-box, présent chez les trois membres de la sous-classe TCF (ternary complex factor), forme un élément structural inductible qui interagit sur l'ADN avec le facteur SRF (serum response factor) (387). Le domaine OST (On-SighT), uniquement présent au sein de la protéine GABPA, participe au recrutement de co-facteurs.

Les différentes sous-familles de protéines ETS possèdent de nombreuses similitudes : elles ont des séquences bien conservées, des fonctions moléculaires et biologiques le plus souvent relativement similaires, elles peuvent agir comme des cibles nucléaires de voies de transduction du signal et fonctionnent généralement en collaboration avec d'autres FTs (381). Cependant, bien qu'il y ait des similitudes générales, les mécanismes de signalisation, les partenaires protéiques et les modes d'expression spatiale et



temporelle peuvent différer pour chaque membre expliquant en partie une spécificité d'action. Divers mécanismes régulent l'activité des facteurs ETS : contrôle transcriptionnel étroit, microARNs et protéines co-régulatrices. La liaison à l'ADN, l'interaction protéine-protéine, l'activation et la répression transcriptionnelles ainsi que la localisation cellulaire des divers facteurs ETS sont eux aussi étroitement contrôlés par diverses modifications post-traductionnelles (phosphorylation, acétylation, sumoylation, ubiquitylation et glycosylation), permettant un contrôle précis du programme transcriptionnel spécifique des ETS (380, 381). L'expression transcriptionnelle des divers ETS a été étudiée dans plus de 20 tissus et lignées cellulaires (388). La moitié des membres de la famille ETS est ainsi exprimée dans tous les types de cellules testés. L'expression de chaque membre de la famille ETS n'est donc pas limitée à un type tissulaire ou cellulaire pour une activité spécifique. Dans certains cas, cependant, une forte expression est confinée à un seul type cellulaire (p. ex. SPDEF dans la prostate). Physiologiquement, les facteurs ETS régissent de multiples processus, comprenant notamment le contrôle du cycle cellulaire, la différenciation, la prolifération, l'apoptose, le remodelage tissulaire et l'angiogenèse (381).





**Figure 8. Domaines structuraux et fonctionnels de la famille des FT ETS (380).** Les rectangles de couleurs identifient le domaine ETS de liaison à l'ADN (rouge), le domaine PNT (vert), le domaine OST (bleu) et la B-box (magenta). Le P encerclé symbolise une région phosphorylée. Ne sont pas identifiées les autres régions impliquées dans l'activation/répression transcriptionnelle, les modifications post-traductionnelles (amarrage des kinases, sumoylation), transport nucléaire, etc.

### b. Domaine ETS

L'élément commun à toute cette famille de FTs est la présence d'un domaine de liaison à l'ADN, le domaine ETS. Ce domaine conservé de 85 acides aminés est un variant du motif de liaison à l'ADN prototypique « hélice-boucle-hélice » (HLH, helix-loop-helix) et se compose de trois hélices  $\alpha$  reposant sur quatre feuillets  $\beta$  antiparallèles (381). Cette structure lie l'ADN sur une région couvrant 12 à 15 paires de bases (bp), mais il affiche une préférence de séquence pour 9 bp seulement, dont 3 sont en amont et 2 en aval d'un noyau central qui ne varie pas, 5'-GGA(A/T)-3'. L'interface protéine-ADN est caractérisée à la fois

par une lecture directe, qui implique un contact par des chaînes latérales du domaine ETS, et par une lecture indirecte, qui utilise le positionnement du squelette phosphodiester. Les interactions directes les plus importantes dans la poche principale impliquent une liaison hydrogène entre les deux arginines invariantes de l'hélice de reconnaissance H3 et les deux guanines du noyau GGA(A/T). D'autres interactions directes et médiées par un groupement H<sub>2</sub>O, hydrophobes et électrostatiques sont assurées principalement par des résidus dans la boucle entre les hélices H2 et H3, l'"aile" entre les brins S3 et S4, et l'extrémité N terminale de l'hélice H1. Il n'y a aucun contact direct avec une bp en dehors du motif central GGA(A/T), ce qui implique qu'il existe une grande importance des régions latérales pour la spécificité de liaison de chaque protéine (389). Des études thermodynamiques ont ainsi révélé une variation de l'affinité de SPI1 pour l'ADN d'une amplitude de 400, selon les séquences entourant le noyau GGA (390).

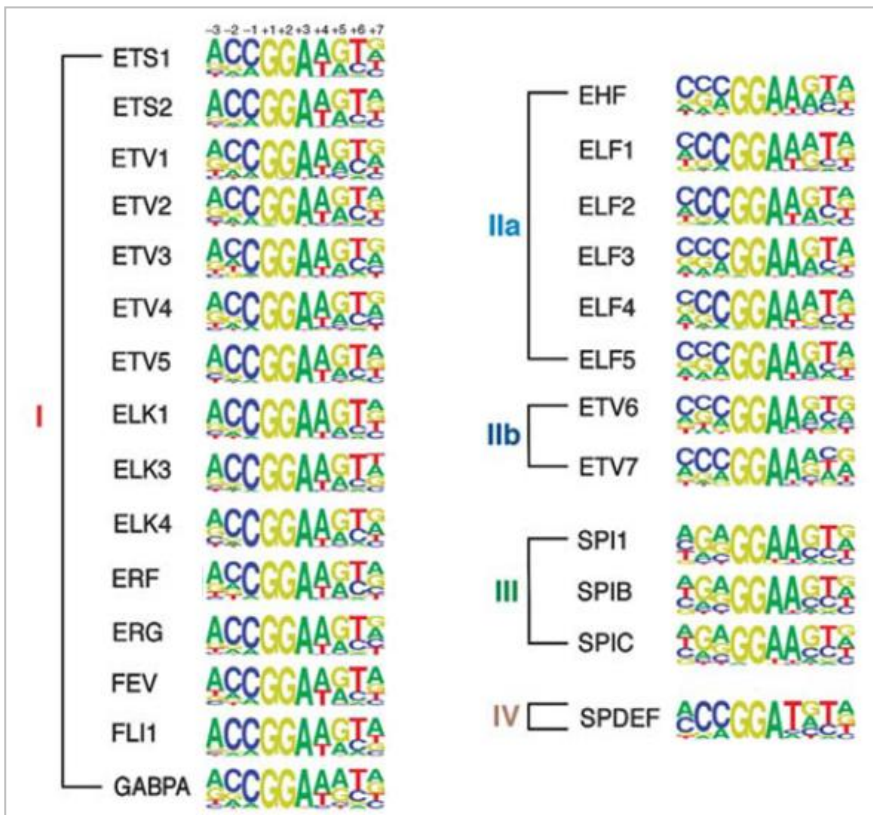
### c. Spécificité de liaison à des séquences ADN préférentielles

Plusieurs études *in vitro* ont pu déterminer la sélectivité de certains membres de la famille ETS et leurs préférences de liaison pour une séquence d'ADN particulière (389, 391-397). L'ensemble des protéines ETS murines et humaines a ainsi été soumis à des études de spécificité de liaison protéique à l'ADN en micropuits et par microarrays afin d'en déterminer les préférences de séquence de liaison (398). La séquence consensus à affinité la plus élevée comporte 9 bp, ACCGGAAGT, comprenant un noyau GGA invariant (**Figure 9**).

Les préférences de séquence observées dans l'analyse systématique de l'ensemble de la famille suggèrent qu'il existe quatre classes de protéines ETS (398). La classe I, qui 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 ; sous-familles PEA3, TCF, ETS, ERF et ERG), présente une séquence consensus identique à la séquence *in vitro* la plus courante (ACCGGAAGT). La classe II comprend huit membres (classe IIa : EHF, ELF1, ELF2, ELF3, ELF4, ELF5 ; classe IIb : ETV6, ETV7 ; sous-familles TEL, ESE et ELF) et a une séquence consensus qui diffère principalement au niveau du premier nucléotide (CCCGGAAGT). 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 (393). La classe IV ne comprend que SPDEF, qui a une préférence unique pour une séquence de noyau GGAT au lieu de GGAA (399). Les études quantitatives de liaison à l'ADN indiquent que l'affinité des différentes protéines ETS pour une séquence consensus définie varie seulement d'un facteur 10 environ, alors que l'affinité de la même protéine pour les différentes séquences consensus diffère de 5 à 80 fois (400). Les mécanismes



expliquant les spécificités de liaison des différentes classes de protéines ETS et leurs influences sur l'activité transcriptionnelle de ces FTs restent encore incomplètement élucidés.



**Figure 9. Spécificités de liaison à l'ADN des FT ETS chez l'homme (398).** Les profils de liaison à l'ADN (colonnes de droite) des différentes protéines ETS (colonnes de gauche), déterminés à l'aide d'essais réalisés en micropuits, sont indiqués. Les logos sont dessinés à l'aide d'enoLOGOS et la hauteur d'une lettre à une position particulière est directement proportionnelle à l'effet de ce nucléotide sur l'affinité de liaison de la protéine.

#### d. Mécanismes de régulation de la liaison à l'ADN

Un certain nombre d'autres facteurs influencent la liaison des protéines ETS à l'ADN, notamment l'interaction avec d'autres FTs (400), la liaison de séquences palindromiques sous forme de dimères (401), et la présence de régions auto-inhibitrices autour du domaine ETS (402, 403). L'activité de liaison à l'ADN est généralement inhibée jusqu'à ce qu'un déclencheur précis soit en place. Les deux types principaux de déclencheurs activant la liaison à l'ADN sont la phosphorylation et les interactions avec un FT co-régulateur.

### i. Auto-inhibition

Plusieurs protéines ETS possèdent des mécanismes intrinsèques qui inhibent leur liaison à l'ADN. Ce phénomène a pu être identifié 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 d'auto-inhibition est celui d'ETS1. Il a été découvert par des analyses de délétion (392, 404, 405) et sa pertinence biologique confirmée par l'existence de deux variants naturels. L'oncoprotéine p135 du rétrovirus E26, ayant permis la découverte d'ETS1, ne comprend ainsi pas de séquences inhibitrices C-terminales (406). Un variant d'épissage d'ETS1, ne possédant pas les séquences inhibitrices N-terminales codées par l'exon 7, se lie à l'ADN avec une affinité plus élevée et une spécificité plus large que les autres isoformes (407). Le mécanisme autorégulateur d'ETS1 a fait l'objet de nombreuses études. Les régions situées à la fois aux extrémités N- et C-terminales du domaine ETS forment un module inhibiteur qui bloque la liaison d'ETS1 à l'ADN (408, 409). Cette auto-inhibition peut être levée suite à un changement conformationnel et au déroulement de l'une de ces hélices  $\alpha$ . La liaison du FT RUNX1 à ETS1 peut lever l'auto-inhibition en perturbant l'action de ce module inhibiteur (410, 411). De même, les interactions avec le domaine de liaison à l'ADN (paired domain) du FT Pax5 sont connues pour favoriser la liaison à l'ADN d'Ets1 et d'un sous-ensemble d'autres protéines ETS à un site de liaison Ets dit 'faible' au sein du promoteur *mb-1* (412). Des études structurales montrent que les interactions entre le paired domain de Pax5 et l'hélice de liaison à l'ADN d'Ets1 entraînent des changements dans les interactions protéine-ADN d'Ets1 et favorisent ainsi sa capacité à reconnaître un site ADN considéré comme sous-optimal. Inversement, la phosphorylation d'Ets1 par CAMKII renforce l'auto-inhibition d'Ets1 par la stabilisation de la conformation inhibitrice (413).

D'autres membres de la famille ETS, notamment ceux de la sous-famille TCF, sont également sujets à l'auto-inhibition. Cette inhibition est médiée par les interactions entre le domaine ETS en N-terminal et la partie C-terminale de la molécule. La phosphorylation du domaine C-terminal par les MAPK entraîne une perte de l'auto-inhibition et, par conséquent, une meilleure liaison à l'ADN. Cette meilleure liaison à l'ADN s'accompagne d'un changement de conformation complexe transmis à travers toute la protéine (414). L'interaction avec un second FT lié au promoteur peut favoriser, comme c'est le cas avec SRF (415, 416) et Pax5 (412, 417), ou dissocier, comme c'est le cas avec les protéines HLH, la liaison à l'ADN des protéines TCF.





## ii. Partenaires protéiques de co-régulation

L'interaction avec des partenaires de co-régulation est partagée par la majorité des FTs ETS (418). Elle permet un contrôle combiné de l'expression des gènes et améliore la spécificité d'action des protéines ETS. L'importance de ces complexes dans l'induction de réponses génétiques spécifiques a été notamment mise en évidence par des études sur Spi1/Pu.1 et son partenaire co-régulateur dans la lignée lymphoïde, Pip/Irf4. Il a été démontré que les gènes dont l'expression est régulée par le complexe Pu.1-Pip étaient différents de ceux régulés par Pu.1 seul (419).

Trois types d'interaction existent : i) celle avec d'autres FTs liés à l'ADN, ii) celle avec des FTs non liés à l'ADN et iii) celle avec des complexes co-activateurs-co-répresseurs. La formation de complexes avec des FTs liés à l'ADN implique le plus souvent des interactions entre le domaine ETS et le domaine de liaison à l'ADN réciproque sur le FT partenaire, permettant la juxtaposition de leurs sites de liaison à l'ADN. Un modèle pour ce type d'interaction est celui entre Ets1 et Runx1, dans lequel les interactions réciproques entre leurs domaines respectifs de liaison à l'ADN favorisent la formation d'un complexe tertiaire (410, 411). Cependant, il existe d'autres types d'interactions aboutissant à la formation de complexes protéiques notamment pour ceux impliquant la sous-famille TCF avec SRF, où les interactions entre les domaines de liaison à l'ADN sont minimales, et l'interaction principale se situe entre une courte région des TCF (la B-box) et le domaine de liaison à l'ADN de SRF (387). Comme la séquence B-box est unique aux membres de la sous-famille TCF, cela garantit que ces protéines sont spécifiquement recrutées sur les sites de liaison composites SRF-ETS. Enfin, la formation d'un complexe se fait également avec des protéines co-régulatrices qui n'établissent que peu ou pas de contact avec l'ADN. Là encore, ces interactions concernent souvent le domaine ETS mais d'autres régions peuvent être impliquées comme les régions répétées riches en ankyrine de GABPB qui se lient au domaine ETS et à une hélice  $\alpha$  C-terminale de GABPA (420).

En plus des partenaires de co-régulation spécifiques, des complexes plus généraux de co-répresseurs et de co-activateurs peuvent être recrutés par les protéines ETS. On peut citer par exemple CREB-binding protein (CBP), qui a été proposé comme co-activateur de plusieurs protéines ETS, telles que ETS1, ETV1, la sous-famille des TCF et SPI1 (381, 421).

## iii. Liens avec les voies de signalisation

De nombreux FTs ETS sont des cibles nucléaires de voies de signalisation, en particulier celles des MAPK. ELK1 a été la première cible directe décrite d'ERK (422). La phosphorylation d'ELK1 par ERK



favorise une meilleure liaison à l'ADN et stimule sa capacité d'activation transcriptionnelle. Par la suite, il a été démontré que tous les membres de la sous-famille TCF sont des cibles de cette voie, et qu'ils répondent de façon différente aux kinases JNK et MAPK (423, 424). Les membres de la sous-famille TCF contiennent des domaines d'amarrage pour les MAPK (380). Ces séquences peptidiques linéaires, appelées motifs D et FxFP, entourent les domaines de transactivation contenant des sites accepteurs de groupement phosphate des protéines TCF. Ces motifs augmentent l'efficacité de phosphorylation en attirant les MAPK près des sites accepteurs, et la phosphorylation qui en résulte régule la liaison à l'ADN, le recrutement de co-activateurs ou de co-répresseurs et la localisation cellulaire des protéines TCF (380, 425).

Ceci n'est pas spécifique à la sous-famille TCF puisque le même type de mécanisme a été identifié pour ETS1 et ETS2, où les domaines PNT pourraient faciliter l'activité ou servir de sites d'amarrage aux MAPK. Bien que l'accent ait été mis sur les voies MAPK, d'autres voies de signalisation régulent également les protéines ETS. Par exemple, la phosphorylation de SPI1/PU.1 par la caséine kinase II favorise les interactions avec son partenaire co-régulateur Pip et les kinases calcium-dépendantes comme CAMKII inhibent la liaison à l'ADN d'ETS1 (413, 426).

#### **e. Localisation des liaisons à l'ADN à l'échelle du génome**

Les techniques de CHIP-array et CHIP-seq ont permis l'analyse à l'échelle du génome de la localisation des sites de liaison des facteurs ETS (427-429). Plusieurs protéines ETS et protéines de fusion ont été analysées (ETS1, GABPA, ELF1, ERG, FLI1, ELK1, EWS-ERG, EWS-FLI1, SPI1, SPDEF, ETV1). Les résultats de ces études montrent que les cibles génomiques des protéines ETS peuvent être schématiquement divisées en deux classes (400, 428). Il y a d'abord les sites de liaison dits « redondants » que l'on trouve dans les promoteurs proximaux des gènes de « ménage », c'est-à-dire exprimés dans la plupart des types cellulaires (400). En effet, les régions ADN liées de façon redondante se trouvent le plus souvent à une courte distance (environ 40 pb) en amont des sites d'initiation de la transcription (TSS, transcription starting site) et présentent des caractéristiques de promoteurs actifs, comme en témoignent la sensibilité à la DNase I et la présence d'une triméthylation de l'histone H3K4. Les sites de liaison de cette classe sont caractérisés par une séquence consensuelle ETS (CCGGAAGT), suggérant que la plupart, sinon la totalité, des protéines ETS ont la capacité de les lier avec une affinité relativement élevée. La fréquence de cette séquence (chez environ 25% des promoteurs humains), l'association avec des gènes couramment exprimés et sa présence près des TSS suggèrent un rôle commun des facteurs ETS pour différentes étapes de la transcription ou le maintien d'une chromatine en configuration active. Il



Il y a d'autre part des sites de liaison dits "spécifiques" que l'on trouve le plus souvent dans les régions enhancères associées à des gènes pouvant avoir des fonctions biologiques spécifiques d'un membre de la famille ETS (400). 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 FTs. Ces sites de faible affinité liés par les ETS serviraient de médiateurs pour des fonctions spécifiques des protéines ETS, nécessitant des interactions avec leurs FTs partenaires.

#### **f. Mécanismes de tumorigénèse impliquant les protéines ETS**

Les mécanismes de tumorigénèse impliquant les protéines ETS sont principalement des mécanismes gain de fonction. L'activation aberrante des protéines ETS est considérée comme un mécanisme clé dans la survenue de nombreux types de tumeurs solides (sarcome d'Ewing ; cancers du sein, de la prostate, de l'estomac ou encore de la thyroïde) et d'hémopathies malignes (430, 431). Les translocations chromosomiques et amplifications géniques sont les principales causes d'expression aberrante de facteurs ETS.

#### **i. Translocations chromosomiques et amplification génique**

##### **1. Tumeurs solides**

Les premières translocations chromosomiques décrites, à l'origine de protéines de fusion oncogéniques, sont EWS-FLI1 (EWS RNA-binding protein 1-friend leukaemia integration 1 transcription factor) (432) et EWS-ERG (433) - qui ont été découvertes dans le sarcome d'Ewing pédiatrique au début des années 1980 - ainsi que ETV6-NTRK3 (ets variant 6-neurotrophic tyrosine kinase receptor type 3), qui a été identifié dans un sous type rare sécrétant de cancer du sein (434). Les translocations chromosomiques actuellement les plus étudiées sont celles identifiées dans le cancer prostatique. L'analyse du génome tumoral du cancer prostatique a révélé que les membres de la famille *ETS*, *ERG* et *ETV1*, étaient fréquemment réarrangés, entraînant des fusions avec la région 5' UTR du gène *TMPRSS2* (transmembrane protease serine 2) aboutissant aux gènes de fusion *TMPRSS2-ERG* et *TMPRSS2-ETV1* (435). Ces réarrangements suppriment les exons 1 à 4 codant pour l'extrémité N-terminale des facteurs ETS et placent les exons restants sous le contrôle d'éléments de réponse de récepteurs aux androgènes de *TMPRSS2* (436). Dans le cas des réarrangements *TMPRSS2-ERG*, l'expression basale d'ERG est augmentée d'environ 2 000 fois et d'environ 6 000 fois après stimulation par les androgènes (435). Par la suite, des réarrangements similaires impliquant *ETV4* et *ETV5* ont été détectés à des fréquences inférieures (437, 438). Au total, environ 50% des tumeurs prostatiques localisées et 40% des tumeurs métastatiques contiennent des



fusions Tmprss2-Ets. Des études fonctionnelles utilisant des modèles murins ont démontré que la surexpression d'*Erg* ou *Etv1* au sein de l'épithélium prostatique pouvaient entraîner des néoplasies intraépithéliales prostatiques et, en combinaison avec la perte du suppresseur de tumeur *PTEN*, des adénocarcinomes (439-442).

Des gains du nombre de copies et des amplifications des gènes *ETS* conduisant à leurs surexpressions ont également décrits dans les tumeurs solides. Le nombre de copies d'*ETV1* est augmenté (> 6 copies) dans environ 13% des mélanomes localisés et 18% des mélanomes métastatiques, et l'expression combinée d'*ETV1* et du mutant *NRAS* G12D conduit à l'immortalisation et la transformation des mélanocytes (443, 444). Les cancers du sein ont des gains du nombre de copies des régions 1q21-23 et 1q32, qui sont associés à l'amplification et à la surexpression d'*ETV3* et *ELF3*, respectivement (445).

## 2. Hémopathies

Les gènes de fusion comportant des *ETS* sont fréquemment identifiés dans les leucémies aiguës et néoplasies myéoprolifératives humaines. *TEL/ETV6* est ainsi impliqué dans plus de 40 translocations différentes où le domaine PNT, le domaine ETS, ou les deux domaines, sont fusionnés en phase avec des FTs partenaires tels que *AML1* (446, 447) et *EVI1* (448), ou des tyrosines kinases, comme *PDGFβR* (449), *TRKc* (450), *ABL* (451) et *JAK2* (383). Dans la plupart des cas des fusions impliquant des tyrosine kinases, le domaine PNT de TEL s'oligomérisent et participe à l'activation constitutive de la protéine tyrosine kinase et à la transformation tumorale. Les protéines de fusion TEL-PDGFRβ, TEL-TRKc, TEL-ABL et TEL-JAK2 sont chacune capables de provoquer une néoplasie myéoproliférative dans des modèles murins de greffe de MO et d'activer des cibles en aval, comme STAT5, qui peut ensuite transactiver un certain nombre de gènes, dont ceux codant pour l'oncostatine M, BCL-XL, la cycline D1 et PIM1, qui régulent tous la prolifération et la survie cellulaire (452). D'autres fusions impliquant TEL et divers partenaires (MN1, ARNT, TTL, BAZ2A, MDS2) contribuent à la transformation tumorale par une perte de l'activité transcriptionnelle de TEL et/ou de ses partenaires (430). Plus récemment, deux gènes de fusions impliquant *SPI1* (*STMN1-SPI1* et *TCF7-SPI1*) ont été identifiés dans les LAL-T pédiatriques de haut risque (453). Les cas de LAL-T arborant ces fusions avaient un immunophénotype et un PEG particuliers ainsi qu'un pronostic très défavorable. Ces protéines de fusion conservaient la capacité de lier l'ADN et une activité transcriptionnelle. Leurs expressions dans des progéniteurs murins augmentaient la prolifération cellulaire et provoquaient un blocage de la maturation T (453).



Au sein des hémopathies lymphoïdes B, ce type d'évènement est plus rare même si des translocations impliquant les gènes des Ig (*IGH* ou *IGL*) et *SPIB* ainsi que des gains de la région 11q24, comprenant *ETS1* et *FLI1*, ont été décrits dans les DLBCL (258, 454, 455). La diminution de l'expression d'*ETS1* et *FLI1* dans des lignées cellulaires de DLBCL entraînait une perte de la viabilité cellulaire et perturbait l'expression de gènes clés de la réaction du CG et de la différenciation plasmocytaire (*XBP1*, *PRDM1*) (454). Plus récemment, des mutations somatiques faux-sens d'*ETS1* et *ETV6* ont été identifiées chez environ 5% des patients atteints de DLBCL (154, 456)

## ii. Autres mécanismes de gain de fonction

L'activité transcriptionnelle des facteurs ETS peut être augmentée suite à la survenue de mutations de régions régulatrices de gènes agissant en cis. Ces mutations génèrent de nouveaux sites de liaison GGAA pour les facteurs ETS. Ce type d'évènements a été décrit de façon très fréquente dans les tumeurs solides (431). Des mutations dans les régions promotrices de *TERT* (telomerase reverse transcriptase), conduisant à une surexpression de *TERT*, ont ainsi été identifiées dans près de 70% des mélanomes (457, 458).

Un mécanisme d'action en trans nouvellement décrit et conduisant à l'activation de facteurs ETS implique les mutations faux-sens du gène *TP53*. Une partie de ces mutations abroge la fonction suppressive de tumeur de p53 tout en engendrant un gain de fonction oncogénique pour la protéine (459). Ces mutations gain de fonction (GOF, gain of function) permettent à p53 de se lier à de nouveaux partenaires protéiques, notamment des FTs, ce qui entraîne une dérégulation de nouveaux gènes cibles. Des expériences de ChIP-seq réalisées dans des cellules de syndrome de Li-Fraumeni ont démontré que les motifs consensus ETS étaient fréquemment identifiés dans les promoteurs liés par les mutants GOF de p53 (460). De plus, *ETS2* forme des complexes avec les mutants GOF de p53 plus efficacement qu'avec les formes sauvages et coopère avec les mutants GOF à la modification du programme transcriptionnel, à l'origine de la croissance tumorale et des mécanismes de métastases et de résistance aux traitements (461). Enfin, les mutants GOF de p53 protège *ETS2* de la destruction médiée par COP1 (E3 ubiquitin-protein ligase constitutive photomorphogenesis protein 1). Bien que les données actuelles indiquent qu'*ETS2* est le principal facteur ETS impliqué dans les interactions avec les mutants GOF de p53, l'implication d'autres membres de la famille ETS ne peut être exclue.

D'autres mécanismes impliquant une augmentation de l'activité et de la stabilité des facteurs ETS ont été décrits. Les facteurs ETS sont souvent soumis à une régulation post-traductionnelle. La



phosphorylation par exemple augmente les interactions des FTs ETS avec leurs co-activateurs ou co-répresseurs transcriptionnels. Les facteurs ETS peuvent donc être activés par une signalisation aberrante en amont survenant dans les cellules cancéreuses (380). Dans les tumeurs stromales gastro-intestinales (GIST), les mutations activatrices du récepteur tyrosine kinase KIT entraînent une surexpression d'ETV1 en augmentant la stabilité d'ETV1 via la voie MEK-ERK, initiant ainsi un programme transcriptionnel oncogénique (462). L'inhibition pharmacologique de KIT ou MEK aboutit à la dégradation de la protéine ETV1 sans modifier son taux d'ARNm (463). KIT est également une cible transcriptionnelle directe d'ETV1, et cette relation crée une boucle de signalisation (462). Lorsqu'il est muté et constitutivement actif, KIT agit en synergie avec la signalisation induite par PDGFR $\alpha$  concourant aussi à la stabilisation de la protéine ETV1 (464).

La régulation de la stabilité des protéines ETS par polyubiquitination et la dégradation via le protéasome qui s'en suit participe à l'homéostasie normale de la cellule et au contrôle des processus associés au cancer. ETV1, ETV4, ETV5, ETS1 et ETS2 peuvent tous être ciblés par COP1 et dégradés par le protéasome (465). La fusion *TMPRSS2-ETV1* entraîne fréquemment la perte de sites d'interaction avec COP1 situés à l'extrémité N-terminale d'ETV1, ce qui augmente sa stabilité dans le cancer prostatique (465). La phosphorylation par SRC de la Tyr283 d'ETS1 empêche la liaison de COP1, aboutissant à une surexpression d'ETS1 dans le cancer du sein triple négatif (466).

#### g. SPI1/PU.1

Le gène codant pour le FT SPI1/PU.1 a été caractérisé quasi simultanément par deux approches. Le locus *Spi1* a été identifié comme étant un site d'insertion récurrent pour le SFFV (spleen focus-forming virus) dans les érythroleucémies murines de Friend (467, 468) et a été ainsi nommé SFFV provirus integration site-1 (*Spi1*). L'intégration du SFFV en amont des régions 5' de *Spi1* entraîne l'expression continue de PU.1 dans les proérythroblastes, bloquant ainsi toute différenciation. Ces proérythroblastes immortalisés subissent des anomalies génétiques supplémentaires qui entraînent la prolifération de ces clones tumoraux. Parallèlement, le criblage d'une banque murine d'ADNc de macrophages a permis de cloner un nouveau FT selon sa capacité à se fixer sur une séquence riche en purines (5'-GGAA-3') contenue dans le promoteur des gènes de classe II du CMH. Ce facteur était également capable d'induire la transcription de gènes rapporteurs sous



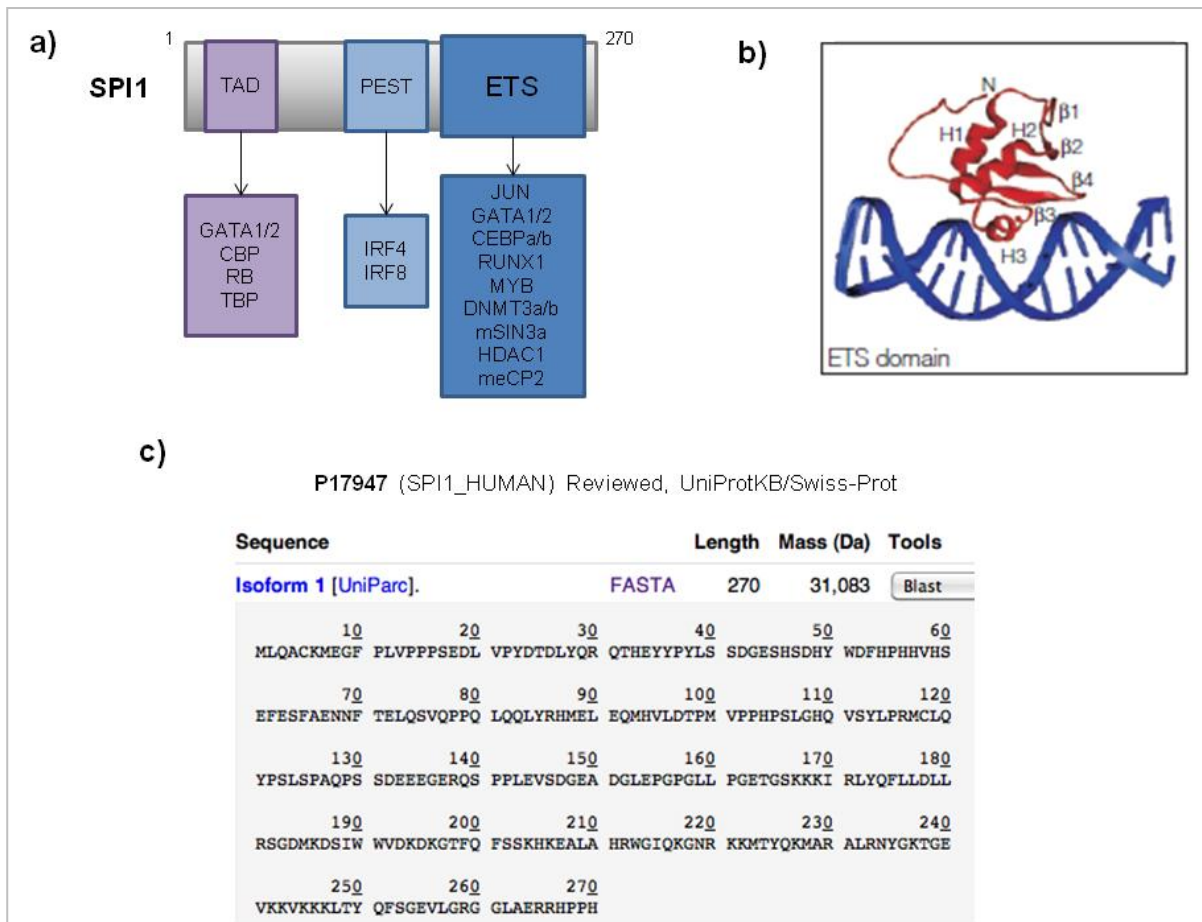
contrôle de divers promoteurs myéloïdes. L'analyse de la séquence de ce facteur a montré qu'il était relié à la famille de protéines ETS (469).

### i. Structure

La séquence codante de *Spi1* a un cadre de lecture de 816 bases, codant une protéine de 272 ou 266 acides aminés (aa), selon le codon d'initiation utilisé pour la traduction. L'homologue protéique humain comporte 264 aa, comprenant cinq exons et quatre introns et présentant une homologie de séquence de 85% avec la séquence murine. Il est situé sur le chromosome 11, en 11p11.22 (470). La protéine de 31 kDa comprend trois domaines fonctionnels distincts, à travers lesquels elle exécute à la fois la liaison à l'ADN et l'activité de transactivation en interagissant avec divers autres régulateurs transcriptionnels (**Figure 10**). Le domaine de liaison de l'ADN, ou domaine ETS, conservé au cours de l'évolution, est confiné à l'extrémité C-terminale de la protéine et sert de médiateur de sa liaison à l'ADN ainsi que de son interaction avec un certain nombre d'autres protéines telles que c-Jun, GATA1, GATA2, C/EBP $\alpha$  et  $\beta$  et RUNX1 (471-473). L'extrémité N-terminale abrite un domaine de transactivation (TAD, transactivation domain) qui contient trois régions riches en aa acides et une région riche en glutamine. Le TAD permet l'interaction avec divers FTs tels que TBP (TATA-Box binding protein), RB, GATA1 et GATA2, CBP ou encore HSP90 (474, 475). Enfin, le domaine PEST, situé entre les deux domaines décrits ci-dessus, est essentiel pour de nombreuses interactions protéine-protéine et doit son nom à une abondance de résidus proline (P), acide glutamique (E), sérine (S) et thréonine (T). Ce domaine interagit avec les facteurs de régulation de l'interféron (IRF), y compris IRF4 et IRF8 (426, 476).







**Figure 10. a)** Domaines fonctionnels de la protéine SPI1 et interaction avec ses protéines partenaires. ETS, domaine de liaison à l'ADN ; PEST, proline-, glutamic acid-, serine-, and threonine-rich domain ; TAD, transactivation domain. **b)** Structure du domaine ETS. Les hélices (H) et feuillets β (β) du domaine ETS sont représentés en rouge. **c)** Séquence amino-acidique de la protéine SPI1 humaine (issue de la base de données de séquences de protéines UniProtKB, [www.uniprot.org](http://www.uniprot.org)).

## ii. Expression et fonctions

SPI1 s'exprime exclusivement dans les cellules hématopoïétiques. Ses niveaux d'expression sont élevés dans les lignées monocytaires, granulocytaires et lymphoïdes B (467, 469, 477, 478). Des niveaux d'expression plus faibles sont observés dans les érythrocytes matures, alors qu'ils ne s'expriment pas dans les cellules T matures (479, 480). Il régule un grand nombre de gènes des lignées myéloïde et lymphoïde. Dans les lignées myéloïdes, les gènes cibles de SPI1 comprennent les récepteurs aux G-CSF, GM-CSF et M-CSF (481, 482). En outre, SPI1 transactive également d'autres gènes impliqués dans l'acquisition d'un phénotype de monocyte mature. Cela permet aux monocytes/macrophages de remplir leurs fonctions immunologiques à la suite d'un signal d'activation. Par exemple, SPI1 contrôle l'expression des

gènes codant pour les FcRI et RIIIA (483, 484), récepteurs d'affinité élevée et faible, respectivement, pour la région constante de l'IgG. Une fois exprimées à la surface de la cellule, ces molécules permettent au macrophage de reconnaître et de phagocyter les bactéries opsonisées par l'IgG. SPI1 induit également l'expression des molécules d'adhésion CD11b, CD18 et CD14 (482, 485, 486) pendant la différenciation myéloïde. CD18 correspond à la chaîne d'intégrines  $\beta$  qui, lorsqu'elles sont associées à CD11b, constituent toutes deux la glycoprotéine membranaire Mac-1, qui assure l'adhésion des monocytes à l'endothélium, ainsi que la phagocytose des particules opsonisées par le complément. CD14 est une glycoprotéine membranaire qui se lie spécifiquement au LPS, un composant structural de la paroi bactérienne. Cette reconnaissance est une étape cruciale dans le déclenchement de la fonction antibactérienne du macrophage.

Une caractéristique commune à tous ces gènes est l'absence de TATA box dans leurs promoteurs ; ils ont un site de liaison SPI1 situé en amont du TSS à une position correspondant classiquement à celle d'une TATA box, accompagné d'un groupe de sites de liaison spécifiques pour d'autres régulateurs ubiquitaires tels que SP1, RUNX1, C/EBP et les membres de la famille CBF (474). Par l'intermédiaire de son TAD, SPI1 peut se lier *in vitro* à TBP et à un ensemble de facteurs associés aux TATA (TAFs). Il peut ainsi activer presque tous les promoteurs myéloïdes en recrutant des TBP et ensuite d'autres composants de la machinerie transcriptionnelle.

### iii. Rôle et régulation de l'expression de SPI1 dans l'hématopoïèse précoce

L'expression du gène SPI1 est étroitement régulée dans le système hématopoïétique, de sorte que son niveau d'expression détermine si les cellules se différencient en monocytes/macrophages, granulocytes ou lymphocytes. Lorsque les CSH et progéniteurs multipotents se différencient, l'expression de SPI1 est maintenue ou régulée à la hausse dans certaines lignées ou bien à la baisse dans d'autres. L'expression de SPI1 dans les CSH est importante pour leur auto-renouvellement et pour leur développement en CMP et CLP (487). Ces deux groupes de progéniteurs se différencient ensuite pour former des cellules sanguines matures, notamment des mégacaryocytes, des globules rouges, des neutrophiles et des macrophages (tous dérivés du CMP) ainsi que des LB et LT (dérivés du CLP). Iwasaki et al. ont montré, dans un modèle murin de foie fœtal KO pour SPI1, que la suppression de l'expression de SPI1 entraîne une diminution du nombre de CSH et une disparition des CMP, CLP et GMP (487). Ces défauts hématopoïétiques sont autonomes sur le plan cellulaire car les CSH KO pour SPI1 ne génèrent pas de colonies myéloïdes lorsqu'elles sont cultivées *in vitro* (487, 488). De même, dans des études de greffes compétitives, les CSH, issues de



foie fœtal et injectées dans des souris irradiées de façon létale, étaient à peine détectables dans le sang périphérique, et leurs faibles nombres dans la MO ne permettaient pas le développement de monocytes, de LB et LT (487, 489, 490).

Des études réalisées à l'aide de systèmes rapporteurs dans lesquels une protéine fluorescente verte (GFP) a été introduite dans le locus de SPI1 ont permis de suivre son modèle d'expression dans toute la différenciation hématopoïétique (209, 487, 491, 492). Bien qu'un faible niveau de SPI1 soit détecté dans les LT-HSC (Lin-Sca1+kit+CD34-), son expression augmente à mesure que les cellules deviennent des MPP (Lin- Sca1+ kit+ CD34+) et se développent ensuite en CMP et CLP (487, 492, 493). A ce stade, le niveau d'expression de SPI1 est décisif pour permettre la progression vers la lignée myéloïde ou lymphoïde (494, 495). La culture de cellules progénitrices de foie fœtal (Lin-) avec des cytokines favorisant la lymphopoïèse génère principalement des cellules pro-B, tandis que la surexpression de SPI1 dans ces cellules biaise considérablement l'engagement vers la lignée myéloïde et entraîne une augmentation marquée de la production des macrophages (496). Ainsi, on estime que les concentrations de SPI1 sont environ huit fois plus élevées dans les macrophages que dans les LB (209, 496).

SPI1 influence non seulement l'engagement des MPP vers la lignée myéloïde ou lymphoïde, mais régule aussi vers quelles voies de différenciation potentielles vont s'engager les CMP et CLP. Le réglage précis des niveaux d'expression de SPI1 dans les CMP est important pour orienter la différenciation vers les MEP ou GMP. SPI1 est exprimé à des niveaux élevés dans les GMP et oriente l'engagement vers les lignées neutrophile et monocyttaire. En revanche, son expression est régulée à la baisse dans les MEP, ce qui est essentiel pour le développement des mégacaryocytes et des globules rouges (12, 209, 492, 493). Des souris "knockdown", chez lesquelles l'expression de SPI1 a été réduite de 20%, développent toutes un blocage du développement des macrophages et des LB, une instabilité génomique et une leucémie aiguë myéloïde (LAM) (488). Dans la lignée lymphoïde, la différenciation des CLP en LB ou LT s'accompagne d'une diminution des niveaux d'expression de SPI1, qui s'éteignent progressivement à mesure que les LT se développent. En revanche, l'expression de SPI1 est maintenue à un niveau faible au cours de la lymphopoïèse B (209, 492, 496). Toutes ces études indiquent que le contrôle des niveaux d'expression de SPI1 est essentiel au développement hématopoïétique normal et déterminant pour l'engagement des progéniteurs vers les lignées myéloïde ou lymphoïde.



Des études *in vitro* de transfection transitoire ont démontré que l'activité de SPI1 dans la lignée myéloïde dépend d'une séquence promotrice de 334 bp allant d'environ 170 bp en amont des principaux TSS à 180 bp en aval. Les principaux sites de liaison dans cette région sont ceux d'OCT1, de SP1, de SPI1 et également de GATA. Celui de SPI1 joue un rôle crucial dans l'expression du gène SPI1 lui-même dans les cellules myéloïdes (468, 497, 498). Des éléments de régulation distaux sont aussi primordiaux pour l'expression de SPI1, notamment une région en amont de 14 kb du gène. En effet, des souris transgéniques comprenant seulement la région promotrice de SPI1 n'exprimaient pas SPI1 à des niveaux similaires à celles comprenant la région promotrice et des séquences de 35 kb en amont et aval du gène (499). Des tests ultérieurs d'hypersensibilité à la DNase I ont identifié un élément régulateur 14 kb en amont du gène (URE, upstream regulatory element) conférant une activité plus de 100 fois supérieure à celle observée avec le promoteur de SPI1 seul (499). Un site de liaison SPI1 a été identifié dans cette région URE et des expériences de mutagenèse de ce site de liaison entraînaient une perte de fonction de l'URE dans les cellules myéloïdes. D'autres FTs peuvent lier la région URE notamment ELF1, FLI1, RUNX1 et C/EBP $\alpha$ . Ces résultats suggèrent qu'il pourrait y avoir deux sites autorégulateurs positifs dans le gène SPI1 : un dans l'URE et un dans la région promotrice.

#### iv. Partenaires

Malgré l'importance de SPI1, ce n'est pas le seul élément déterminant du développement hématopoïétique. Diverses protéines ont été identifiées comme partenaires interagissant avec SPI1. Ces protéines partenaires comprennent d'autres FTs, des cofacteurs non liés à l'ADN, des facteurs de remodelage de la chromatine et des protéines impliquées dans la régulation du cycle cellulaire. L'activation médiée par SPI1 de programmes transcriptionnels spécifiques de lignées hématopoïétiques est généralement associée à des interactions synergiques avec d'autres régulateurs transcriptionnels. SPI1 peut interagir physiquement entre autres avec (i) des facteurs généraux de transcription (TBP), (ii) des FTs de l'hématopoïèse précoce (GATA2 et RUNX1), (iii) de l'érythropoïèse (GATA1) et (iv) d'autres lignées (C/EBP $\alpha$ , C/EBP $\beta$ , IRF4/8 et JUN). Une liste des protéines interagissant avec SPI1 est présentée dans la **Figure 10**.

SPI1 peut soit activer soit réprimer la transcription des gènes, notion soutenue par une étude *in vivo* montrant que suite à la délétion génétique de SPI1 dans les CSH, l'expression de 225 gènes était régulée à la hausse et celle de 97 à la baisse (500). Ce contrôle différentiel de l'expression des gènes est en partie médié par la capacité de SPI1 à construire différents complexes avec un certain nombre de partenaires protéiques, stimulant ou inhibant ainsi la transcription. L'un des premiers partenaires



identifiés de SPI1 a été TBP (474). Grâce à TBP, SPI1 est en mesure de recruter la machinerie transcriptionnelle sur ses sites de liaison au sein des promoteurs et enhancers. En effet, de nombreux gènes cibles de SPI1 n'ont pas de séquences TATA-box et sont donc incapables de se lier directement à TBP. SPI1 peut former des complexes avec un certain nombre de protéines qui peuvent modifier directement la chromatine ou l'ADN. Par exemple, SPI1 peut se lier aux histones acétyltransférases CREB (CBP) et p300 afin d'activer la transcription (421, 501). Il peut aussi se lier à un complexe d'histones désacétylases composé d'HDAC1 et de Sin3a (mSIN3a), et, avec ce complexe, à MeCP2 qui se lie au CpG méthylé pour réprimer les gènes (502-504). Un lien possible entre les séquences d'ADN méthylé et SPI1 a également été suggéré par l'observation que SPI1 est capable de former un complexe avec les ADN méthyltransférases DNMT3A/B (503). Plus récemment, il a été montré que TET2 et TET3 pouvaient agir en synergie avec les FTs E2A et SPI1 afin de promouvoir la déméthylation de l'ADN et l'accessibilité de la chromatine au niveau de régions enhancers spécifiques des LB (505). SPI1 pourrait agir comme une plateforme recrutant, au niveau de ses gènes cibles, différentes protéines modifiant la chromatine, ayant des fonctions activatrices ou répressives de la transcription. L'affinité de liaison des facteurs ETS en général semble pouvoir être modifiée par l'état de méthylation de l'ADN. Il n'existait ainsi pas de différence d'affinité de liaison à un ADN non méthylé entre SPI1 et ETS1 alors que de grandes variations étaient en revanche observées sur des ADN hémi- ou totalement méthylés (506).

Aux premiers stades de la différenciation hématopoïétique (CSH et progéniteurs multipotents), les protéines de la famille C/EBP et le FT RUNX1 jouent un rôle important de concert avec SPI1. Plusieurs gènes codant pour des récepteurs à des facteurs de croissance essentiels de la lignée myéloïde, tels que M-CSF ou GM-CSF, sont ainsi régulés par le couple SPI1-C/EBP $\alpha$  (482, 507). RUNX1 est un FT critique pour le développement du pool définitif de CSH et a également un rôle dans la spécification de certaines lignées hématopoïétiques. De plus, l'expression de SPI1 est elle-même activée par l'interaction protéique SPI1-RUNX1, via leurs liaisons sur l'élément régulateur en amont de 14 kb du gène *SPI1* (508).

La différenciation complète des progéniteurs myéloïdes jusqu'aux cellules les plus matures requiert l'augmentation constante de l'expression de SPI1 (509). Des niveaux d'expression élevés de SPI1, conjointement avec l'action de C/EBP $\alpha$  ou  $\beta$ , peuvent permettre de transdifférencier certaines lignées en macrophages (510, 511). Au cours de la myélopoïèse, SPI1 semble collaborer dynamiquement avec différents partenaires afin orchestrer les différentes étapes de différenciation. Ainsi, la collaboration de SPI1 avec C/EBP $\alpha$  permet l'expression d'un certain nombre de gènes importants pour la



différenciation myéloïde précoce, notamment jusqu'aux stades de CMP et GMP, mais n'est plus nécessaire par la suite (482, 512, 513). Une fois cette étape de différenciation franchie, SPI1 interagirait avec C/EBP $\beta$ , facteur primordial pour les étapes ultérieures de différenciation myéloïde (514). Un complexe répressif connu pour être important dans le développement des neutrophiles et des macrophages est celui entre SPI1 et GFI1 (515). GFI1 peut réprimer les gènes cibles induits par SPI1, spécifiques de la lignée macrophagique mais n'a que peu d'action sur ceux de la lignée neutrophile ; cet effet distinct étant lié aux concentrations de SPI1 nécessaires pour activer ces réseaux de gènes.

Un autre exemple d'interaction de SPI1 avec un FT spécifique de lignée est son complexe avec GATA1. GATA1 est exprimé au sein des lignées érythroïdes et mégacaryocytaires (516) et un acteur primordial de la différenciation érythroïde, les souris KO pour *Gata1* étant dépourvues de progéniteurs érythroblastiques (517). SPI1 et GATA1 se répriment mutuellement (518). Dans les progéniteurs myéloïdes, SPI1 se lie à GATA1 et bloque physiquement l'activité de liaison de GATA1 à l'ADN (519). L'expression de SPI1 est régulée à la baisse au stade érythroblastique (478), et l'échec de sa régulation à la baisse entraîne un blocage de la différenciation érythroïde et le développement ultérieur de leucémie aiguë érythroblastique. Une condition préalable importante pour la différenciation érythroïde est la capacité de GATA1 à inhiber la liaison des co-activateurs à SPI1 afin ainsi de bloquer son activité transcriptionnelle.

L'importance dans la lignée lymphoïde B de l'interaction de SPI1 avec IRF4/8 est décrite dans les paragraphes 2.g.viii et 2.h.i.

## **v. Rôle dans la lymphopoïèse B**

### **1. Précoce**

SPI1 est un régulateur clé de l'engagement entre les lignées myéloïde et lymphoïde (13, 496). La concentration de SPI1 est la plus élevée dans les cellules myéloïdes, où elle joue un rôle majeur dans l'induction du programme génétique spécifique de cette lignée (218). L'expression de SPI1 est réduite d'environ 10 fois au début de la lymphopoïèse B, et cette faible expression est maintenue tout au long du processus de différenciation des LB (209, 492). La régulation appropriée de l'expression de SPI1 est essentielle au processus d'engagement de la lignée, sa dérégulation entraînant dans certaines lignées un blocage de différenciation et favorisant l'apparition de leucémie aiguë (488, 510, 520, 521). Les concentrations distinctes de SPI1 dans les progéniteurs myéloïdes et lymphoïdes activent différemment un réseau régulateur de gènes impliquant SPI1, IKAROS mais aussi EGR1 et GFI1 (494, 522, 523).





Dans ce modèle, l'activité d'IKAROS et de GF11 sont à l'origine d'une faible concentration de SPI1, ce qui entraîne l'activation d'EBF et du programme lymphocytaire B. D'autres FTs dont E2A, MYB et MEF2C participent aussi à cette régulation (524-526). SPI1 régule également l'expression de FLT3 et IL7R, deux récepteurs cytokiniques clés pour la différenciation myéloïde et lymphoïde (13, 527).

Les embryons murins KO pour *Spi1* ou les souris adultes avec un KO conditionnel de *Spi1* au stade de CSH présentent un défaut de production de LB matures (528). L'inactivation conditionnelle de *Spi1* en aval des CLP (par une approche de transduction rétrovirale *in vitro*) (487) ou des pré-B (CD19-Cre) permet le développement des LB matures (214, 215). Les CLP KO pour *Spi1* génèrent des LB et expriment des gènes spécifiques de la lignée lymphoïde B à des niveaux comparables à ceux de CLP sauvages en culture. *In vivo*, la délétion ciblée de *Spi1* après le stade pré-B ne perturbe pas l'expression des Ig ni la réponse des cellules à divers agents mitogènes. De plus, des analyses par CHIP-seq des sites de liaison de SPI1 ont montré que SPI1 lie plus de sites ADN au stade de CLP que de pro-B (45 000 vs. 17 000) (218). Ces données suggèrent que c'est entre les stades CSH et CLP que l'expression de SPI1 est indispensable pour l'initiation de la lymphopoïèse B, ce qui a été confirmé récemment (529). SPI1 n'est ainsi pas essentiel pour la maturation des LB une fois qu'il a dirigé les progéniteurs vers le stade de CLP. Une explication possible est que d'autres membres de la famille des FTs ETS peuvent avoir une redondance fonctionnelle avec SPI1 dans les LB (530, 531). C'est le cas par exemple de SPIB, qui lie une séquence nucléotidique similaire à SPI1. Les souris *Spi1*<sup>+/-</sup> *SpiB*<sup>-/-</sup> ont un nombre nettement réduit de LB (532), et l'inactivation de *Spi1* et *SpiB* dans les progéniteurs lymphoïdes B entraîne un blocage au stade pré-B et le développement de LAL pré-B (530). Les mêmes conséquences sont observées dans les cellules B dépourvues de SPI1/IRF4 ou IRF4/IRF8, démontrant qu'un complexe ETS/IRF contrôle la lymphopoïèse B précoce (533, 534).

## 2. Tardive

Bien que SPI1 continue d'être exprimé tout au long de la différenciation lymphoïde B, il ne semble pas qu'il soit indispensable au processus de maturation des LB (cf. ci-dessus). Deux des partenaires d'interaction de SPI1, IRF4 et IRF8, jouent un rôle essentiel dans différentes étapes de la différenciation des LB, en particulier la réponse du CG et la différenciation en plasmocytes, potentiellement par leurs liaisons avec SPI1/SPIB et aux séquences d'ADN composites ETS-IRF (108, 109, 180, 535). Des éléments composites ETS-IRF ont été identifiés dans plusieurs éléments de régulation importants spécifiques des LB, notamment les gènes *IGH* et *IGL*, ainsi qu'*IRF4*, *MEF2C* et *PAX5* (181, 236, 426, 526). SPI1 et SPIB régulent





l'expression de nombreux composants du complexe BCR et sont nécessaires à une signalisation BCR normale (224, 532). Ces deux FTs ETS sont exprimés de façon constitutive tout au long du développement des LB, l'expression de SpiB s'éteignant rapidement lors de la différenciation des plasmocytes, tandis que Spi1 reste exprimé à une faible concentration et peut, au moins en culture, encore se lier à plusieurs milliers de sites ADN dans le génome (536, 537). Les souris double KO pour *Spi1* et *Spib* dans les LB matures (CD23-Cre) ne génèrent pas de CG et d'Ac de haute affinité après immunisation. Les LB issus de ces souris présentent un défaut de survie après stimulation via les récepteurs CD40 ou TLR, malgré une différenciation plasmocytaire paradoxalement augmentée. Spi1 et Spib régulent l'expression de nombreux composants de la voie de signalisation du BCR (Lyn, Syk, Btk, Ptnp6, Card11) et des récepteurs des ligands CD40L, BAFF et TLR (224). Spi1 et Spib permettent aux LB de répondre de façon adéquate aux signaux environnementaux et participent ainsi aux événements cellulaires contrôlant la différenciation lymphoïde B et l'immunité humorale.

SPI1 a également été impliqué dans la différenciation terminale des LB en ASC. Le développement des ASC est coordonné par les interactions antagonistes entre d'un côté PAX5 et BCL6, qui favorisent le programme lymphoïde B, et de l'autre IRF4 et BLIMP1, qui sont nécessaires au développement des ASC (30, 31). SPI1 est exprimé dans les LB activés avant d'être régulé à la baisse, sans être complètement éteint dans les plasmocytes (181). Cette régulation à la baisse de SPI1 semble être importante sur le plan fonctionnel, car les souris dépourvues de miR-155, un microARN ciblant directement la région 3'UTR de SPI1, ont une concentration augmentée de SPI1, des anomalies de la production d'Ac et de la différenciation en ASC (222, 223). miR-155 diminue l'expression de SPI1, entraînant une diminution de celle de PAX5 et une augmentation de celle de BLIMP1, permettant l'initiation du programme des ASC. Dans des modèles murins de KO conditionnel (CD19-Cre), les LB KO pour *Spi1* n'avaient qu'une augmentation modeste du nombre d'ASC alors que la délétion de *Spi1* et d'*Irf8* entraînait une augmentation spectaculaire, suggérant l'importance des éléments composites Ets-Irf dans la différenciation tardive des LB en ASC. Irf8/Spi1 contrôlait cette transition en simultanément activant les composants du programme lymphoïde B, y compris *Bcl6*, et en réprimant les facteurs de promotion en ASC tels que *Prdm1* (181). Compte tenu du rôle établi d'Irf4 comme 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. Ces études (181, 223) confirment que SPI1 fonctionne comme un régulateur négatif de la différenciation terminale des LB.

#### vi. Mécanismes de transformation impliquant *SPI1*

Plusieurs études ont suggéré que des variations dans le niveau d'expression de SPI1 pouvaient contribuer aux mécanismes de leucémo/lymphomagénèse (488, 521). Une modification ciblée de l'URE de *Sfpi/Spi1*, réduisant significativement son niveau d'expression, entraînait le développement de LAM chez la souris (488). Des taux élevés de miR-155, dont *SPI1* est une cible directe, ont été observés dans certains syndromes myélo- et lympho-prolifératifs (LLC, DLBCL, HL) (538, 539). SPI1 pourrait également avoir un rôle indirect dans les cas de mutations de gènes codant pour des FTs partenaires avec lesquels il entretient une interaction dynamique primordiale pour différentes étapes de l'hématopoïèse, comme c'est le cas de GATA1, et participer ainsi aux mécanismes de leucémogénèse (518).

Dans un modèle murin de LAM induites par irradiation, des délétions du chromosome 2 (où se situe le locus génique de *Sfpi/Spi1* chez la souris) ainsi que des mutations du domaine Ets de *Sfpi* sur l'allèle non délété ont été identifiées dans 86% des cas (540). Alors que leurs présences étaient controversées chez l'homme (541-543), une étude récente a confirmé l'existence de mutations acquises perte de fonction de *SPI1* principalement dans le sous-groupe des LAM avec réarrangement de *MLL* (544). 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*).

Plus récemment, deux gènes de fusions impliquant *SPI1* (*STMN1-SPI1* et *TCF7-SPI1*) ont été identifiés dans les LAL-T pédiatriques de haut risque (453). Les cas de LAL-T arborant ces fusions avaient un immunophénotype et un PEG particuliers ainsi qu'un pronostic très défavorable. Ces protéines de fusion conservaient la capacité de lier l'ADN et une activité transcriptionnelle. Leurs expressions dans des progéniteurs murins augmentaient la prolifération cellulaire et provoquaient un blocage de la maturation T (453).

En revanche, aucun évènement génétique impliquant *SPI1* n'a été identifié jusqu'à présent dans les hémopathies lymphoïdes B.



## OBJECTIFS de notre travail

Le cancer résulte de l'accumulation progressive d'anomalies génétiques et épigénétiques responsables d'une croissance et d'une prolifération anormales. Le traitement approprié des processus cancéreux exigent une compréhension approfondie de l'histoire naturelle de la maladie, y compris l'identification et l'ordre d'apparition des mutations, de leurs conséquences fonctionnelles ainsi que des voies de signalisation dérégulées. L'identification et la caractérisation des conséquences fonctionnelles des mutations acquises dans le clone tumoral peuvent avoir des implications thérapeutiques importantes, leur ciblage pouvant conduire dans certaines situations à un meilleur contrôle de la maladie.

Au cours des dix dernières années, l'avènement des techniques de NGS a permis une meilleure caractérisation et compréhension du paysage mutationnel de nombreux cancers et des mécanismes concourant à la transformation tumorale. C'est notamment le cas pour les hémopathies lymphoïdes B. Les analyses de NGS ont en effet permis d'identifier dans ces maladies des mutations dans des gènes impliqués physiologiquement dans la différenciation lymphocytaire B et la biologie du CG. Ainsi, les voies de signalisation que les LB utilisent normalement pour répondre à une stimulation antigénique sont fréquemment détournées par les cellules tumorales afin de favoriser leur survie et leur prolifération. Ces anomalies génétiques peuvent également perturber les programmes transcriptionnels caractéristiques de la contrepartie lymphoïde B physiologique de ces maladies et altérer les régulateurs épigénétiques de la structure chromatinienne et de l'expression génique de ces cellules.

Même si il s'agit d'un cancer rare, la MW a pu également bénéficier de ces avancées technologiques, avec la description de mutations hautement récurrentes dans les gènes *MYD88* et *CXCR4*, chez respectivement 95% et 30-40% des patients atteints de MW. Cependant, le nombre d'échantillons de MW analysés par WES ou WGS, rapportés dans la littérature, reste faible (< 40) et il est donc possible que d'autres mutations récurrentes d'intérêt existent. D'autre part, le lien entre les anomalies génétiques décrites dans la MW, les spécificités de la contrepartie lymphoïde B physiologique de cette maladie et leur impact sur la différenciation lymphoïde B n'est pas clairement établi.

Nous avons souhaité, par une première approche de WES réalisée sur des échantillons tumoraux de patients atteints de MW, rechercher la présence d'autres mutations récurrentes et caractériser leurs conséquences fonctionnelles et leur éventuel rôle dans la différenciation lymphoïde B. Nous avons ainsi identifié une mutation récurrente dans le gène codant pour le facteur de transcription SPI1 de



la famille ETS, survenant chez 6% des patients MW, et décrit les conséquences fonctionnelles de ce mutant sur sa capacité à lier à l'ADN et ses motifs de liaison, sur la prolifération cellulaire et la différenciation lymphocytaire B.

Les résultats de notre travail sont présentés sous la forme d'un article récemment soumis et actuellement en révision.



## RESULTATS

### Title: A recurrent activating missense mutation in Waldenström macroglobulinemia affects the DNA binding sequence of the ETS transcription factor SPI1 and enhances cellular proliferation

Damien Roos-Weil <sup>1,2,3,4,5</sup>, Hussein Ghamlouch <sup>1,2,3,4</sup>, Véronique Della-Valle <sup>1,2,3,4</sup>, M'boyba K Diop <sup>1,2,4,6</sup>, Cécile Hérate <sup>1,2,3,4</sup>, Diane Lara <sup>1,2,3,4</sup>, Eric Durot <sup>1,2,3,4</sup>, Rima Haddad <sup>7</sup>, Elena Mylonas <sup>1,2,3,4,8</sup>, Frederik Damm <sup>8</sup>, Françoise Pflumio <sup>7</sup>, Paresh Vyas <sup>9</sup>, Olivier Elemento <sup>10</sup>, Philippe Dessen <sup>1,2,3,4,6</sup>, Valérie Camara-Clayette <sup>1,2,6</sup>, Véronique Leblond <sup>5</sup>, Vincent Ribrag <sup>1,11</sup>, Pascale Cornillet-Lefebvre <sup>12</sup>, Philippe Rameau <sup>6</sup>, Nabih Azar <sup>5</sup>, Pierre Morel <sup>13</sup>, Bilyana Stoilova <sup>9</sup>, Marlen Metzner <sup>9</sup>, Thomas Mercher <sup>1,2,3,4</sup>, Said Aoufouchi <sup>2,3,4,14</sup>, Nathalie Droin <sup>1,2,3,4</sup>, Christel Guillouf <sup>1,2,3,4</sup>, Florence Nguyen-Khac <sup>15,16,17,\*</sup>, Olivier A Bernard <sup>1,2,3,4,\*</sup>

#### Affiliations

<sup>1</sup>INSERM U1170, Gustave Roussy, Villejuif, France

<sup>2</sup>Gustave Roussy, Villejuif, France

<sup>3</sup>Université Paris-Saclay, Villejuif, France

<sup>4</sup>Equipe labellisée Ligue Nationale Contre le Cancer

<sup>5</sup>Sorbonne Universités, UPMC Univ Paris 06, AP-HP, GRC-11, Groupe de recherche clinique sur les hémopathies lymphoïdes (GRECHY), Hôpital Pitié-Salpêtrière, APHP, Paris, France

<sup>6</sup>AMMICA, INSERM US23/CNRS UMS3655, Gustave Roussy, Villejuif, France

<sup>7</sup>Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA) DSV-IRCM-SCSR-LSHL ; Université Paris Diderot Sorbonne Paris Cité UMR 967, Fontenay-aux-Roses, France ; Université Paris Sud UMR 967, Fontenay-aux-Roses, France

<sup>8</sup>Department 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

<sup>9</sup>MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, NIHR Oxford Biomedical Research Centre Haematology Theme, Radcliffe Department of Medicine and Department of Haematology, Oxford University and Oxford University Hospitals NHS Foundation Trust, UK

<sup>10</sup>Caryl and Israel Englander Institute for Precision Medicine, Institute for Computational Biomedicine, Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA.

<sup>11</sup>DITEP Gustave Roussy, Villejuif, Paris, France

<sup>12</sup>Laboratoire d'hématologie, Pôle de biologie, CHU de Reims - Hôpital Robert Debré, Avenue du Général Koenig, Reims, France

<sup>13</sup>Centre Hospitalier Dr. Schaffner, Lens, France

<sup>14</sup>CNRS UMR8200, Gustave Roussy, Villejuif, France

<sup>15</sup>INSERM UMRS 1138, Cell Death and Drug Resistance in Lymphoproliferative Disorders Team, Centre de Recherche des Cordeliers, Paris, France.



<sup>16</sup>Sorbonne Universités, UPMC Univ Paris 06, UMR\_S 1138, Centre de Recherche des Cordeliers, Paris, France.

<sup>17</sup>Service d'Hématologie Biologique, GH Pitié-Salpêtrière, AP-HP, Paris, France.

**Running title:** Activating mutation in *SPI1/PU.1*

**Keywords:** PU.1, SPI1, ETS, B-cell differentiation, Waldenström macroglobulinemia.

**Additional information:**

Financial support: This work was supported by grants from INSERM, Institut National du Cancer (INCa), 2013-PLBIO-09, 2016-PLBIO-068, INCa-DGOS-INSERM\_12551, Association Laurette Fugain and Comité Val-d'Oise de la Ligue contre le cancer. PV 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. oxfbrc-2012-1)

\*Shared senior authorship

Corresponding authors:

Olivier A. Bernard, INSERM U1170, Gustave Roussy, 39 rue Camille Desmoulins, 94805 Villejuif Cedex, France; e-mail: olivier.bernard@inserm.fr

Florence Nguyen-Khac, Unité de Cytogénétique Hématologique, Service d'Hématologie Biologique, GH Pitié-Salpêtrière/Charles Foix, 83 Bd de l'Hôpital, Paris 75013, France.

The authors declare no potential conflicts of interest

Word count, total number of figures and tables: 6016 words; 4 figures.



**Abstract:**

The ETS-domain transcription factors are divided into subfamilies based on protein similarities, DNA binding sequences and interaction with cofactors. They are regulated by extracellular clues and contribute to a variety of cellular processes, including proliferation and transformation. *ETS* genes are targeted by oncogenic processes through chromosomal translocations and copy number gains. The *PU.1/SPI1* gene is also targeted by inactivating point mutations in human myeloid malignancies. We investigated a recurrent somatic missense mutation (Q226E) of the *PU.1/SPI1* gene in Waldenström macroglobulinemia, a human B-cell lymphoproliferative disorder. The mutation changes DNA binding of the mutant protein from classical SPI1 to ETS1-like sequences, shifting the balance from binding to promoter regions from enhancers. Increased binding by mutant SPI1 at promoters activates gene expression of intracellular signaling pathways typically promoted by other ETS factor family members. The functional consequences are decreased terminal B-cell differentiation in a model cell line and primary samples. In summary, we describe oncogenic subversion of transcription factor function through subtle alteration DNA binding specificity leading to differentiation arrest.

**Significance:** The demonstration that a somatic point mutation subtly changes the balance of genome binding provides a mechanistic paradigm for how missense mutations in transcription factor genes may be oncogenic in human tumors.





## 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 (30, 31). 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 (545-547). The ETS family of transcription factors share a conserved winged helix-turn-helix DNA binding domain (DBD) and bind to a consensus DNA sequence containing a core 5'-GGA(A/T)-3' motif (381). ETS factors regulate gene expression through binding to ETS motifs or to composite sites through interaction with other transcription factors, including IRF4, IRF8, C/EBPa and b, and c-Jun (380). SPI1 is an ETS-family member that plays a crucial role in hematopoiesis (487, 488, 528, 540), including different steps of B-cell development, coordinating the proliferation and differentiation of B-lymphoid progenitors (548), but also acting as an essential repressor of terminal B-cell differentiation (181, 223, 224). In humans, inactivating *SPI1* mutations have been described in acute myeloid leukemia (AML) (541, 544) including the *MLL (KMT2A)*-rearranged subgroup and more recently, gene fusions involving *SPI1* have been identified in high-risk pediatric T-cell acute lymphoblastic leukemia (453).

Biological processes driving GC reaction or generation of antibody-secreting plasma cells can be somatically altered leading to GC-derived B-cell lymphomas (199) and multiple myeloma (31). Molecular studies have identified activating mutations that target genes encoding proteins that convey proliferative signals downstream of antigen exposure and BCR signaling and include the cytosolic signaling adapter protein MYD88 (71, 92). 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- $\kappa$ B and JAK signaling pathways (92). Specific expression of the mutant protein in B-cell differentiation triggers with a long latency lymphoproliferative diseases and occasional clonal B-cell transformation, which displays features of post GC diffuse large B-cell lymphomas (DLBCL) (352).

Waldenström's macroglobulinemia (WM) is defined as an IgM-secreting and bone marrow-infiltrating lymphoplasmacytic lymphoma, presumed to be related to memory B-cells. A MYD88 mutation is present in more than 95% of WM cases (94). Although it is an indolent disease and despite recent advances in treatment, WM remains incurable with currently available therapies, most patients dying of disease progression. There is a need to develop more specific targeted therapies, and 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 to 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 an associated block in B-cell differentiation.



## Methods

### Patients and materials

A total of 85 patients meeting the diagnostic criteria for Waldenström's macroglobulinemia (WM) (549), were enrolled in this study. Written consent to bone marrow (BM) and biological analyses were obtained 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 (333).

### Cell sorting and nucleic acid extraction.

Samples were processed as described (115). Tumoral cells were defined by the positivity of CD19 and monotypic light chain (LC) while control samples were CD3-positive T lymphocytes. The cellular populations were sorted using an Influx cell sorter (BD Biosciences, San Jose, CA, USA). The sorted cells were subjected to an additional round of sorting to provide a cell fraction purity of more than 98% and subjected to DNA/RNA extraction using the AllPrep DNA/RNA Kit (Qiagen).

### Whole-exome sequencing.

Whole-exome sequencing was conducted on paired-samples from 16 patients as described previously (115). DNA from sorted CD19+ LC+ tumor cells and CD3+ non-tumor cells were used for exome capture using SureSelect All Exon V4, V5 or clinical research exome kit (Agilent Technologies). Reads were mapped to the reference genome hg19 using the Burrows–Wheeler Aligner (BWA) alignment tool version 0.7.10. PCR duplicates were removed using Picard tools – MarkDuplicates (version 1.119). Local realignment around indels and base quality score recalibration were performed using GATK 3.2 (Genome Analysis ToolKit). Reads with a mapping quality score < 30 were ignored. The detection of candidate somatic mutations was performed according to previously described algorithms with minor modifications (115). The number of reads containing single-nucleotide variations (SNV) and indels, in both tumor and reference samples, was enumerated using VarScan (v2.3.7). The mean coverage in the targeted regions was respectively 100 and 93X for tumoral and control samples (**Supplementary Table S2**). Candidate non-synonymous exonic mutations were validated using Sanger sequencing and/or targeted deep resequencing (MiSeq instrument, Illumina). We designed primers surrounding the candidate somatic variants sites using Primer3 plus (<http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>). For targeted sequencing fragments were PCR-amplified from genomic DNA from mononuclear blood cells or from sorted fractions. In total, we validated 265/299 (89%) somatic mutations (**Supplementary Table S3**). Coding mutations and the distribution of nucleotide interchanges are shown in **Supplementary Table S2** and **Supplementary Figure S1a-c**, respectively. Sequencing data deposition is in progress.

### Targeted sequencing.

Twenty genes found recurrently mutated in our WES cohort and/or in the literature (94, 306) were selected for sequencing an extended series of WM samples. The whole coding regions of these genes were submitted for Ion AmpliSeq Designer (ThermoFisher Scientific). The designed panel included 539



amplicons covering 91 393 bases. Our sequence data covered 98% of these bases (**Supplementary Fig. S1d** and **Supplementary Table S4**). Ten ng of genomic DNA per primer pool from 72 BM CD19+LC+ tumor cells and 39 paired BM CD3+ lymphocytes were analyzed, including 3 tumors controls pairs analyzed by whole-exome sequencing. Libraries were generated with addition of paired-end adaptors (NEXTflex, Bio Scientific) before paired-end sequencing (2 x 250 bp reads) using an Illumina MiSeq flow cell and the onboard cluster method (Illumina, San Diego, CA). Targeted sequencing was analyzed similarly to WES. The mean read depth within the targeted regions was 1799X.

### **RNA sequencing.**

RNA-seq was performed as described (115, 550). Sequences were aligned to the reference genome with TopHat2 version 2.0.9 (or 2.0.14 for mice datasets) and Bowtie2 version 2.1.0. The number of reads per genes (RefSeq database) was counted with HTSeq-count using the "union" mode. The counts were then normalized with the DESeq2 method, which takes into account the library size of each sample. DESeq2 was also used to detect differential expression signatures between two conditions. We filtered genes with a fold-change superior to 1.5 and an adjusted *P*-value of 0.01. Molecular pathway and gene set enrichment analyses (GSEA) of differentially expressed genes were performed using GSEA (2.2.3), the Molecular Signatures Database (6.2, Broad Institute) and the Lymphoid Signature Database (551).

### **Cell culture and drugs.**

All cell lines were maintained in a humidified 5% CO<sub>2</sub> incubator at 37°C in standard conditions. OCI-Ly10 cells (a kind gift of K. Leroy (Institut Cochin, Paris, France)) were checked for the presence of reported MYD88 p.L265P and CD79A p.L191\_G208del mutations. Ba/F3 cells (a kind gift of P. Dubreuil (CRCM, Marseille, France)) were supplemented with 1 ng/ml of rmlL3. The bromodomain and extraterminal domain (BET) inhibitor JQ1 is a kind gift of J. Bradner, (Boston, USA). Lenalidomide was purchased from Aldrich Chemistry (reference CDS022536).

### **Vectors, retroviral and lentiviral production, transduction and transplantation.**

*SPI1* (NM\_003120.2) and *MYD88* (NM\_002468.4) cDNAs were subcloned into murine stem cell virus (MSCV)-IRES-GFP and into pcDNA3. Mutations were introduced using the Quick Change Kit, following the manufacturer's instructions. Every PCR-amplified or mutagenized fragment was checked by sequencing. Viral particles and transduction procedures were performed as described previously (552). The lentiviral pV81 vector with doxycycline inducible expression of WT or mutant SPI1 and GFP (pV81/GFP-SPI1) was constructed after a series of cloning steps. The pV81 backbone was described elsewhere (553). Lentiviral particles were produced in HEK293T cells, which were co-transfected with plasmids of interest, along with packaging and VSV G envelope expression plasmids using jetPRIME (Polyplus transfection, Ozyme). Supernatants were collected at 48h and concentrated by ultra-centrifugation. Bone marrow transplantation assays and hematopoietic differentiation analyses were performed as described previously (554).

OCI-Ly10 cells were transduced with WT or mutated pV81/ GFP-SPI1 and flow-sorted twice gating on viable GFP+ cells 24h after addition of doxycycline (1000 ng/mL) (Sigma, reference D-9891) to reach more than 90% of GFP+ cells. These flow-sorted cells were used for subsequent analyses. In this model,

### **Université Paris-Saclay**

Espace Technologique / Immeuble Discovery  
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France



both mutant and WT proteins were expressed at similar levels and the GFP-fusion decreased similarly upon doxycycline removal (data not shown).

### **Electrophoretic mobility shift assay (EMSA).**

The *SPI1*-consensus binding site, containing two GGAA motifs (PU.1 box, underlined) (5' GAG GAA CCG CGG AAT CAG GAA GAA CTG GCC GGC GC 3') has been described elsewhere (393). Biotinylated probes were incubated in the presence or absence of *in vitro* translation (IVT) protein extracts from pcDNA3 plasmids encoding wild-type (WT) or mutated (QE) SPI1 and empty pcDNA3 (unprogrammed lysate, up. lysate), prepared using TNT Quick Coupled Transcription/Translation Systems (Promega). Reactions were further incubated with or without a 160X excess of a specific or non-specific competitor. Specific competitor is the non-biotinylated probe and nonspecific competitor is the non-biotinylated sequence mutated in one GGAA motif (underlined: 5' GAG GAA CCG CGG AAT CAC CAA CTG GCC GGC GC 3') (393). Samples were separated by electrophoresis; DNA and protein complexes were transferred to HyBond N+ membranes (Amersham). UV cross-linked membranes were visualized according with LightShift® Chemiluminescent EMSA Kit (Thermo-Fisher) and images were captured and quantified using an ImageQuant LAS4000 detector (GE Healthcare Life Sciences). Two independent experiments were performed.

### **Luciferase assay.**

The Dual-Luciferase Reporter Assay System (Promega) was used for luciferase assays. HEK293T cells were transiently transfected with equimolar amounts of empty pcDNA3 vector (control), pcDNA3 vectors encoding WT or mutant SPI1, along with a *cFes-luc* reporter construct (393) and a renilla luciferase control reporter. Cells were harvested 72h after transfection and the Dual Luciferase Reporter Assay was used for detecting luciferase activity (Promega). Three independent experiments in optimized conditions were performed (once in duplicate, twice in triplicate).

### **Cellular proliferation assays.**

For cellular proliferation, doxycycline (1000 ng/mL) was added to exponentially growing OCI-Ly10 cells. Viable GFP+ cells were flow-sorted after 24h and plated in replicates in 48-well plates, 150,000 cells/well, in 300 µL of media containing doxycycline (1000 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 were checked by flow-cytometry at H48. Two independent experiments were performed in triplicate.

### **Drug sensitivity assays.**

OCI-Ly10 cells were seeded at 300,000 cells/well in triplicate in 300 µL of fresh media, in a 48-well plate. JQ1 and lenalidomide were dissolved in DMSO and equal volumes of diluted drug were added to cells to reach the indicated final concentrations. OCI-Ly10 cells were treated with doxycycline (1000 ng/mL) and the indicated concentrations of JQ1 or lenalidomide for respectively 24h and 48h. Doxycycline, JQ1 and



lenalidomide were added to the media at the same time (H0). Three independent experiments were performed in triplicate.

### **Flow cytometry and cell cycle analyses.**

Cells analyzed by flow cytometry were stained in PBS 1X 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. Whole BM, blood, thymic and spleen cells from mice were harvested, subjected to red blood cell lysis, and stained for FACS analysis. Cells were analyzed on a FACS Cantoll with FACS Diva software (BD Biosciences). For cell-cycle analyses, cells were fixed in 70% cold ethanol and stained with propidium iodide (PI) and 7-AAD. Quantification of different cycle phases was performed using Watson pragmatic algorithm (FlowJo). Alternatively, cells were labeled by the addition of 10 mM BrdU (5-bromodeoxyuridine) to the culture medium for 1 h 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 was adapted from the MagnaChIP kit protocol (Millipore) and (555). Immunoprecipitated and input DNA fragments were end-repaired, extended, ligated with indexed adaptors (NEXTflex, Bio Scientific), size selected, amplified by PCR for 10 cycles and sequenced using Illumina HiSeq 2000 (Illumina, San Diego, CA). Reads were aligned *in toto* to human genome hg19 with BWA alignment tool (v0.7.5a) and peak calling assessed using MACS 2.0 with a p value  $<10^{-20}$ . Annotation and motif analyses were performed with HOMER (v4.7.2), with a p value  $<10^{-30}$ . Genomic localization of peaks in OCI-Ly10 and GM12878 cell lines was performed using Refseq database and ChromHMM tool (556).

### **Murine B-cell transduction and culture.**

All experiments were performed on 8-10 weeks 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 1X ; 2% FBS), CD43<sup>-</sup> cells were separated using AutoMACS Pro Separator instrument (Miltenyi) with anti-CD43 antibody coated magnetic beads and comprised more than 95% B220<sup>+</sup> cells (**Supplementary Fig. S5c**). CD43<sup>-</sup> cells were activated by *in vitro* stimulation with anti-IgM (7 µg/ml) and anti-CD40 (10 µg/ml) for 24h before transduction with the retroviral vectors, and transduced B cells were cultured in media containing anti-CD40 (10 µg/ml) for 36h after spin-infection. The number of viable cells was determined by counts of Trypan blue–negative cells and the percentage of GFP<sup>+</sup> Sytox<sup>-</sup> cells was determined by flow cytometry. Washed, transduced cells were then cultured in 48-well plates at  $10^6$  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<sup>+</sup> cells during culture was determined by counting Trypan blue–negative cells and measure the percentage of GFP<sup>+</sup> cells, CD19<sup>+</sup> and/or CD138<sup>+</sup> GFP<sup>+</sup> cells by flow cytometry.



### Western blotting.

Total cell lysates from different cultured cells were prepared by direct lysis into 2X Laemmli Sample Buffer (Bio-Rad), betamercaptoethanol and sonication. Proteins lysates were separated by a 4–12% NuPAGE Novex Bis-Tris gradient gel (Invitrogen), transferred to nitrocellulose membranes (GE Healthcare) by semidry transfer (Bio-Rad) and probed with the following primary antibodies: anti-SPI1 (Santa-Cruz, sc-352), anti-HA.11 tag (Eurogentec, clone 16B12) and anti-MYD88 (Abcam, ab133739). Antibody to  $\beta$ -actin (A3853, Sigma-Aldrich) was used as an internal control. HRP-conjugated antibodies [anti-rabbit IgG (NA934V, GE) and anti-mouse IgG (NA931V, GE)], and ECL Plus Kit (RPN2132, GE) were used for detection.

### Immunofluorescence.

Immunofluorescence experiments were performed as previously described (557). Briefly, OCI-Ly10 were fixed (in suspension) for 15 minutes using a solution of 4% paraformaldehyde in PBS at 37°C, or for 15 minutes using a solution of 10% trichloroacetic acid in water at 4°C. Cells were washed in PBS, incubated in a solution of 0.05 M  $\text{NH}_4\text{Cl}$  in PBS for 15 minutes and permeabilised with washing buffer (0.1% Triton X-100, 3% BSA, in PBS) for 10 minutes. Cells were resuspended in 20  $\mu\text{l}$  fluorescence-preserving mounting medium (Dako North America, Inc., Carpinteria, CA) containing 2.5  $\mu\text{g/ml}$  Hoechst 33342 (Molecular Probes). Images were processed using ImageJ software.

### Statistics.

Data are depicted as mean  $\pm$  s.d. as precised. *P*-values were calculated with the two-tailed unpaired Student's *t*-test unless otherwise specified. All statistical analyses were performed with Prism software version 6.0 (GraphPad Software, Inc., La Jolla, CA, USA). The size of animal cohorts was based on our previous studies as well as the published literature. Neither randomization nor blinding was used, since all animal experiments were performed with homogeneous strain, age and similar variance.

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 was defined as time between diagnosis and first treatment and compared across groups using log-rank tests. Overall survival was defined as 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 as statistically significant. The SAS 9.3 (SAS, Inc.) and R 3.0.2 software packages were used.





## Results

### Identification of a recurrent *SPI1* mutation in WM

To gain insights into WM physio-pathology, we used whole exome sequencing (WES) to compare tumor and CD3+ germline coding sequences of a discovery cohort of 16 WM patients and identified a total of 265 somatic mutations affecting 230 different genes, corresponding to a median of 16.5 mutations per patient (**Methods** and **Supplementary Table S1, S2, S3** and **Supplementary Fig. S1a-c**). In addition, we analyzed an extended cohort of 69 WM patients, by targeted sequencing of the exonic regions of 20 genes (see **Supplementary Table S4** for genomic coordinates of targeted exonic regions and **Supplementary Fig. S1d** for the list of genes). The distribution and frequency of gene alterations in the whole cohort of 85 WM patients is shown in **Figure 1a**. As previously described (94, 306), *MYD88* (81/85, 95%), *CXCR4* (24/85, 28%) and *ARID1A* (7/85, 8%) were confirmed to be part of the most frequently mutated genes in WM. Among the 13 genes found mutated in at least 2 patients, 12 (*MYD88, IGLL5, CXCR4, ARID1A, CD79B, TP53, LTB, HIST1H1E, TRAF2, ROBO2, TRRAP, EZH2*) had been previously identified in the context of WM or other B-lymphoid malignancies (94, 154, 306, 558).

We focused on a novel mutation in the transcription factor gene *SPI1* (**Fig. 1b**), which was observed in five of 85 WM patients (5.8%) and in both the discovery (n = 2/16) and validation (n = 3/69) cohorts. All patients carried the same single nucleotide change (C for 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**, middle). The Q residue at position 226 within the alpha 3 helix of the DBD is a well conserved amino acid specific for the class III subgroup of the ETS family of proteins, while E is present at this position in classes I and IIa proteins (**Fig. 1b**, bottom). The mutations were observed on both DNA and RNA sequences from flow-sorted WM tumor cells, either at clonal or subclonal levels, with a variant allele frequency (VAF) ranging from 5 to 70%; **Fig. 1b**, top and middle). The transcription levels of *SPI1* RNA expression were similar in *SPI1* WT and QE samples (**Fig. 1c**).

This mutation was not reported in the COSMIC databases, suggesting it might be specific to WM. Its stereotyped and heterozygous nature suggests it may be among the very few *SPI1* activating mutations in humans.

### *SPI1* mutant maintains DNA binding capacity and enhances transactivation and cellular proliferation

We first used electrophoretic mobility shift assays (EMSA) to investigate *in vitro* the binding of *SPI1* QE protein to the ETS-DNA motif derived from a known *SPI1*-target (**Fig. 1d** and **Supplementary Fig. S2a**) (393). *SPI1* QE binds DNA and estimation of the intensity of the shifted complex suggests a stronger binding of *SPI1* QE, with respect to *SPI1* WT.

To evaluate the transcriptional potential of the *SPI1* mutant protein, we performed luciferase assays with a reporter gene containing the *Fes* promoter region (393). Dose-dependent transcriptional activity was observed for both WT and QE proteins (**Fig. 1e** and **Supplementary Fig. S2b**). Up to two fold more activation was observed for *SPI1* QE.





We next investigated the effect of mutant SPI1 on cellular parameters of a *MYD88* LP human B lymphoma cell line (OCI-Ly10). OCI-Ly10 cells 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 48h. Both proteins were expressed at similar levels, as judge by GFP fluorescence and western-blotting evaluation (**Supplementary Fig. S2c**). In those conditions, both SPI1 WT and QE-expressing cells expanded cell number significantly more than cells harboring the Ctl 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. 2a**)). Cell cycle analyses showed that ectopic SPI1 expression decreased the percentage of cells in G1 phase and proportionately increased cells in G2/M phase (**Fig. 2b–e**). Similar variations of higher amplitude were observed in cells expressing SPI1 QE. These data indicate that mutant SPI1 promotes cell expansion and this is likely via cellular proliferation in manner that is similar to, but greater than SPI1 WT.

To characterize the genome wide effects of the mutant SPI1, we used RNA sequencing (RNA-seq) to compare OCI-Ly10 cells expressing SPI1 WT or QE proteins. By gene set enrichment analysis (GSEA), we observed that, as compared to 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), toll-like receptor (TLR) and phosphoinositide-3-kinase (PI3K) signaling pathways (**Fig. 2f**, see **Supplementary Table S14** for the list of enriched signatures). When expression signatures from the Lymphochip database (551) were used, we noted differences in target genes of transcription factors that are key regulators of B-cell differentiation, including XBP1, PRDM1 and IRF4 (**Fig. 2f**). These data suggest that SPI1 QE affects cellular proliferation to a higher extent than SPI1 WT.

### **SPI1 mutation changes the DNA binding specificity of the protein**

We also assessed the DNA binding profiles of both proteins by ChIP-seq in OCI-Ly10-transduced cells using anti-GFP and anti-SPI1 antibodies. Results with both antibodies were similar. Anti-GFP results are presented in **Fig. 2g-i**, **Supplementary Figures S3a-b** and **Supplementary Tables S6-9** and those with anti-SPI1 are summarized in **Supplementary Figures S3c-d** and **Supplementary Tables S10-13**. SPI1 WT and QE proteins showed different genome-wide DNA-binding patterns as genomic regions bound by SPI1 QE were significantly enriched in promoters (defined by RefSeq (**Fig. 2g**), or in active promoters and strong enhancers (defined by ChromHMM annotations (556) (**Fig. 2h**)). DNA motif analyses of SPI1 WT- and QE-specific peaks identified different preferential motifs. As expected, the classical SPI1 motif was the top of the motif list bound by SPI1 WT (**Fig. 2i**) whereas it ranked in 12th position for SPI1 QE (**Supplementary Tables S6-9**). The top hit of motifs recognized by SPI1 QE is a common class I/IIa (ETS1/GABPA/ELF1/ELK1) motif (**Fig. 2i**). These two motifs differ in a single nucleotide (G or C) at position -2 of the core GGA(A/T). Classes I/IIa motifs were observed in 77 and 95 % (Chi2  $P = 2.2 \times 10^{-90}$ ) and class III motifs in 85 and 57% (Chi2  $P = 8.0 \times 10^{-163}$ ) of SPI1 WT and QE peaks, respectively (**Supplementary Tables S6-9**). These data are in line with 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 (398). Our data are also consistent with frequent binding of classes I/IIa proteins at promoters, in contrast with what is observed for SPI1, in an EBV-transformed B-cell line (**Supplementary Fig. S3e-f**).



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

### ***SPI1* mutation cooperates with *MYD88* L265P and interferes with terminal B-cell maturation**

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<sup>+</sup> cells in bone marrow and myeloid differentiation at the expense of B-lymphoid differentiation (**Supplementary Fig. S4a-c**). 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 to *SPI1* WT (Hardy's fraction A vs. B and C; **Supplementary Fig. S4d**). These data are consistent with previous reports describing *SPI1* as a determinant of lymphoid and myeloid fate choices; high *SPI1* levels favoring the development of myeloid cells while B-cell development require its down-regulation (559). We however did not observe lymphoid transformation within an 18-month follow-up.

Because *SPI1* QE mutations are detected in association with *MYD88* LP mutations, we constructed retroviral vectors driving expression of *MYD88* and *SPI1* (WT or mutant) either alone or in combination (**Fig. 3a**) to investigate potential oncogenic cooperation. All constructs expressed at comparable levels of proteins when tested in murine cells (**Supplementary Fig. S5a&b**). Naïve B-lymphocytes were purified from WT mice, stimulated *in vitro* and transduced by retroviral constructs (**Figure 3b**) and the proportion of fluorescent cells was monitored in the culture, as described in (560). 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 (**Figure 3c**). As already reported (560), 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 co-expressed with *MYD88* LP, indicating cooperation between those two mutants. We also investigated the consequences on lymphoid maturation toward plasmablastic/plasmacytic cells by examining the proportion of GFP<sup>+</sup> B-cells (CD19) expressing CD138, a marker of terminal B-cell differentiation. Expression of *MYD88* did not markedly affect the proportion of CD138<sup>+</sup> B-cells, whereas expression of *SPI1*, WT or QE, alone or in combination with *MYD88* LP, decreased the proportion of this cell population (**Fig. 3d-e**). The drop in the ratio of CD138<sup>+</sup> B-cells was more important when *SPI1* was expressed alone. Conversely to cellular growth, WT or mutant *SPI1* did not significantly affect the ratio of differentiation 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. S5d-f**).

To confirm our observations that *SPI1* QE interferes with B-cell differentiation and their potential relevance in primary WM samples, we performed GSEA analyses of RNA-seq data specifically for immunologic



signatures comparing *SPI1* WT (n = 29) and QE (n = 3) WM samples with available RNA (**Supplementary Table S15**). The top hit was a gene set comparing naïve B cells to plasma cells (**Fig. 3f**). The prevalence of this signature was further exemplified by 7 of the top 20 most enriched gene sets (out of more than 4000 total gene sets) that represent a comparison of plasma cells with either naïve 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-cells 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. 3g**). 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 on the left side of the 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 also indicated that the *SPI1* QE samples were markedly less differentiated than the other WM samples (**Fig. 3h**). We observed similar results when we compared *SPI1* QE samples along with the four WM samples, which segregated with them in PCA, to the rest of the cohort (**Supplementary Fig. S6a**).

### **Proliferation of *SPI1* QE-expressing cells requires MYC and IRF4 pathways**

To support our findings on the functional consequences of *SPI1* mutation, we extended GSEA analyses 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. 4a**). Most of these enrichments were also confirmed when we performed GSEA analyses only considering genes bound by SPI1 in OCI-Ly10 active promoters' regions (**Supplementary Fig. S7a-c**). Notably, signatures of MYC targets were among the most enriched signatures in *SPI1* QE-expressing OCI-Ly10 cells and *SPI1* QE WM samples. We thus interrogated if *SPI1* QE-expressing OCI-Ly10 cells were more sensitive to MYC inhibition, using a BET inhibitor (JQ1) treatment. As previously described (561), JQ1 treatment induced significant cell death in OCI-Ly10 cells from 48h of exposition (data not shown). While increasing concentrations of JQ1 did not affect viability of control cells at 24h, the proportion of viable *SPI1* WT and QE-expressing OCI-Ly10 cells decreased with a significantly more pronounced effect in *SPI1* QE condition (**Fig. 4b**). We confirmed that JQ1 treatment lowered MYC expression in this context (data not shown). These results indicate that MYC pathway is of key importance in *SPI1* QE cell proliferation, which can be abrogated by the use of BET inhibitor. A similar validation approach was performed to probe the activation of the IRF4 pathway (**Fig. 4a**). Because of its promising clinical effects and its targeting of IRF4 (261), we evaluated the impact of lenalidomide (Len) effect on cell growth. Len treatment of cell lines revealed a higher sensitivity of *SPI1* QE-, with respect to *SPI1* WT-expressing cells (**Fig. 4c**).

### **Clinical outcomes of *SPI1* mutated patients**

Finally, we evaluated the association of *SPI1* mutations with the clinical outcome of WM patients (**Fig. 4d**). When compared to *SPI1* WT cases, *SPI1* mutated cases exhibited significantly shorter overall survival (OS), with median survival times that were respectively not reached and of 20.7 months (hazard ratio (HR)



= 7.39, 95% confidence interval (95% CI) = 1.5-89, log-rank  $P = 0.0004$ ). *SPI1* mutated patients showed a trend to shorter time to relapse and to next treatment (TNT) compared to *SPI1* WT cases although it was not significant (log-rank  $P = 0.13$  and  $0.36$ ). OS of *SPI1* mutated patients were also shorter than that of *CXCR4* mutated patients, a previously described aggressive subgroup of WM patients (**Supplementary Fig. S8**) (log-rank  $P = 0.0003$ ).

## 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. ETS proteins bind to a core GGA(A/T) motif. Q226 is present in class III proteins and responsible for the binding of a G at position -2 of the core, whereas the E, present at this position in classes I and IIa proteins, binds to C at position -2 to the core (398). Of note, mutagenesis of the Q226 into E facilitated the crystallization of the ETS domain of SPI1, suggesting that this specific amino-acid change improves DNA binding of SPI1 (562). In keeping with that, we showed that the mutant Q226E protein retained DNA binding capacity, exhibited a 20% stronger transcriptional activity as compared to the SPI1 WT form, and increased the proliferation rate in different cellular contexts. The mutation modifies the DNA binding specificity of the protein in terms of sequence, reminiscent of specificities harbored by others 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 class I/IIa type-profile, with binding sites more frequently located in promoters (40-60%) with active promoters (531, 563).

An important difference between SPI1 and classes I/IIa ETS factors is that DNA binding and transcription activities of the latter are tightly regulated by intracellular signaling pathways, such as calcium and MAPK activity (413, 564) whereas the activity of SPI1 is not, suggesting that promoters bound by the SPI1 QE may escape normal regulation. Mutant SPI1 expression was associated with enrichment in transcription signatures related to cell cycle, proliferation, intra-cellular signaling pathways and also MYC and IRF4-related transcriptional programs. Moreover, as SPI1 interacts with epigenetic regulators, such as CBP (565) and TET2 (505, 566), SPI1 QE may also recruit these epigenetic regulators to promoter regions from enhancers and in this way consolidate altered gene expression through epigenetic changes.

Missense mutations in transcription factor genes in hematopoietic malignancies, such as *ELF4*, *EGR2*, *MEF2C* or *IRF4* in mature B-cell disorders (115, 123, 567), 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*) (568), 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 (432, 435). However, abnormalities of ETS proteins have rarely been identified in mature B-cell



malignancies, with the exception of *ELF4* mutations in chronic lymphocytic leukemia (567), translocations involving *SPIB* and gain of chromosome 11q24, encompassing *ETS1* and *FLI1*, in DLBCL (258, 454, 455).

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 (546, 563), Fli1 (569), Gabpa (570), and the three class III factors Spi1, SpiB, and SpiC (571). Survey of their expression during human and murine B-cell maturation suggest the highest activity of the class I/IIa ETS/Ets factors in GC B-cells, together with the highest level of ETS1/Ets1 transcription (**Supplementary Fig. S6b-e**).

Loss of Ets1 leads to premature B-cell differentiation into antibody-secreting cells (225) and recent reports have emphasized that Spi1 and SpiB physiologically act as negative regulators of plasma cell development and sensors to environmental stimuli (181, 223, 224). Considering that WM is characterized by a differentiation block at the transition from post-GC lymphocyte to plasma cell, we investigated and 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, in line with recent reports identifying Spi1 as a gatekeeper for this terminal maturation step. Moreover, our gene expression analyses revealed that mutated *SPI1* WM cases were markedly 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 analyses of RNA-seq from normal B-cell populations show a significant enrichment of classes I/IIa motifs in naïve, GC or memory B-cell populations versus plasma cells (**Supplementary Fig. S9**). It also remains to be shown if others missense mutations in genes of ETS family (*ETS1*, *ELK1*) (154) or transcriptions factors (*EGR1*, *EGR2*, *MEF2C*) (115, 123, 572) identified in others B-cell lymphoproliferative disorders may have the same impact on DNA recognition and B-cell differentiation.

Finally, *SPI1* Q226E mutations are associated with *MYD88* LP, but not with del6q, the most frequent cytogenetic abnormality of WM. The impact of *SPI1* Q226E on biology is underscored by the association of mutation acquisition to poor prognosis, although this impact on survival needs to be validated. The enrichment of MYC pathway associated with *SPI1* Q226E-expressing cases is of interest, as targeting of this pathway by BET inhibitor could potentially be of therapeutic benefit in these high-risk patients. Similarly, *in vitro* experiments show that cell expressing *SPI1* Q226E are sensitive to lenalidomide, which is known to indirectly target the important B-lymphocyte transcription factors IRF4, IKZF1 and IKZF3 (573). Our results thus identify mutant *SPI1* as a new genetic biomarker for poorer prognosis WM patients and point to novel treatment options for this patient group.



## Authors' Contributions

**Conception and design:** D Roos-Weil, F Nguyen-Khac, O A Bernard

**Development of methodology:** D Roos-Weil, H Ghamlouch, V Della-Valle, C Hérate, R Haddad, E Mylonas, F Damm, O Elemento, T Mercher, S Aoufouchi, F Pflumio, N Droin, P Vyas, C Guillouf, F Nguyen-Khac, O A Bernard

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** D Roos-Weil, H Ghamlouch, V Della-Valle, M K Diop, C Hérate, D Lara, E Durot, R Haddad, E Mylonas, F Damm, P Dessen, V Camara-Clayette, P Cornillet-Lefebvre, V Ribrag, P Rameau, N Azar, P Morel, B Stoilova, M Metzner, O Elemento, T Mercher, S Aoufouchi, V Leblond, F Pflumio, N Droin, P Vyas, C Guillouf, F Nguyen-Khac, O A Bernard

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** D Roos-Weil, H Ghamlouch, V Della-Valle, M K Diop, C Hérate, E Mylonas, P Dessen, O Elemento, T Mercher, S Aoufouchi, F Pflumio, N Droin, C Guillouf, F Nguyen-Khac, O A Bernard

**Writing, review, and/or revision of the manuscript:** D Roos-Weil, H Ghamlouch, V Della-Valle, M K Diop, C Hérate, D Lara, E Durot, E Mylonas, F Damm, P Dessen, V Camara-Clayette, P Cornillet-Lefebvre, V Ribrag, P Rameau, N Azar, P Morel, B Stoilova, M Metzner, O Elemento, T Mercher, S Aoufouchi, V Leblond, F Pflumio, N Droin, P Vyas, C Guillouf, F Nguyen-Khac, O A Bernard

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** D Roos-Weil, H Ghamlouch, V Della-Valle, M K Diop, C Hérate, R Haddad, E Mylonas, F Damm, P Dessen, O Elemento, T Mercher, S Aoufouchi, F Pflumio, N Droin, P Vyas, C Guillouf, F Nguyen-Khac, O A Bernard

**Study supervision:** D Roos-Weil, F Nguyen-Khac, O A Bernard

## Acknowledgements:

We thank the members of the Gustave Roussy platforms, including Patrick Gonin for excellent mouse care, and Yann Lecluse and Philippe Rameau for flow cytometry. This work was supported by grants from INSERM, Institut National du Cancer (INCa), 2013-PLBIO-09, 2016-PLBIO-068, INCa-DGOS-INSERM\_12551, Association Laurette Fugain and Comité Val-d'Oise de la Ligue contre le cancer. HG is a recipient of successive grants from région Ile de France and Fondation de France (No 00067113). CH is a recipient of a région Ile de France grant. ED is a recipient of Fondation pour la Recherche Médicale grant. We thank the Blueprint consortium for access to RNA-seq data, and W. Cook (La Trobe University, Melbourne, Australia) and R. DeKoter (Western University in London, Ontario) for cell lines.





## References

1. De Silva NS, Klein U. Dynamics of B cells in germinal centres. *Nat Rev Immunol* **2015**;15(3):137-48 doi 10.1038/nri3804.
2. Nutt SL, Hodgkin PD, Tarlinton DM, Corcoran LM. The generation of antibody-secreting plasma cells. *Nat Rev Immunol* **2015**;15(3):160-71 doi 10.1038/nri3795.
3. Recaladin T, Fear DJ. Transcription factors regulating B cell fate in the germinal centre. *Clinical and experimental immunology* **2016**;183(1):65-75 doi 10.1111/cei.12702.
4. Garrett-Sinha LA. Review of Ets1 structure, function, and roles in immunity. *Cell Mol Life Sci* **2013**;70(18):3375-90 doi 10.1007/s00018-012-1243-7.
5. Bories JC, Willerford DM, Grevin 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**;377(6550):635-8 doi 10.1038/377635a0.
6. Sharrocks AD. The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol* **2001**;2(11):827-37 doi 10.1038/35099076.
7. Hollenhorst PC, McIntosh LP, Graves BJ. Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu Rev Biochem* **2011**;80:437-71 doi 10.1146/annurev.biochem.79.081507.103945.
8. Scott EW, Simon MC, Anastasi J, Singh H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **1994**;265(5178):1573-7.
9. 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(12):3437-44 doi 10.1182/blood-2004-06-2234.
10. 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(6):624-30 doi 10.1038/ng1361.
11. 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**;106(5):1590-600 doi 10.1182/blood-2005-03-0860.
12. Nutt SL, Kee BL. The transcriptional regulation of B cell lineage commitment. *Immunity* **2007**;26(6):715-25 doi 10.1016/j.immuni.2007.05.010.
13. 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(11):2169-81 doi 10.1084/jem.20140425.
14. 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(11):2183-98 doi 10.1084/jem.20140338.
15. 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**;8(1):1426 doi 10.1038/s41467-017-01605-1.
16. 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(3):998-1007.
17. Lavalley VP, Baccelli I, Krosch 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(9):1030-7 doi 10.1038/ng.3371.





18. 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(8):1274-81 doi 10.1038/ng.3900.
19. Basso K, Dalla-Favera R. Germinal centres and B cell lymphomagenesis. *Nat Rev Immunol* **2015**;15(3):172-84 doi 10.1038/nri3814.
20. 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**;463(7277):88-92 doi 10.1038/nature08638.
21. Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, *et al.* Oncogenically active MYD88 mutations in human lymphoma. *Nature* **2011**;470(7332):115-9 doi 10.1038/nature09671.
22. 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**;127(22):2732-41 doi 10.1182/blood-2015-11-684183.
23. 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(9):826-33 doi 10.1056/NEJMoa1200710.
24. 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(2):110-5 doi 10.1053/sonc.2003.50082.
25. 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(4):649-54 doi 10.3324/haematol.2012.070458.
26. 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(9):1088-101 doi 10.1158/2159-8290.CD-14-0104.
27. Hunter ZR, Xu L, Yang G, Zhou Y, Liu X, Cao Y, *et al.* The genomic landscape of Waldenstrom macroglobulinemia is characterized by highly recurring MYD88 and WHIM-like CXCR4 mutations, and small somatic deletions associated with B-cell lymphomagenesis. *Blood* **2014**;123(11):1637-46 doi 10.1182/blood-2013-09-525808.
28. 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(8):1513-35 doi 10.1084/jem.20151775.
29. 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 doi 10.1111/j.0105-2896.2006.00373.x.
30. 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(10):4220-6 doi 10.1182/blood-2008-01-136366.
31. Calvo J, BenYoucef A, Baijjer J, Rouyez MC, Pflumio F. Assessment of human multi-potent hematopoietic stem/progenitor cell potential using a single in vitro screening system. *PLoS One* **2012**;7(11):e50495 doi 10.1371/journal.pone.0050495.
32. 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(1):25-38 doi 10.1016/j.ccr.2011.06.003.
33. 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**;11(2):303-13.



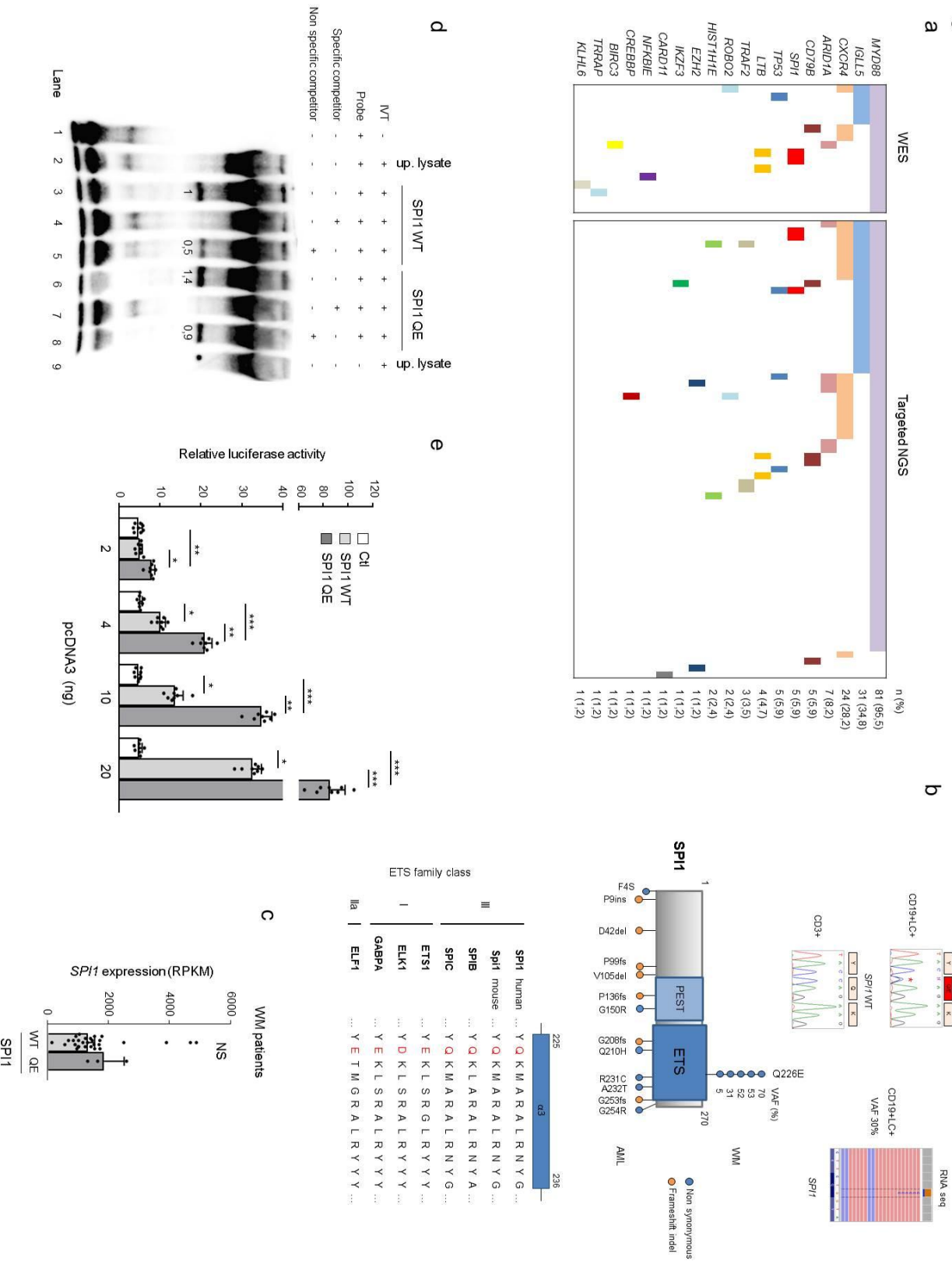
34. Ridinger-Saison M, Boeva V, Rimmele P, Kulakovskiy I, Gallais I, Levavasseur B, *et al.* Spi-1/PU.1 activates transcription through clustered DNA occupancy in erythroleukemia. *Nucleic Acids Res* **2012**;40(18):8927-41 doi 10.1093/nar/gks659.
35. Ernst J, Kellis M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat Protoc* **2017**;12(12):2478-92 doi 10.1038/nprot.2017.124.
36. David MD, Petit D, Bertoglio J. The RhoGAP ARHGAP19 controls cytokinesis and chromosome segregation in T lymphocytes. *J Cell Sci* **2014**;127(Pt 2):400-10 doi 10.1242/jcs.135079.
37. 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(2):481-94 e15 doi 10.1016/j.cell.2017.09.027.
38. Kasar S, Kim J, Improgo R, Tiao G, Polak P, Haradhvala N, *et al.* Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. *Nat Commun* **2015**;6:8866 doi 10.1038/ncomms9866.
39. 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(13):2147-60 doi 10.1038/emboj.2010.106.
40. 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(4):228-35 doi 10.1016/j.smim.2008.08.003.
41. 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(3):413-26 doi 10.1084/jem.20131424.
42. 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(31):11365-70 doi 10.1073/pnas.1411701111.
43. 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(6):723-37 doi 10.1016/j.ccr.2012.05.024.
44. Kodandapani R, Pio F, Ni CZ, Piccialli G, Klemsz M, McKercher S, *et al.* A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. *Nature* **1996**;380(6573):456-60 doi 10.1038/380456a0.
45. 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 doi 10.1186/s12864-015-1303-0.
46. 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 doi 10.3389/fimmu.2017.00383.
47. Cowley DO, Graves BJ. Phosphorylation represses Ets-1 DNA binding by reinforcing autoinhibition. *Genes Dev* **2000**;14(3):366-76.
48. Foulds CE, Nelson ML, Blaszczyk AG, Graves BJ. Ras/mitogen-activated protein kinase signaling activates Ets-1 and Ets-2 by CBP/p300 recruitment. *Mol Cell Biol* **2004**;24(24):10954-64 doi 10.1128/MCB.24.24.10954-10964.2004.
49. 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(11):3729-43.
50. Lio CW, Zhang J, Gonzalez-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**;5 doi 10.7554/eLife.18290.
51. de la Rica L, Rodriguez-Ubrea J, Garcia M, Islam AB, 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 doi 10.1186/gb-2013-14-9-r99.
52. 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(7574):525-30 doi 10.1038/nature15395.
  53. 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**;476(7360):298-303 doi 10.1038/nature10351.
  54. 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(1):108-25 doi 10.1158/2159-8290.cd-17-0532.
  55. 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(5748):644-8 doi 10.1126/science.1117679.
  56. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, *et al.* Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature* **1992**;359(6391):162-5 doi 10.1038/359162a0.
  57. 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(3):633-43 doi 10.1084/jem.20062041.
  58. 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(13):2233-41 doi 10.1182/blood-2013-01-475772.
  59. 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**;43(9):830-7 doi 10.1038/ng.892.
  60. Zhang XK, Moussa O, LaRue A, Bradshaw S, Molano I, Spyropoulos DD, *et al.* The transcription factor Fli-1 modulates marginal zone and follicular B cell development in mice. *J Immunol* **2008**;181(3):1644-54.
  61. Xue HH, Bollenbacher-Reilley 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(4):421-31 doi 10.1016/j.immuni.2007.03.010.
  62. 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(12):7374-84 doi 10.4049/jimmunol.1001413.
  63. 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(2):951-62 doi 10.1074/jbc.M705262200.
  64. Walker BA, Boyle EM, Wardell CP, Murison A, Begum DB, Dahir NM, *et al.* Mutational Spectrum, Copy Number Changes, and Outcome: Results of a Sequencing Study of Patients With Newly Diagnosed Myeloma. *J Clin Oncol* **2015**;33(33):3911-20 doi 10.1200/JCO.2014.59.1503.
  65. Lu G, Middleton RE, Sun H, Naniong M, Ott CJ, Mitsiades CS, *et al.* The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. *Science* **2014**;343(6168):305-9 doi 10.1126/science.1244917.



Figure 1



Figures



Figure 2

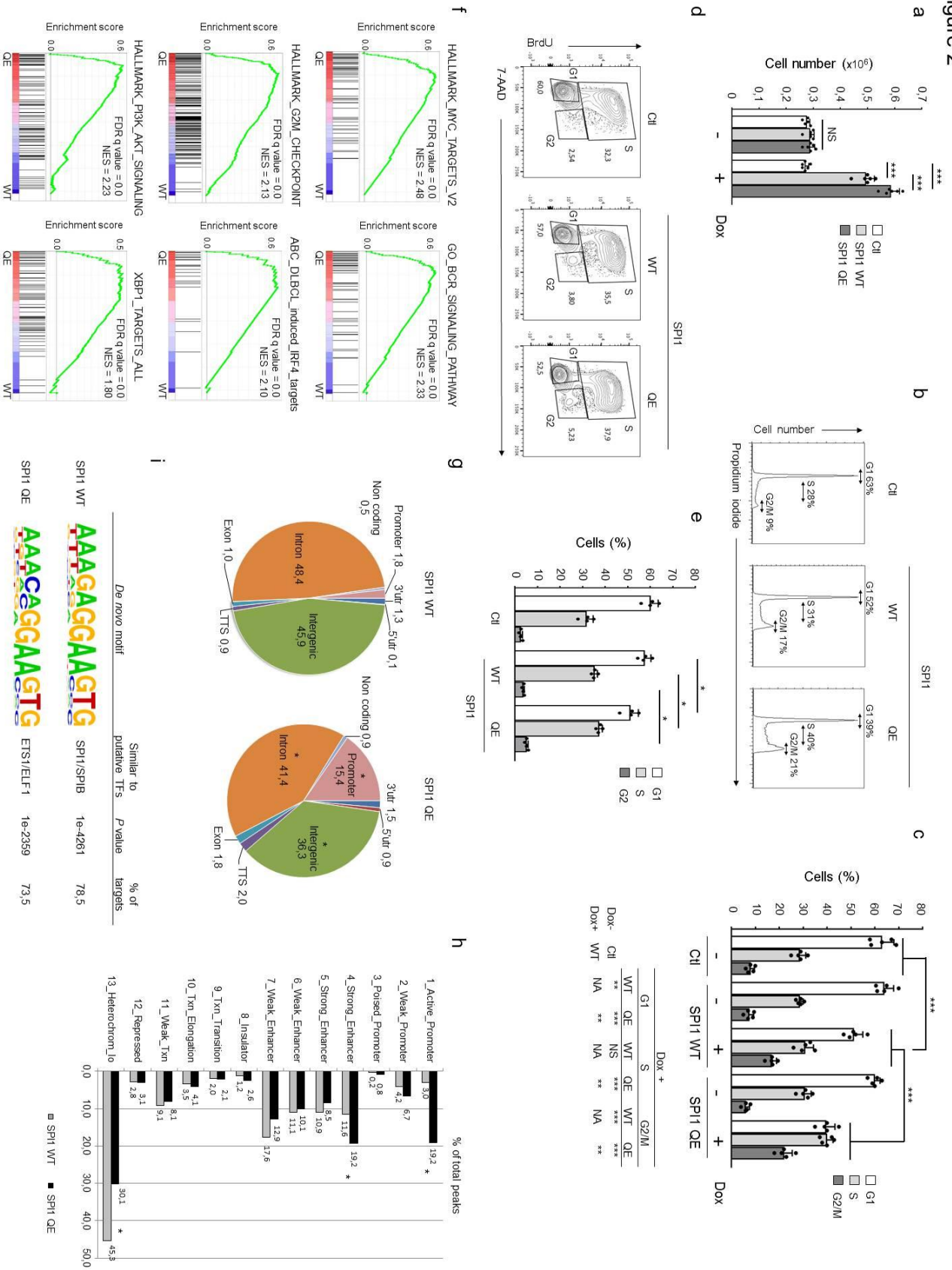




Figure 3

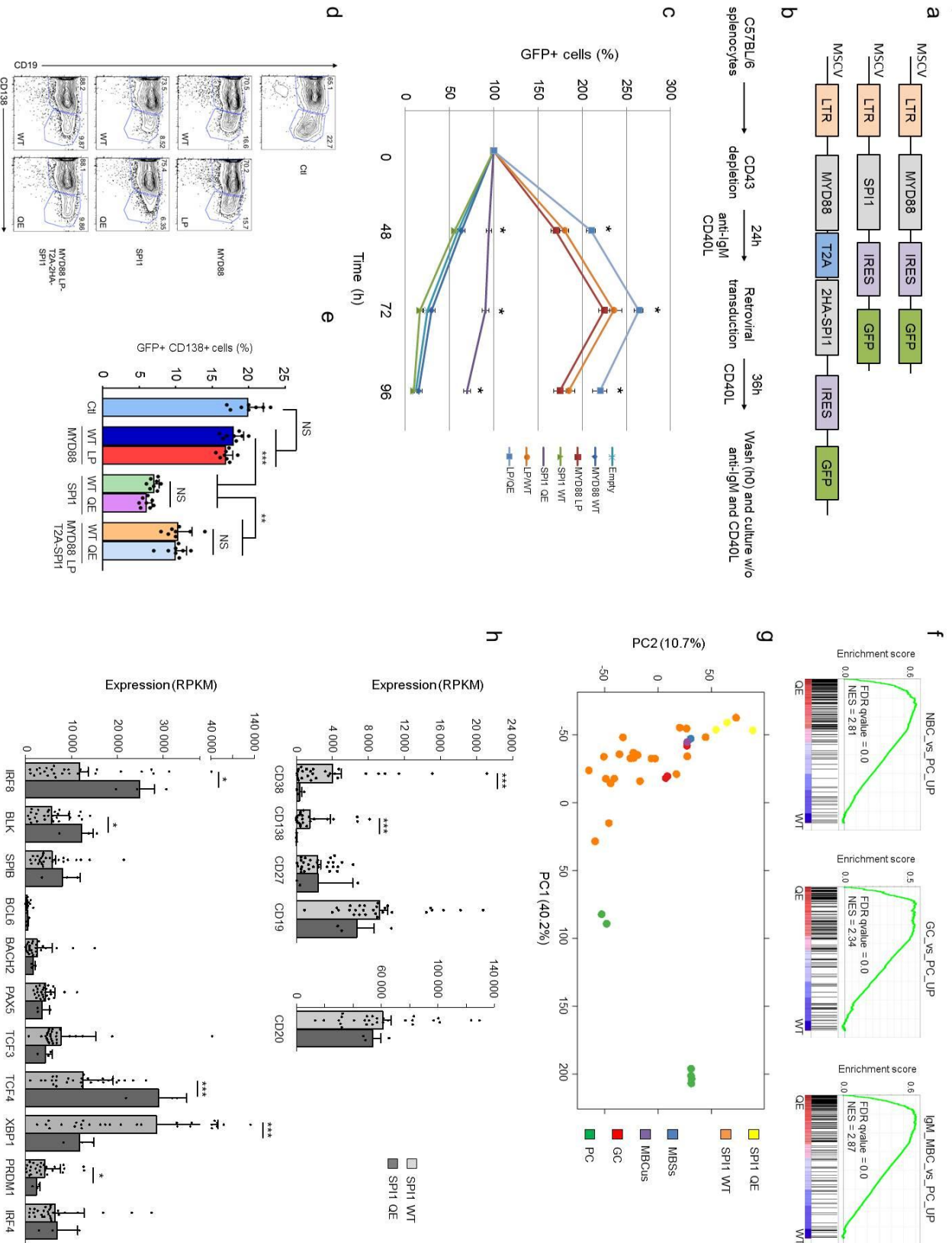
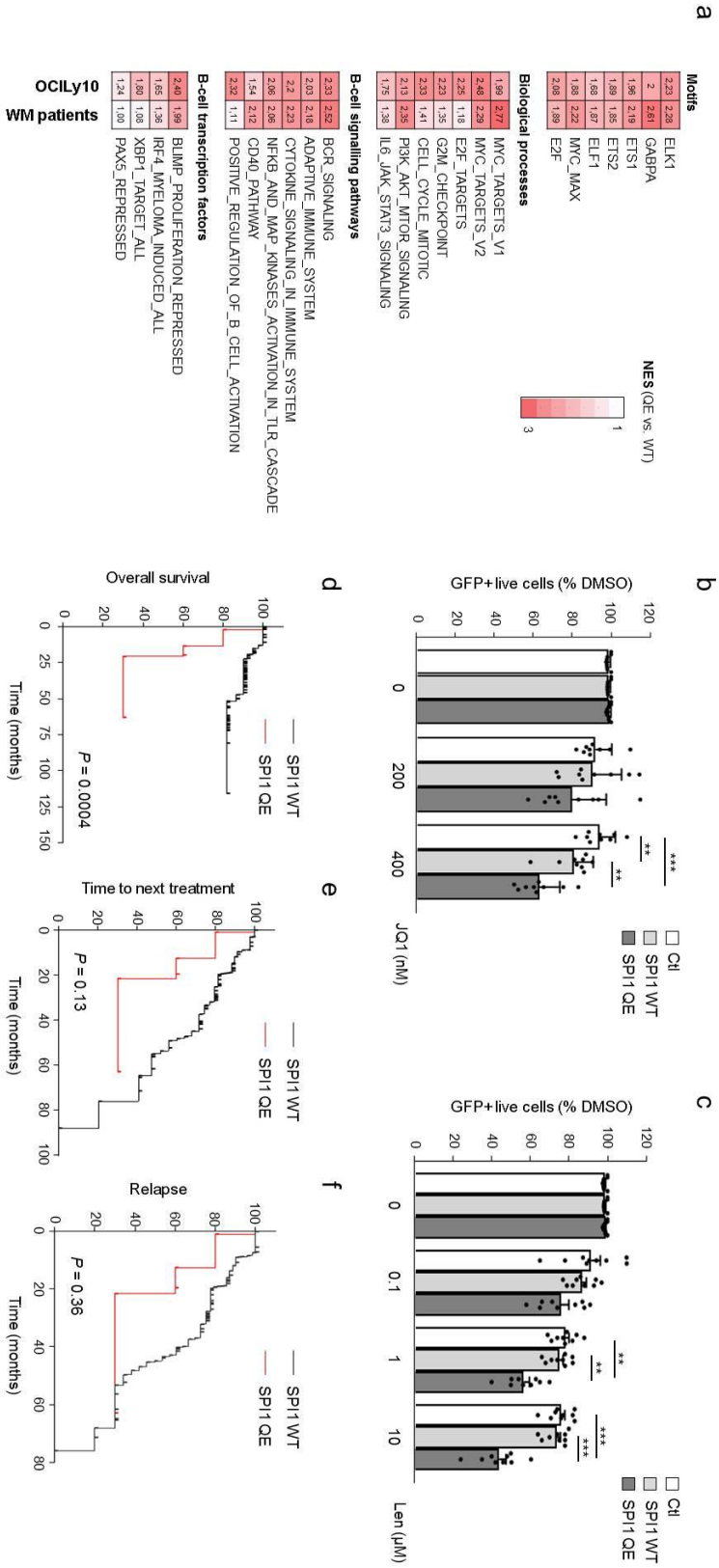


Figure 4





## Legends to the figures

**Figure 1.** Mutational analyses of Waldenström macroglobulinemia (WM) identify an acquired recurrent *SPI1* missense mutation.

(a) Mutational landscape of 85 WM patients analyzed by whole exome (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. WES, whole exome sequencing; NGS, next generation sequencing.

(b) *SPI1* mutations in hematological malignancies.

Top, 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 positive T-cells of the same patient (bottom, left). A red star indicates the mutant base. Expression of the mutant copy in RNA-seq data from the same patient is shown (right). Gln, glutamine; Glu, glutamic acid; VAF: variant allele frequencies.

Middle, locations of *SPI1* mutations identified in WM ( $n = 85$ ) and acute myeloblastic leukemia (AML) ( $n = 541$ ) samples (541, 544). *SPI1* mutations are indicated at their amino acid position, color-coded by type of mutation (orange: frameshift; blue: non synonymous). Inactivating mutations observed in AML are shown beneath the protein, whereas WM QE mutations appear above. The VAF estimated in the 5 WM patients and identified functional domains are indicated. PEST, proline-, glutamic acid-, serine-, and threonine-rich domain; ETS, ETS DNA-binding domain.

Bottom, 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.

(c) Expression levels of *SPI1* transcripts in *SPI1* WT ( $n = 29$ ) and QE ( $n = 3$ ) WM samples. Normalized read counts mapping to *SPI1* exons 1-5 are plotted. Histograms represent median and bars 95% confidence interval. Statistics were performed using Student *t* test.

(d) Mutant *SPI1* binds to ETS motif sequences. Electrophoretic mobility shift assay (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 non-specific competitor, as indicated. The ETS motif-containing probe and competitors are from (393). Sequences are provided in Material and Methods. Estimation of the DNA binding complex intensity in the different conditions is represented. Protein quantities produced by the indicated IVT lysates were similar (**Supplemental Fig. S2a**). EMSA were performed three times. IVT, *in vitro* translation; up., unprogrammed.

(e) Mutant *SPI1* transactivates an ETS–reporter construct.



HEK293T cells transfected with increasing amounts of the indicated (control, SPI1 WT or QE) expression plasmids and fixed amounts of a luciferase reporter gene (carrying three copies of the ETS motif assayed in (d)). Three independent experiments were performed ( $n = 8$ ; mean  $\pm$  s.d.). Relative luciferase activities were normalized with Renilla activity from a co-transfected vector. Levels of SPI1 expression in different conditions regarding type and concentration of transduction were verified (**Supplementary Fig. S2b**). All errors bars represent standard deviation. \* $P \leq 0.05$ , \*\* $P \leq 0.01$ , \*\*\* $P \leq 0.001$ . Ctl, control; s.d: standard deviation.

**Figure 2.** Biological effects of mutant SPI1 in human cell lines.

(a) Induction of SPI1 stimulates cellular proliferation. Quantitation of *in vitro* proliferation for OCI-Ly10 cells transduced with doxycyclin-inducible constructs encoding GFP (control, ctl), or SPI1-GFP fusion proteins (either WT or QE) ( $n = 6$ , mean  $\pm$  S.D.), with (+) or without (-) the 48h addition of doxycycline. The levels of SPI1 proteins and their nuclear localization were checked in OCI-Ly10 bearing the different constructs (**Supplementary Fig. S2c-d**), 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. All errors bars represent standard deviation.

(b-e) Cell cycle analyses in OCI-Ly10 cells transduced with doxycyclin-inducible constructs encoding GFP (control, ctl), or SPI1-GFP fusion proteins (either WT or QE). Representative flow cytometry analyses (b) of propidium iodide (PI) 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 24h 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 24h of doxycycline stimulation. G1, S, G2 cell cycle phases are indicated in representative flow cytometry analyses. The results are shown as means  $\pm$  standard deviation. Statistics were performed using two-tailed Student *t* test. \* $P \leq 0.05$ , \*\* $P \leq 0.01$ , \*\*\* $P \leq 0.001$ . Ctl, control; 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 S14**. FDR, false discovery rate; NES, normalized enrichment score.

(g-i) ChIP-seq analyses in OCI-Ly10 cells transduced with doxycycline-inducible vectors expressing SPI1 WT or QE. Genomic localization, defined by RefSeq (g) or ChromHMM (556) (h) annotations, of DNA sequences (peaks) bound by SPI1 WT (left or grey) and QE (right or black) (GFP) proteins. Different genomic regions and their respective percentage (in respect to the total number of peaks) are indicated. Statistics comparing the repartition of genomic regions were performed using Chi2 test. (i) Commonest DNA motif recognized by SPI1 WT and QE proteins. Homer motif *de novo* 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 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. TF, transcription factors. TTS, transcription termination sites. UTR, untranslated regions.

**Figure 3.** Biological effects of mutant SPI1 in primary murine and human cells.

(a) Graphical illustration of murine stem cell virus (MSCV) constructs used for murine B-cells experiments. RNA and protein expressions of MYD88 and SPI1 from the different constructs were verified and are illustrated in **Supplementary Figure S5a-b**. LTR, long terminal repeats sequence; IRES, internal ribosome entry site.

(b) Scheme of murine B-cells experiment. Procedure details are provided in Material and 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: MYD88 LP, green: SPI1 WT, violet: SPI1 QE, orange: MYD88 LP/SPI1 WT (LP/WT), light blue: MYD88 LP/SPI1 QE (LP/QE)) and cultured for 96h 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  $\pm$  s.d). Error bars indicates standard deviation. Statistics were performed using Student  $t$  test. The LP/QE condition was statically different from MYD88 LP and LP/WT conditions, while SPI1 QE was different from MYD88 WT, SPI1 WT and empty ones.  $*P \leq 0.05$ . LP, L265P.

(d) Representative flow cytometry analyses of CD19+CD138+ cells after 72h of culture. CD19 and CD138 expression analysis is gated on viable (Sytox-) transduced (GFP+) cells.

(e) Percentages of CD19+CD138+ GFP+ cells as means  $\pm$  s.d ( $n = 8$ ) are shown in histograms for each construct.  $**P \leq 0.01$ ,  $***P \leq 0.001$ . Ctl, control; NS, non significant.

(f) GSEA enrichment plots for immunologic signatures using MSigDB, comparing *SPI1* WT ( $n = 29$ ) versus QE ( $n = 3$ ) WM RNAseq samples. The complete list is available in **Supplementary Table S15**. GC, germinal center; MBC, memory B-cells. NBC, naïve B-cells. PC, plasma cells. FDR, false discovery rate. NES, normalized enrichment score.

(g) Principal component analysis of normal B-cells and WM samples. This analysis included plasma cells (PC, green;  $n = 6$ ), switch (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) Expressions of key genes controlling normal B-cell differentiation in *SPI1* WT (light grey,  $n = 29$ ) and QE (dark grey,  $n = 3$ ) WM RNA-seq samples. Error bars indicate standard deviation. Statistics were performed using Student  $t$  test.  $*P \leq 0.05$ ,  $**P \leq 0.01$ ,  $***P \leq 0.001$ .



**Figure 4.**

(a) Integration of GSEA analyses 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 transcriptions factors motifs, intracellular signaling pathways and B-cell biology are shown in the figure. The complete lists are available in **Supplementary Tables S14-15**. NES values are indicated inside each box.

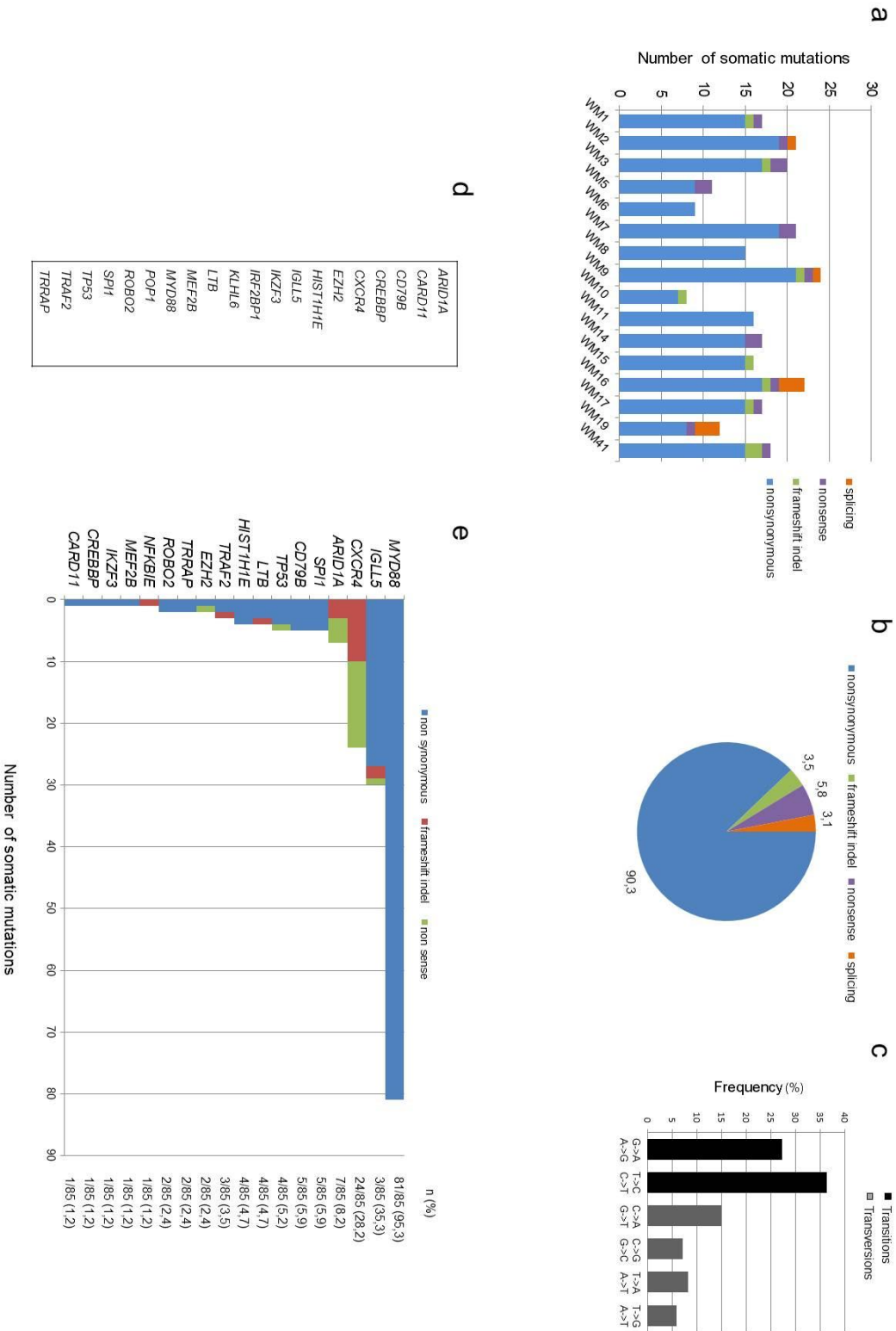
(b) JQ1 and MYC inhibition decrease cellular proliferation of SPI1 QE transduced OCI-Ly10 cells. Viability of OCI-Ly10 cells bearing empty (control) (white), SPI1 WT (light grey) or QE (dark grey) constructs ( $n = 9$ , mean  $\pm$  S.D.). OCI-Ly10 cells were treated for 24h with doxycycline and the indicated concentrations of BET inhibitor (JQ1). Cell viabilities are represented relative to those without JQ1 treatment, set as 100%. 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. All errors bars represent standard deviation.

(c) SPI1 QE-expressing cells are sensitive to lenalidomide. Viability of OCI-Ly10 cells bearing empty (control) (white), SPI1 WT (light grey) or QE (dark grey) constructs ( $n = 9$ , mean  $\pm$  S.D.). OCI-Ly10 cells were treated for 24h with doxycycline and the indicated concentrations of lenalidomide. Cell viabilities are represented relative to those without lenalidomide treatment, set as 100%. Three independent experiments have been performed.  $**P \leq 0.01$ .  $***P \leq 0.001$ . Statistics were performed using two-tailed Student  $t$  test. All errors bars represent standard deviation. Len, lenalidomide.

(d-f) Clinical impact of *SPI1* Q226E mutations in WM. Kaplan–Meier survival curves of overall survival (d), time to next treatment (e) and relapse-free survival (f) for patients with ( $n = 5$ ) and without ( $n = 80$ ) *SPI1* Q226E mutations.  $P$  values are indicated and based on the log-rank test.

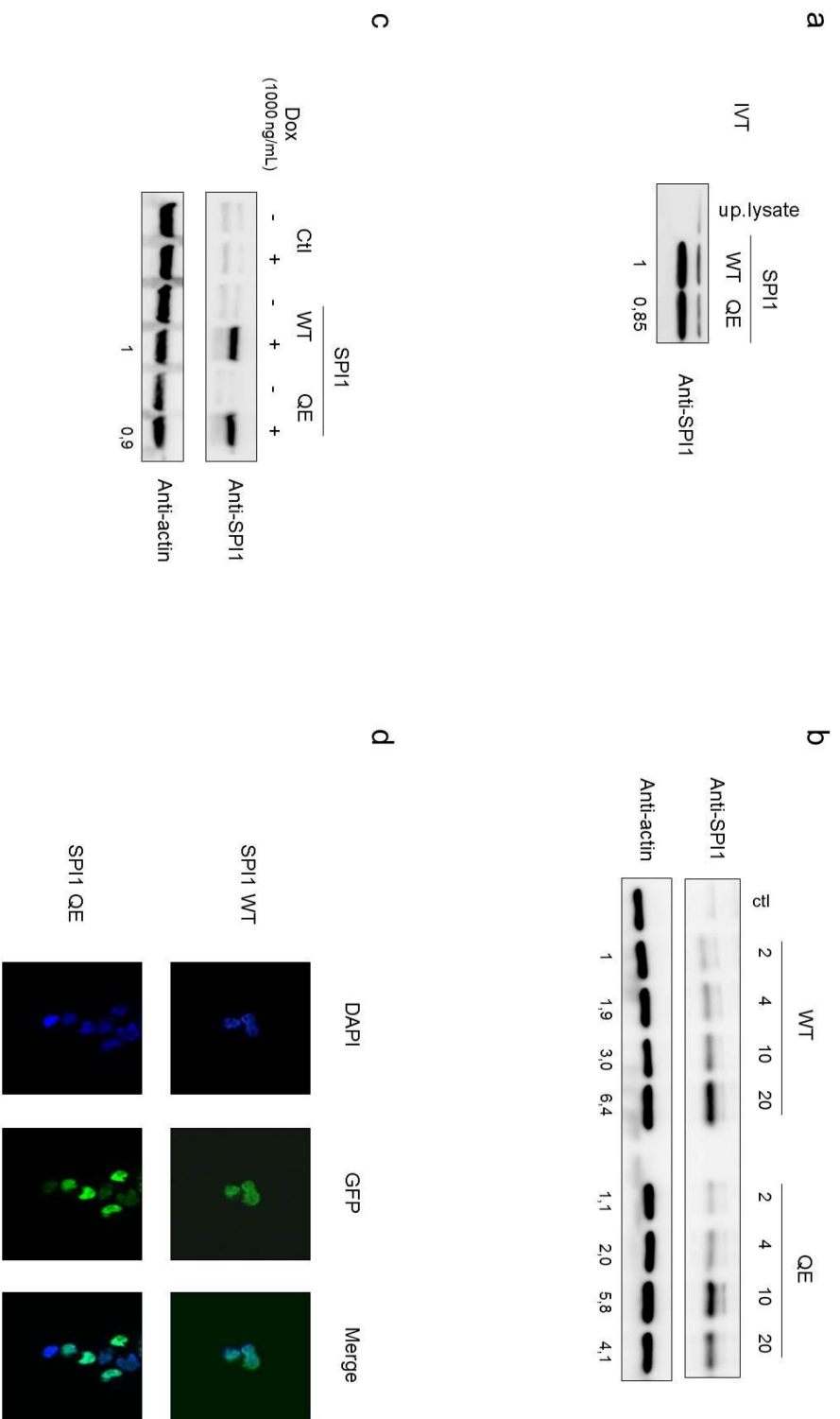


Supplementary Figure S1



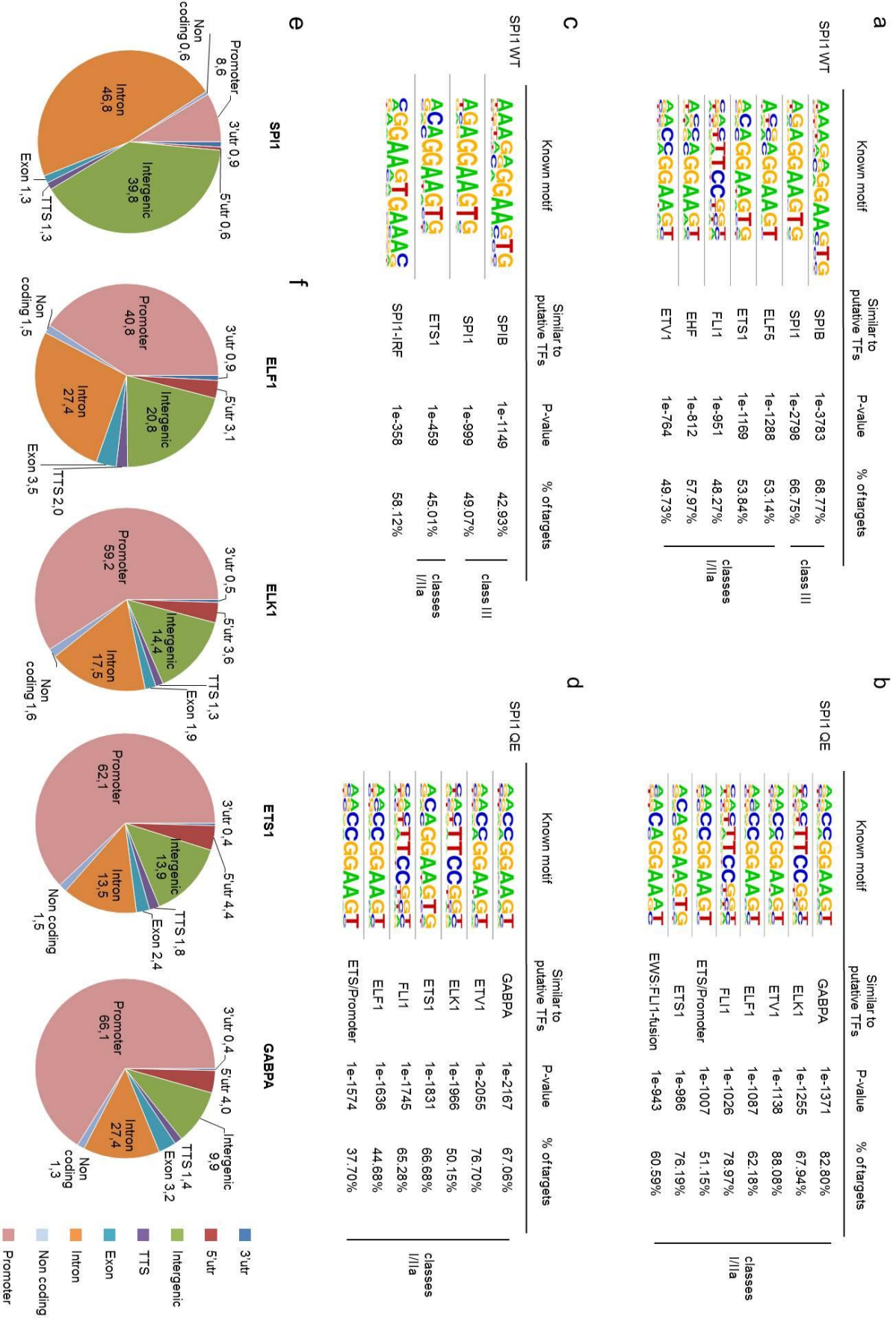
Supplementary Figures





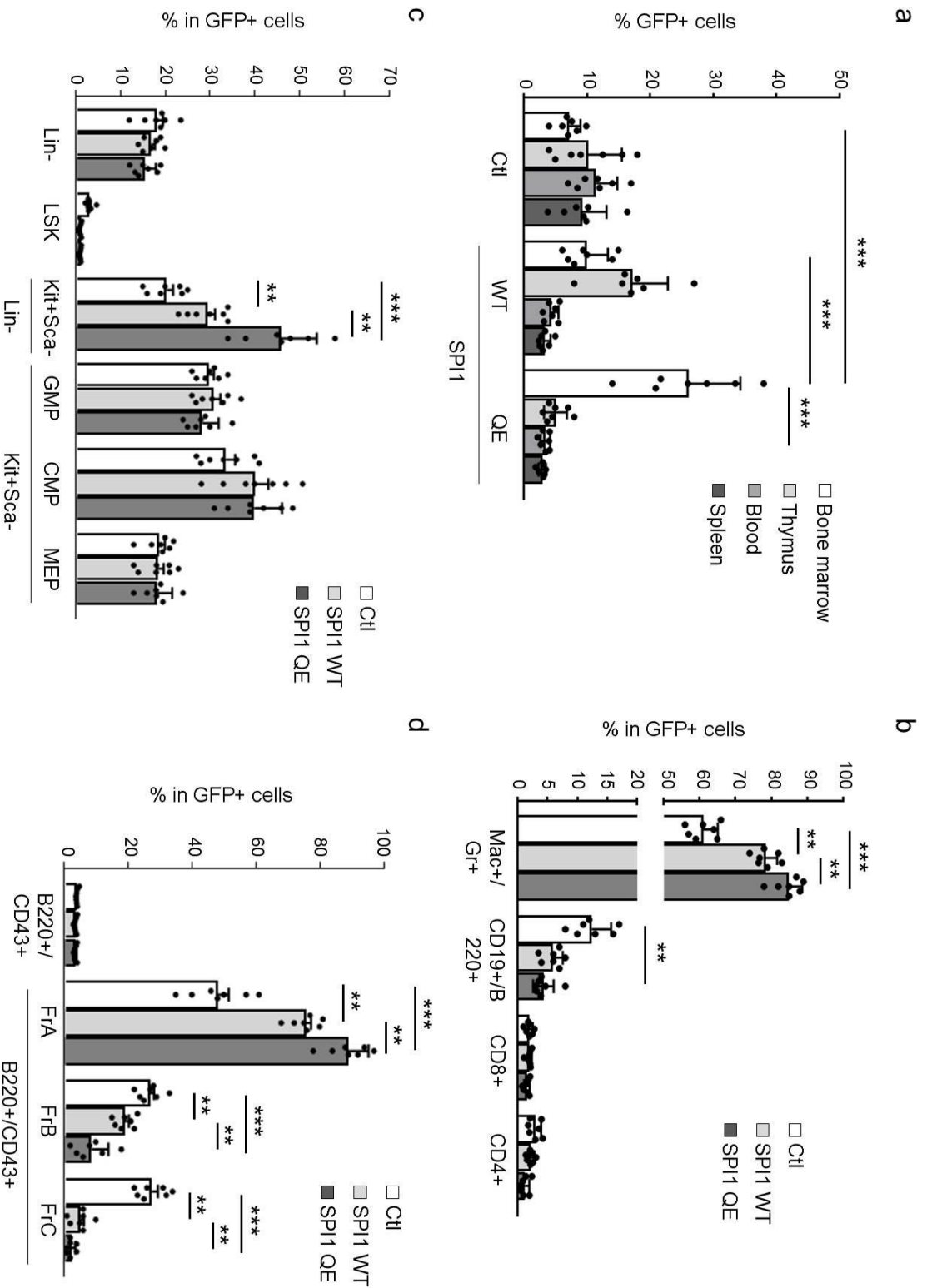


## Supplementary Figure S3

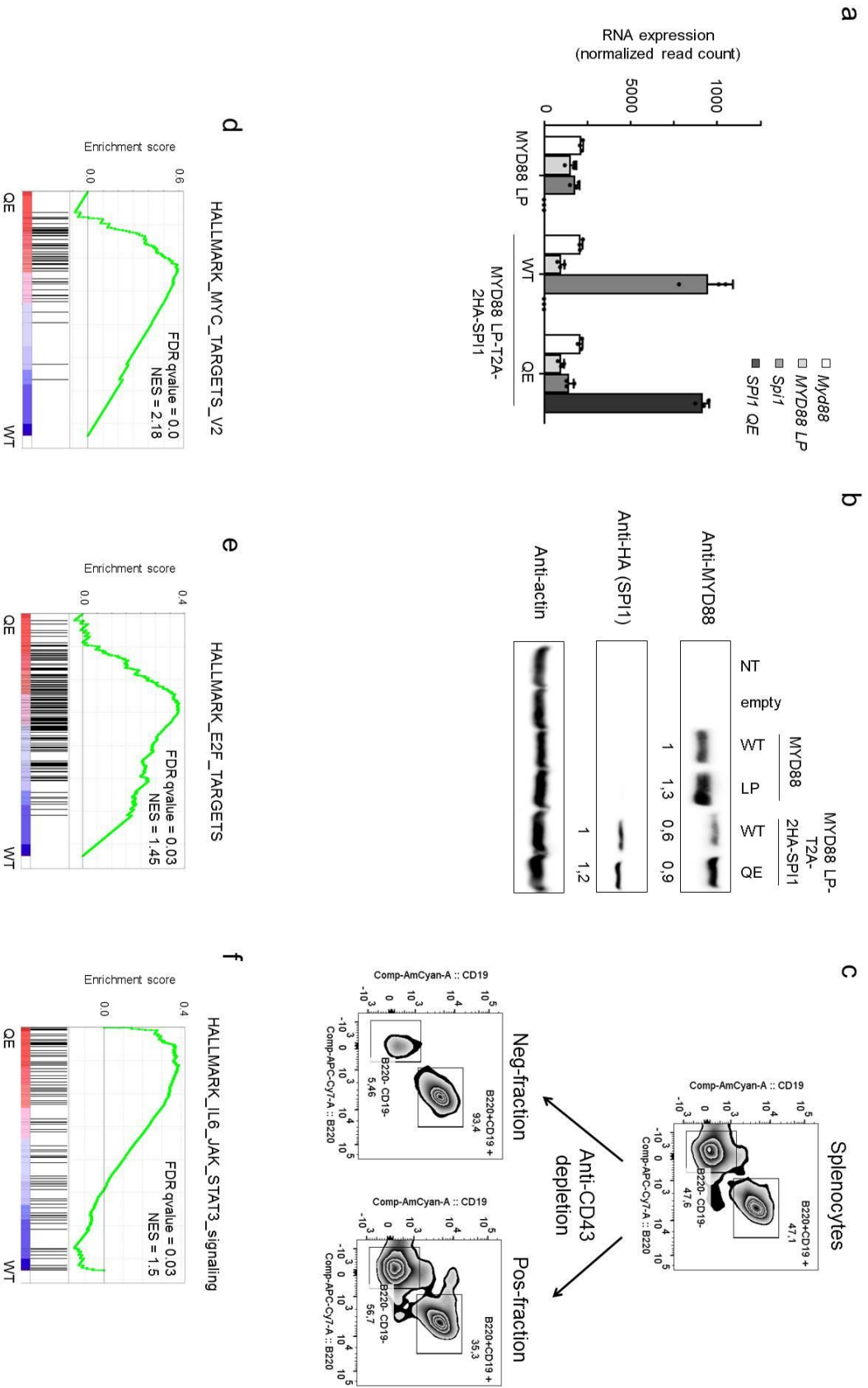




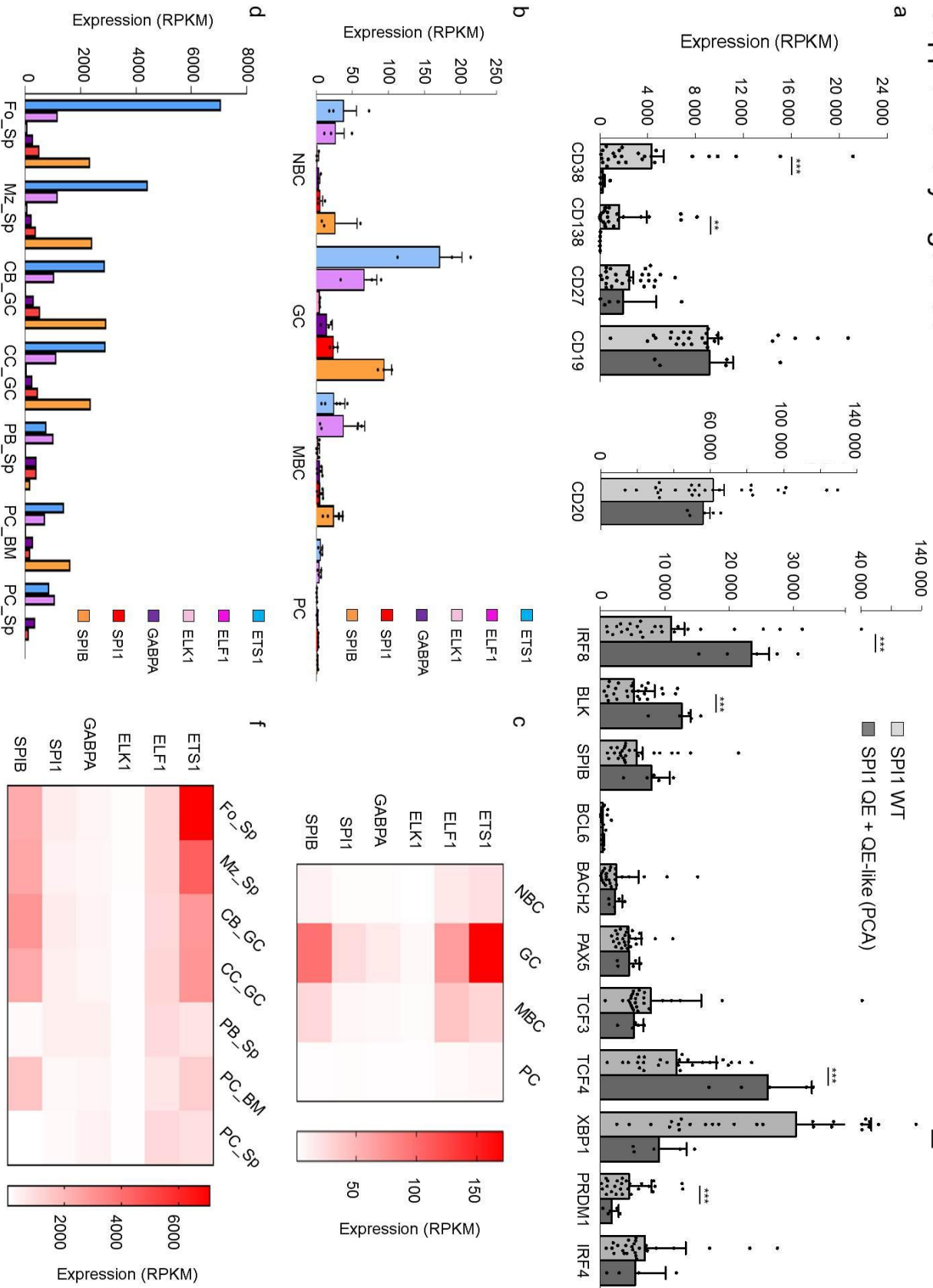
Supplementary Figure S4



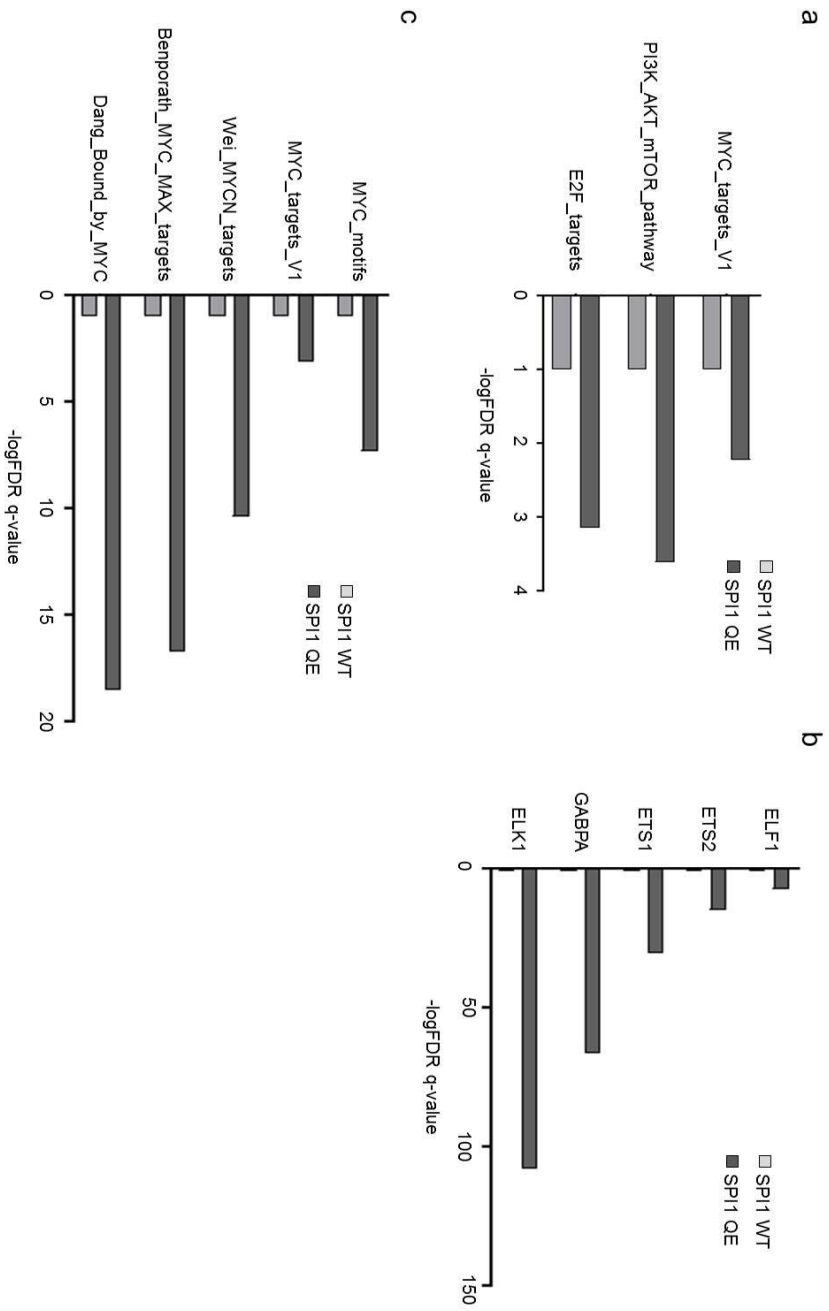
Supplementary Figure S5



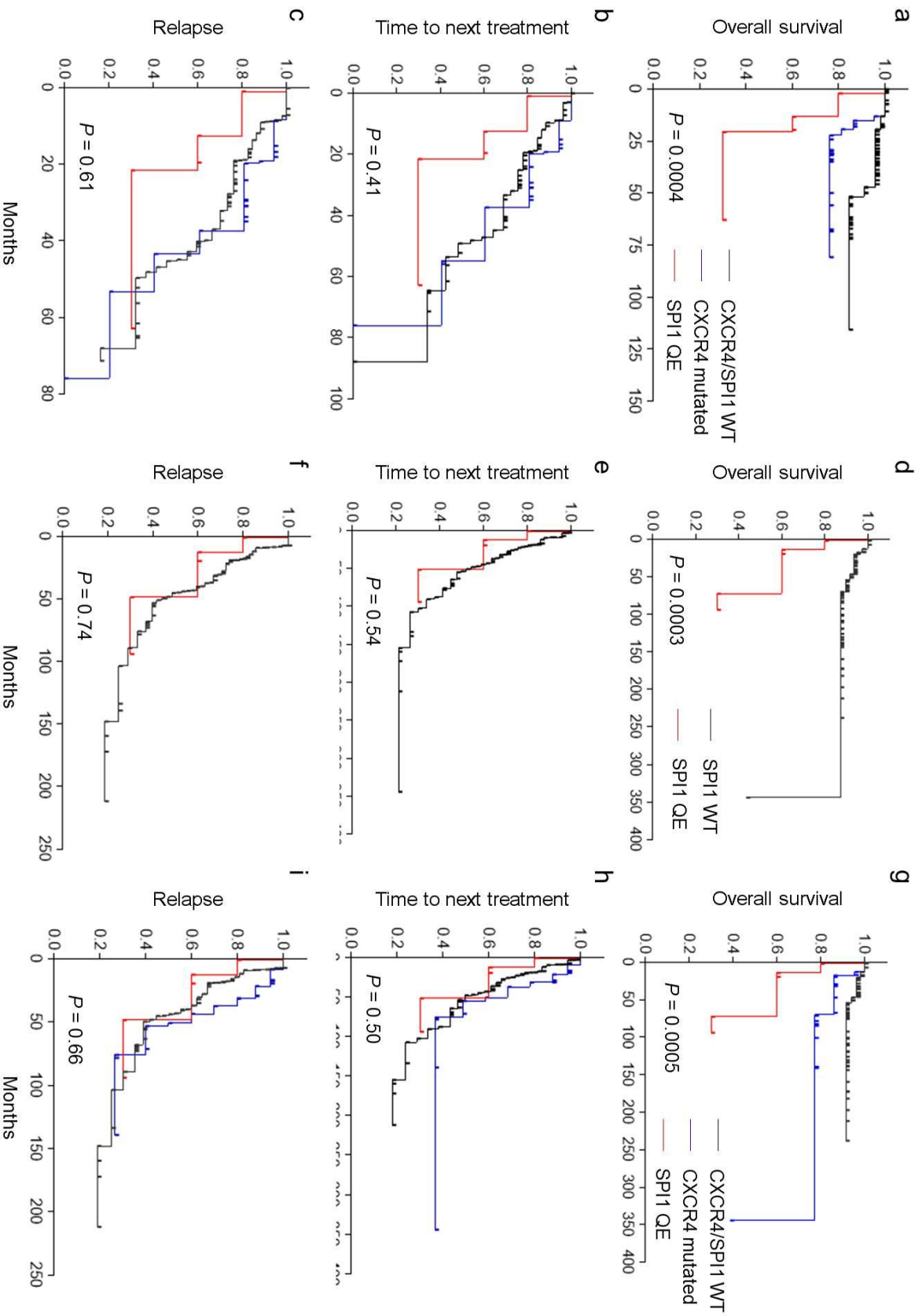
Supplementary Figure S6



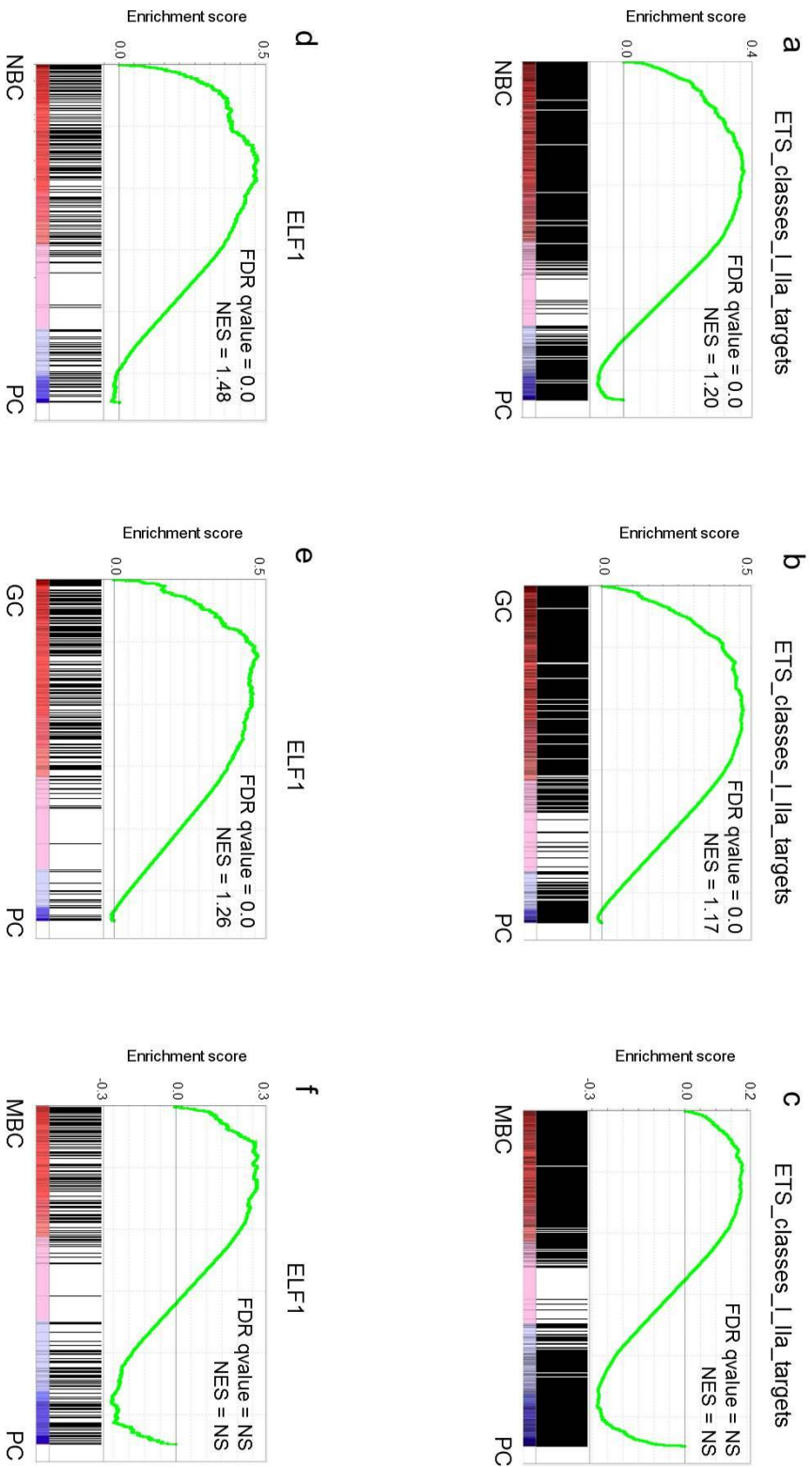
## Supplementary Figure S7



## Supplementary Figure S8



## Supplementary Figure S9



### 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).

### Supplementary Figure S2.

(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) Western blot evaluation of SPI1 and  $\beta$ -actin protein expressions in HEK 293T cells transfected with increasing amounts of the indicated (control, SPI1 WT or QE) expression plasmids. Shown is a representative example of three independent experiments. Relative densitometric values (shown for each band) were normalized against the levels of the internal control  $\beta$ -actin conditions. The condition SPI1 WT (2 ng) was set as 1. Ctl, control.

(c) Western blot evaluation of levels of SPI1 and  $\beta$ -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 for 48h. Relative densitometric values were normalized against the levels of the internal control  $\beta$ -actin conditions and set as 1 for SPI1 WT condition. Ctl, control. Dox, doxycycline.

(d) 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.

### Supplementary Figure S3.

(a-d) ChIP-seq analyses, performed with anti-GFP-Trap (a-b) and anti-SPI1 (c-d) 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 (a,c) and QE proteins (b,d). 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.

(e-f) ChIP-seq analyses and genomic localization, defined by ENCODE annotations, of DNA sequences bound by SPI1 (class III) (e), and ELF1, ELK1, ETS1 and GABPA (classes I and IIa) (f), 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 S4.

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 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 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 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(574). Ctl, control; Fr, fraction.

### Supplementary Figure S5.

(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. 3a**) and sorted on GFP expression.

(b) Expression levels of MYD88, HA-tagged SPI1 and  $\beta$ -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  $\beta$ -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 S6.

(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 B-cell differentiation: naïve 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 (c) and representative heat map (d) of ETS classes I and IIa (ETS1, ELF1, ELK1, GABPA) and class III (SPI1, SPIB) proteins at different steps of 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.(575)

### Supplementary Figure S7.

(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.

### Supplementary Figure S8.

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

### Supplementary Figure S9.

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

## References

1. Hardy, R.R., Carmack, C.E., Shinton, S.A., Kemp, J.D. & Hayakawa, K. Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *J Exp Med* **173**, 1213-25 (1991).
2. Heng, T.S., Painter, M.W. & Immunological Genome Project, C. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol* **9**, 1091-4 (2008).



## Supplemental Tables

Supplementary Table S1. Clinical and biological characteristics of the whole WM population, and of *SPI1* WT and QE subgroups.

	Total	<i>SPI1</i> WT	<i>SPI1</i> QE	<i>P</i> value
	n=85	n=80	n=5	
<b>Median age at WM diagnosis, years (range)</b>	65,6 (35-91)	65,4 (35-88)	74,3 (60-91)	0,14
<b>Median age at NGS evaluation, years (range)</b>	68,6 (41-91)	68,5 (41-88)	76,3 (60-91)	0,24
<b>Sex, male (%)</b>	56 (65,9)	52 (65)	4 (80)	0,45
<b>Median time from WM diagnosis- months (range)</b>	61,6 (0-812)	61,8 (0-812)	2 (0-130)	0,05
<b>Previous therapy for WM</b>				
Yes (%)	17 (20)	16 (20)	1 (20)	0,85
<b>WM IPSS score, n (%)</b>				
Low	13/66 (20)	13/61 (21)	0/5 (0)	0,03
Intermediate	29/66 (44)	27/61 (44)	2/5 (40)	0,84
High	24/66 (36)	21/61 (34)	3/5 (60)	0,03
<b>IgM levels</b>				
Median (range), g/L	21,1 (4,6-84)	20,7 (4,6-84)	32,7 (13,5-35,5)	0,26
> 40 g/L, n (%)	12/77 (16)	12/72 (17)	0/5 (0)	0,21
<b>Symptomatic disease, yes (%)</b>	71/80 (89)	66/75 (88)	5/5 (100)	0,12
<b>Clinical and biological parameters</b>				
Adenopathy, yes (%)	19/77 (25)	18/72 (25)	1/5 (20)	0,66
Splenomegaly, yes (%)	11/77 (14)	9/72 (12,5)	2/5 (40)	0,18
Hemoglobin, mean in g/dL	10,3	10,6	9,2	0,19
Hemoglobin < 11,5 g/dL, yes (%)	53/77 (69)	48/72 (67)	5/5 (100)	-
Platelets, mean in G/L	234	237	176	0,12
Platelets < 100 G/L, yes (%)	13/77 (17)	11/72 (15)	2/5 (40)	-
IgM, mean in g/L	25	24,8	28,6	0,61
Medullar infiltration, median (%)	47,7	46,8	64,3	0,05
<b>Cytogenetic abnormalities</b>				
Complex karyotype (≥ 3 abnormalities)	13/56 (23)	12/51 (23,5)	1/5 (20)	0,72
del6q	27/60 (46)	26/55 (47)	1/5 (20)	0,04
del17p	4/57 (7)	3/52 (6)	1/5 (20)	0,18
del13q	12/60 (20)	11/55 (20)	1/5 (20)	0,85
del11q	7/60 (12)	7/55 (13)	0/5 (0)	0,08
tri12	9/62 (14,5)	7/57 (12)	2/5 (40)	0,24
tri4	8/62 (12,5)	5/57 (8,5)	3/5 (60)	0,02
tri18	4/62 (6,5)	3/57 (5)	1/5 (20)	0,11
<b>Mutations</b>				
<i>MYD88</i>	81/85 (95,5)	76/80 (89,5)	5/5 (100)	0,41
<i>IGLL5</i>	31/85 (36,5)	28/80 (35)	3/5 (60)	0,17
<i>CXCR4</i>	24/85 (28,2)	22/80 (26,5)	2/5 (40)	0,19
<i>ARID1A</i>	7/85 (8,2)	7/80 (8,5)	0/5 (0)	0,12



## DISCUSSION

### Analyse moléculaire d'une cohorte de patients MW

Notre travail a porté sur une cohorte de patients MW dont les principales caractéristiques cliniques, biologiques et cytogénétiques étaient concordantes avec celles rapportées dans la littérature. Elles sont décrites dans le **tableau S1 (article)**. Les médianes d'âge au diagnostic de la MW et de l'analyse NGS étaient respectivement de 65 et 68 ans, deux tiers des patients étaient des hommes. Vingt, 44 et 38% des 66 patients analysables pour cette donnée avaient un score IPSSWM respectivement faible, intermédiaire et élevé. Un cinquième des échantillons analysés l'était après une ou plusieurs lignes de traitement. Une délétion 6q était observée dans 46%, une trisomie 4 dans 12.5%, une délétion 17p dans 7% et un caryotype complexe dans 23% des cas.

Notre analyse NGS de 85 échantillons de WM (16 WES, 69 NGS ciblé sur les régions exoniques de 20 gènes) (**figure 1, tableau S4, figure S1**) (**article**) a confirmé le paysage mutationnel dessiné par les premières analyses NGS publiées (94, 340, 341, 355, 377), décrivant des mutations récurrentes dans les gènes *MYD88*, *CXCR4*, *ARID1A* et *CD79B*. Parmi les 13 gènes mutés chez au moins deux patients dans notre série, 12 (*MYD88*, *IGLL5*, *CXCR4*, *ARID1A*, *CD79B*, *TP53*, *LTB*, *HIST1H1E*, *TRAF2*, *ROBO2*, *TRRAP*, *EZH2*) avaient déjà été identifiés comme mutés dans la MW ou d'autres hémopathies lymphoïdes B. Les mutations de *MYD88*, les plus fréquentes, présentes chez plus de 90% des patients, touchent principalement la lysine en position 265 (L265P) mais d'autres mutations, précédemment décrites dans les DLBCL-ABC (92), peuvent également être observées, comme c'est le cas de trois échantillons dans notre série (M232T, I220T, D208E). Ces mutations se situent toutes sur un feuillet  $\beta$  d'une boucle  $\beta$ - $\beta$  du domaine TIR de *MYD88* et sont prédites pour modifier la structure tertiaire de *MYD88* et conduire à l'activation de la voie NF- $\kappa$ B. Les quatre patients non mutés pour *MYD88* avaient une mutation soit de *CXCR4*, de *CD79B*, d'*EZH2* ou de *CARD11*. Des mutations de *CXCR4* étaient identifiées dans près de 30% des échantillons de notre cohorte. Il s'agissait de mutations hétérozygotes, décalant le cadre de lecture ou à l'origine d'un codon stop, prédites pour aboutir à une perte des sérines régulatrices dans la portion intracytoplasmique de *CXCR4* et laisser intactes les régions responsables de la liaison du SDF-1 $\alpha$  et de la signalisation d'aval via les protéines G. Comme rapporté dans une étude récente (355), la moitié des mutations survenait sur la sérine S338. Les autres gènes mutés de façon récurrente étaient impliqués dans les voies de signalisation d'aval (NF- $\kappa$ B, PI3K/AKT) de récepteurs primordiaux dans la biologie du LB (BCR, CD40, TLR) :



*CD79A/B*, *CARD11*, *TRAF2*, *TRRAP*, *LTB*, *NFKBIE* ; mais touchaient aussi des régulateurs épigénétiques (*ARID1A*, *EZH2*, *CREBBP*, *TRRAP*) ou encore des facteurs de transcription dont le rôle est connu dans la lymphopoïèse B (*MEF2B*, *IKZF3*). Les mutations d'*IGLL5* n'ont pas été décrites jusqu'à présent dans la MW mais elles ont été identifiées dans diverses hémopathies lymphoïdes B : DLBCL (456, 576), CLL (558) et myélome (577). *IGLL5* est un homologue d'*IGLL1* (*lambda5*), qui code pour une protéine de type « surrogate light-chain » composant le pré-BCR. Le profil des mutations d'*IGLL5* suggère qu'il puisse s'agir de mutations issues d'une activité off-target d'AID (558). Le rôle de ces mutations dans les mécanismes de lymphomagénèse n'est pas encore établi.

Le nombre d'échantillons analysés par WES ou WGS reste relativement faible (< 50) dans la MW et il est probable que lorsque ce nombre augmentera d'autres mutations récurrentes dans des gènes d'intérêt ou des régions régulatrices (comme cela a pu être par exemple décrit dans les CLL avec les mutations de régions enhancer de *PAX5* (153)) pourront être découvertes.

Les principales particularités du paysage mutationnel de la MW sont la fréquence très élevée des mutations de *MYD88* (quasi-exclusivement L265P) et la présence de mutations de *CXCR4* (qui ne sont que très rarement rapportées dans les autres hémopathies lymphoïdes B). Le profil mutationnel de la MW se rapproche le plus de celui des DLBCL-ABC (et en particulier des lymphomes extra-nodaux, de type lymphome primitif du système nerveux central) du fait de la fréquence élevée des mutations de *MYD88* et *CD79A/B* et de leur sensibilité aux inhibiteurs de BTK (81, 315). De ce point de vue, il a récemment été démontré que les lignées cellulaires de DLBCL-ABC, mutées pour *MYD88* et *CD79A/B*, étaient dépendantes d'une signalisation oncogénique contrôlée par un complexe multiprotéique comprenant BCR, TLR9 et *MYD88*, co-localisant avec mTOR dans les endolysosomes et conduisant à l'activation des voies NF- $\kappa$ B et mTOR (578). La présence de ce complexe multiprotéique était également observée chez les patients MW/LPL, suggérant que la MW puisse être également dépendante de cette signalisation oncogénique particulière.

De nombreuses questions restent cependant non élucidées concernant les mécanismes d'oncogénèse de la MW. Quelles sont les voies de signalisations indispensables à la survie et la prolifération des cellules tumorales de MW ? Quelle est la contrepartie physiologique et la cellule d'origine de la MW ? Quels sont les événements expliquant le blocage de différenciation vers le plasmocyte dans la MW ? Quelles sont les anomalies génétiques et épigénétiques concourant à la transformation tumorale ? Une partie des



réponses à ces questions pourrait être apportée par une analyse large, réalisée sur une grande série de patients et intégrant plusieurs approches (WES/WGS, RNAseq, méthylation), comprenant notamment l'étude de populations lymphocytaires B à différents stades de maturation ainsi que différentes hémopathies lymphoïdes B afin de mieux appréhender les éventuelles spécificités transcriptomiques, épigénétiques de la MW et le rôle du profil des mutations les plus récurrentes dans les mécanismes de transformation tumorale.

### **Description d'une mutation récurrente dans le gène *SPI1***

Outre la confirmation de mutations récurrentes dans les gènes *MYD88*, *CXCR4*, *ARID1A* et *CD79B*, nous avons identifié une mutation somatique faux-sens récurrente (c.676C>G, p.Q226E), non décrite jusqu'à présent, dans le domaine de liaison à l'ADN de *SPI1*, un facteur de transcription du sous-groupe de classe III de la famille ETS. Elle a été retrouvée chez cinq patients (5.9%) et était observée à la fois sur les séquences ADN et ARN d'échantillons tumoraux triés, à des VAF variées (5-70%). Elles étaient toujours associées à une mutation de *MYD88* L265P, avec une VAF équivalente ou inférieure, suggérant que l'acquisition de la mutation *SPI1* Q226E (QE) ne précède pas celle de *MYD88*. Sous réserve de leur nombre limité, les cas mutés pour *SPI1* semblaient associés à une maladie plus agressive (60% de score IPSSWM élevé, infiltration médullaire plus importante) et à un pronostic péjoratif (**figure 4, tableau S1, figure S8**) (**article**). Il s'agit de mutations touchant dans tous les cas le même acide aminé (Q226) et consistant en la même substitution nucléotidique (C>G). Ces éléments étaient fortement évocateurs d'une mutation gain de fonction. Aucune mutation dans le gène *SPI1* n'avait été décrite jusqu'à présent dans les hémopathies lymphoïdes B alors que seules de rares mutations ont été identifiées dans les hémopathies en général. Des mutations inactivatrices de *SPI1* ont été décrites dans le sous-groupe de LAM avec réarrangement de *MLL* (3/31) (544). Deux de ces mutations se situaient dans le domaine de liaison à l'ADN (R231C, A232T) et la dernière était prédite pour aboutir à la perte du domaine ETS (P99fs). Elles étaient associées à une diminution du niveau d'expression de gènes cibles connus de *SPI1* (*CSF1R*, *CSF3R* et *AIF1*). Deux gènes de fusions impliquant *SPI1* (*STMN1-SPI1* et *TCF7-SPI1*) ont été identifiés dans les LAL-T pédiatriques de haut risque (453). Ces protéines de fusion conservaient la capacité à lier l'ADN et une activité transcriptionnelle. Leurs expressions dans des progéniteurs murins augmentaient la prolifération cellulaire, suggérant dans ce cas un gain de fonction (453).





## Conséquences fonctionnelles du mutant SPI1

La glutamine (Q) en position 226 est présente chez les protéines ETS de classe III (SPI1, SPIB, SPIC) et responsable de la liaison à une guanine (G) en position -2 du motif commun GGA(AT) reconnu par toutes les protéines ETS, tandis que l'acide glutamique (E), présent à cette position chez les protéines de classes I et IIa (**figure 1**) (**article**), se lie à une cytosine (C) en position -2 (39). Il est à noter que la mutagenèse du Q226 en E a facilité la cristallisation du domaine ETS du SPI1, suggérant que cette modification spécifique pourrait améliorer la liaison de SPI1 à l'ADN (44). Nous avons montré que la protéine mutante SPI1 QE conservait sa capacité de liaison à l'ADN, avait une activité transcriptionnelle 20% plus importante que la forme sauvage de SPI1, et augmentait la prolifération dans différents contextes cellulaires, en accord avec l'hypothèse d'une mutation gain de fonction.

Les analyses de ChIP-seq, réalisées dans une lignée de DLBCL *MYD88* L265P (OCI-Ly10) ont révélé que la mutation QE modifiait la spécificité de liaison à l'ADN de la protéine en matière de séquence et de localisation sur le génome. Le motif ADN le plus fréquemment lié par la forme sauvage de SPI1 correspondait au motif classique reconnu par les ETS de classe III, GAGGAA, alors que celui lié par la forme mutée était un motif CAGGAA rappelant les motifs de liaison des protéines de classes I et IIa (**figure 2i**) (**article**). Cela est en accord avec le rôle évoqué ci-dessus de l'acide aminé (Q ou E) en position 226 dans la spécificité de reconnaissance du nucléotide en position -2 du motif commun. Il faut noter que nos expériences de bandshift et de transactivation ont été réalisées avec une séquence cible de liaison de SPI1 comprenant à la fois des motifs GAGGAA et CAGGAA, soit un motif de classe III et un de classe I/IIa. De nouvelles expériences de transactivation et de titration calorimétrique isotherme, utilisant des séquences ADN cibles comprenant soit uniquement des motifs de type classe III soit uniquement de classes I/IIa, en présence de la forme sauvage ou mutée de SPI1, pourraient être réalisées afin de confirmer ces données de ChIP-seq et évaluer de façon précise l'affinité de liaison protéique selon les types de séquences ADN. Il existait également des différences dans la localisation génomique des régions ADN liées par la forme sauvage et mutée de SPI1. Les régions génomiques liées par le mutant de SPI1 étaient enrichies en régions promotrices actives et enhancers forts, rappelant là aussi le modèle de liaison à la chromatine observé avec les protéines de classe I/IIa (sites de liaison plus fréquemment situés dans les promoteurs (actifs) (40-60%)), en comparaison à celui des protéines de classe III (sites situés de préférence dans les enhancers et seulement 5-15% dans les promoteurs) (45,46). Ces modèles de liaison des protéines ETS proviennent de données de ChIP-seq générées sur des lignées de cellules B transformées par



l'EBV (**supplemental figure S3e**) (**article**) et de cellules primaires murines. Il serait intéressant de les confirmer sur la lignée humaine OCI-Ly10 et sur des échantillons primaires humains de MW.

Une différence importante entre SPI1 et les facteurs ETS de classes I/IIa est que les activités de liaison à l'ADN et de transcription de ces derniers sont étroitement régulées par les voies de signalisation intracellulaire, telles que celle de la MAPK et celle dépendante du calcium (47,48), alors que l'activité de SPI1 ne l'est pas. Ces éléments suggéreraient que les promoteurs liés par le mutant de SPI1 puissent ainsi échapper à une régulation normale. De plus, comme SPI1 interagit avec certains régulateurs épigénétiques, comme CBP (49) et TET2 (50,51), SPI1 QE pourrait aussi recruter ces régulateurs épigénétiques des régions enhancer vers les régions promotrices et ainsi consolider, par des changements épigénétiques, les modifications d'expression génique.

L'expression du mutant de SPI1 a été associée dans notre travail à un enrichissement en des signatures transcriptionnelles liées au cycle cellulaire, à la prolifération et à plusieurs voies de signalisation intracellulaire impliquées dans les mécanismes physiologiques et pathologiques de la biologie du LB (BCR, CD40, TLR, MYC, PI3K). Cet enrichissement a été observé à la fois dans la lignée OCI-Ly10 et dans les échantillons primaires de patients MW (**figure 4**) (**article**). Lorsque les signatures transcriptionnelles de la base de données Lymphochip (29) ont été utilisées, nous avons également noté des différences dans les gènes cibles de FTs qui sont des régulateurs clés de la différenciation lymphoïde B, notamment XBP1, PRDM1 et IRF4 (**figure 2f**) (**article**), suggérant que le mutant de SPI1 puisse perturber la différenciation lymphoïde B terminale. Divers membres de la famille ETS régulent la différenciation et la fonction des LB. Le rôle de six protéines Ets (Ets1 (4,46), Fli1 (60), Gabpa (61), et les trois facteurs de classe III Spi1, SpiB et SpiC (62)) a été étudié en détail dans les LB murins. L'étude de leurs expressions au cours de la maturation des LB humains et murins suggère que les facteurs ETS/Ets de classe I/IIa sont plus actifs dans les LB du CG, avec le niveau d'expression le plus élevé pour ETS1/Ets1 (**supplemental figure S6b-e**) (**article**). La perte d'Ets1 entraîne une différenciation prématurée des LB en ASC (63) et des études récentes ont souligné que Spi1 et SpiB agissent physiologiquement comme des régulateurs négatifs du développement plasmocytaire mais également comme des transmetteurs des stimuli environnementaux, en régulant l'expression de nombreux composants des voies de signalisation en aval des récepteurs BCR, TLR, CD40 et BAFF (13-15). Considérant que la MW est caractérisée par un blocage de différenciation à la transition entre un lymphocyte post-GC et un plasmocyte, nous avons étudié et confirmé l'impact des mutations de *SPI1* sur la différenciation lymphoïde B terminale. En effet, l'expression de SPI1 dans



des LB murins entraînait une différenciation moins efficace en plasmocytes dans un modèle *in vitro*, conformément aux récentes études identifiant Spi1 comme le gardien de cette étape terminale de maturation. De plus, nos analyses d'expression génique ont révélé que les cas de MW mutés pour SPI1 avaient un PEG plus proche de populations lymphocytaires B naïves, du CG ou mémoires que de plasmocytes et exprimaient moins de marqueurs (CD38, CD138, XBP1, PRDM1) liés à la différenciation plasmocytaire que les échantillons non mutés.

Des analyses supplémentaires seront nécessaires pour caractériser de manière exhaustive la relation entre les différentes classes de protéines ETS et les étapes de la différenciation lymphoïde B ainsi que leur rôle éventuel dans le blocage de différenciation observé dans la MW. A cet égard, il convient de noter que les analyses GSEA effectuées sur les séquences ARN des populations de LB normaux montraient un enrichissement significatif des motifs des classes I/IIa dans les populations de LB naïfs, du CG et mémoires par rapport aux plasmocytes (**supplemental figure S9**) (article). Concernant le rôle du mutant de SPI1 sur la différenciation lymphoïde B, il pourrait être intéressant d'évaluer quels sont les mécanismes impliqués dans la dérégulation de l'expression des composants des voies BCR, TLR, CD40 et des FTs clés de la différenciation. S'agit-il de mécanismes de régulation directe ou indirecte, par exemple par des modifications épigénétiques ou des complexes qu'il forme avec ses cofacteurs IRF4 et IRF8 ? Influence-t-il le déroulement du cycle cellulaire ce qui pourrait perturber la différenciation lymphoïde B comme cela a été suggéré dans les cellules souches (579) ?

Les signatures transcriptionnelles comprenant les cibles de MYC et d'IRF4 étaient parmi les signatures les plus enrichies dans les cellules OCI-Ly10 exprimant le mutant de SPI1 et/ou dans les échantillons MW mutés SPI1 QE. Nous avons confirmé l'importance de ces voies dans la prolifération des cellules OCI-Ly10 exprimant SPI1 QE, plus sensibles que les cellules exprimant la forme sauvage, aux traitements par inhibiteur de BET (JQ1) et par lénalidomide, utilisés dans notre travail afin d'inhiber respectivement la voie MYC et IRF4. Cependant, ces inhibiteurs ne sont pas spécifiques et des expériences complémentaires utilisant des approches ciblées de réduction de l'expression de MYC et IRF4, par exemple par shRNA ou CRISPR, pourraient être réalisées afin de valider ces observations.

Enfin, de façon générale, nos résultats pourraient ou devront être validés dans d'autres modèles cellulaires *in vitro* et *in vivo* : lignées cellulaires de MW (BCWM.1, BCWM.2, MWCL-1 et/ou RPCI-WM1), modèle murin knock-in SPI1 Q226E, échantillons primaires de MW mutés ou non pour SPI1.



## Conclusion et perspectives

Les FTs ETS sont divisés en sous-familles en fonction de leurs similitudes en matière de séquence protéique, de séquences de liaison à l'ADN et d'interactions avec différents cofacteurs. Ils sont régulés par des signaux extracellulaires et contribuent à divers processus cellulaires, dont la prolifération cellulaire et la transformation tumorale. Les gènes de la famille *ETS* sont fréquemment ciblés par des processus oncogéniques que ce soit des translocations chromosomiques ou des gains du nombre de leurs copies. Les propriétés oncogéniques des protéines ETS ont été largement décrites dans les tumeurs solides. Les réarrangements des loci des *ETS* sont une caractéristique du cancer de la prostate, environ la moitié des tumeurs présentant des altérations à un locus d'un gène codant pour un membre de la famille ETS (55,56). Le gène *PU.1/SPI1* est également ciblé par des mutations ponctuelles inactivatrices dans les hémopathies myéloïdes humaines. Cependant, des anomalies des protéines ETS ont rarement été identifiées dans les hémopathies lymphoïdes B, à l'exception des mutations d'*ELF4* dans la CLL (52) et des translocations impliquant *SPIB*, des gains 11q24, régions comprenant les gènes *ETS1* et *FLI1* et enfin plus récemment des mutations d'*ETS1* dans les DLBCL (57-59) (154).

Nous avons identifié, chez environ 6% des patients atteints de MW, une nouvelle mutation somatique récurrente (c.676C>G, p.Q226E) dans le domaine de liaison à l'ADN de SPI1, un FT du sous-groupe de classe III de la famille ETS. La mutation 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 enhancer 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. Les conséquences fonctionnelles de cette mutation sont une augmentation de la prolifération cellulaire et une diminution de la différenciation lymphoïde B terminale dans une lignée cellulaire modèle et des échantillons primaires de MW.

Des mutations ponctuelles dans des gènes codant pour des FTs, comme celles par exemple d'*ELF4*, *EGR2*, *MEF2C* ou *IRF4* identifiées dans les hémopathies lymphoïdes B (26,52,53), pourraient affecter la liaison à l'ADN de la protéine et leurs motifs de reconnaissance. Avec la description récente des mutations du gène *KLF5* (54) qui modifient les propriétés de liaison à l'ADN, nos travaux fournissent un cadre pour comprendre les conséquences fonctionnelles de ces mutations.



Nous décrivons ici un mécanisme de subversion oncogénique de la fonction d'un FT suite à la modification subtile de la spécificité de liaison à l'ADN de la protéine mutante, aboutissant à des conséquences fonctionnelles participant à la transformation tumorale. La démonstration qu'une mutation somatique ponctuelle peut modifier l'équilibre de liaison d'un FT à l'échelle du génome fournit un paradigme mécanistique sur la façon dont les mutations faux sens dans les gènes codant pour des FTs pourraient être oncogéniques dans les tumeurs humaines. Il reste ainsi à démontrer si d'autres mutations erronées dans les gènes de la famille *ETS* (*ETS1*, *ELK1*) (37) ou des FTs (*EGR1*, *EGR2*, *MEF2C*) (26,53,64), identifiés dans d'autres syndromes lymphoprolifératifs B, peuvent avoir le même impact sur les motifs de reconnaissance de l'ADN, la différenciation lymphoïde B et/ou les mécanismes de transformation tumorale.

Enfin, la pertinence biologique du mutant Q226E de SPI1 est appuyé par l'association à un mauvais pronostic, même si cet impact sur la survie nécessite bien sûr d'être validé sur de plus grandes séries de patients. L'enrichissement des voies MYC et IRF4 associé aux cas exprimant le mutant de SPI1 est de ce point de vue intéressant, puisque le ciblage de ces voies par un inhibiteur de BET et le lénalidomide pourrait avoir valeur thérapeutique chez ces patients de mauvais pronostic. Nos résultats identifient la mutation *SPI1* Q226E comme un éventuel nouveau biomarqueur génétique associé à un pronostic défavorable et indiquent de nouvelles options thérapeutiques pour ce groupe de patients.



## REFERENCES

1. Laurenti E, Gottgens B. From haematopoietic stem cells to complex differentiation landscapes. *Nature*. 2018;553(7689):418-26.
2. Vedi A, Santoro A, Dunant CF, Dick JE, Laurenti E. Molecular landscapes of human hematopoietic stem cells in health and leukemia. *Ann N Y Acad Sci*. 2016;1370(1):5-14.
3. Crane GM, Jeffery E, Morrison SJ. Adult haematopoietic stem cell niches. *Nat Rev Immunol*. 2017;17(9):573-90.
4. Weissman IL, Anderson DJ, Gage F. Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations. *Annu Rev Cell Dev Biol*. 2001;17:387-403.
5. Adolfsson J, Mansson R, Buza-Vidas N, Hultquist A, Liuba K, Jensen CT, et al. Identification of Flt3+ lymphomyeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell*. 2005;121(2):295-306.
6. Alhaj Hussien K, Vu Manh TP, Guimiot F, Nelson E, Chabaane E, Delord M, et al. Molecular and Functional Characterization of Lymphoid Progenitor Subsets Reveals a Bipartite Architecture of Human Lymphopoiesis. *Immunity*. 2017;47(4):680-96 e8.
7. Schlissel MS. Regulating antigen-receptor gene assembly. *Nat Rev Immunol*. 2003;3(11):890-9.
8. Weisel F, Shlomchik M. Memory B Cells of Mice and Humans. *Annu Rev Immunol*. 2017;35:255-84.
9. Martin VG, Wu YB, Townsend CL, Lu GH, O'Hare JS, Mozeika A, et al. Transitional B Cells in Early Human B Cell Development - Time to Revisit the Paradigm? *Front Immunol*. 2016;7:546.
10. LeBien TW, Tedder TF. B lymphocytes: how they develop and function. *Blood*. 2008;112(5):1570-80.
11. Wei Q, Frenette PS. Niches for Hematopoietic Stem Cells and Their Progeny. *Immunity*. 2018;48(4):632-48.
12. 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;14(2):69-80.
13. 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;282(19):14194-204.
14. Papanthasiou P, Attema JL, Karsunky H, Hosen N, Sontani Y, Hoyne GF, et al. Self-renewal of the long-term reconstituting subset of hematopoietic stem cells is regulated by Ikaros. *Stem Cells*. 2009;27(12):3082-92.
15. Kikuchi K, Lai AY, Hsu CL, 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;201(8):1197-203.
16. Maeda T, Merghoub T, Hobbs RM, Dong L, Maeda M, Zakrzewski J, et al. Regulation of B versus T lymphoid lineage fate decision by the proto-oncogene LRF. *Science*. 2007;316(5826):860-6.
17. 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;6(8):836-43.
18. Fuxa M, Skok J, Souabni A, Salvaggio G, Roldan E, Busslinger M. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev*. 2004;18(4):411-22.
19. Lassoued K, Nunez CA, Billips L, Kubagawa H, Monteiro RC, LeBlen TW, et al. Expression of surrogate light chain receptors is restricted to a late stage in pre-B cell differentiation. *Cell*. 1993;73(1):73-86.
20. Wardemann H, Yurasov S, Schaefer A, Young JW, Meffre E, Nussenzweig MC. Predominant autoantibody production by early human B cell precursors. *Science*. 2003;301(5638):1374-7.
21. Nemazee D, Buerki K. Clonal deletion of autoreactive B lymphocytes in bone marrow chimeras. *Proc Natl Acad Sci U S A*. 1989;86(20):8039-43.
22. Goodnow CC, Crosbie J, Adelstein S, Lavoie TB, Smith-Gill SJ, Brink RA, et al. Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature*. 1988;334(6184):676-82.
23. Gay D, Saunders T, Camper S, Weigert M. Receptor editing: an approach by autoreactive B cells to escape tolerance. *J Exp Med*. 1993;177(4):999-1008.
24. Sabouri Z, Perotti S, Spierings E, Humburg P, Yabas M, Bergmann H, et al. IgD attenuates the IgM-induced energy response in transitional and mature B cells. *Nat Commun*. 2016;7:13381.
25. Sims GP, Ettinger R, Shirota Y, Yarboro CH, Illei GG, Lipsky PE. Identification and characterization of circulating human transitional B cells. *Blood*. 2005;105(11):4390-8.
26. Cerutti A, Cols M, Puga I. Marginal zone B cells: virtues of innate-like antibody-producing lymphocytes. *Nat Rev Immunol*. 2013;13(2):118-32.
27. Seifert M, Kuppers R. Human memory B cells. *Leukemia*. 2016;30(12):2283-92.
28. Weill JC, Weller S, Reynaud CA. Human marginal zone B cells. *Annu Rev Immunol*. 2009;27:267-85.
29. Bagnara D, Squillario M, Kipling D, Mora T, Walczak AM, Da Silva L, et al. A Reassessment of IgM Memory Subsets in Humans. *J Immunol*. 2015;195(8):3716-24.
30. De Silva NS, Klein U. Dynamics of B cells in germinal centres. *Nat Rev Immunol*. 2015;15(3):137-48.
31. Nutt SL, Hodgkin PD, Tarlinton DM, Corcoran LM. The generation of antibody-secreting plasma cells. *Nat Rev Immunol*. 2015;15(3):160-71.
32. Mesin L, Ersching J, Victora GD. Germinal Center B Cell Dynamics. *Immunity*. 2016;45(3):471-82.
33. Batista FD, Harwood NE. The who, how and where of antigen presentation to B cells. *Nat Rev Immunol*. 2009;9(1):15-27.
34. Goodnow CC, Vinuesa CG, Randall KL, Mackay F, Brink R. Control systems and decision making for antibody production. *Nat Immunol*. 2010;11(8):681-8.
35. Cyster JG. B cell follicles and antigen encounters of the third kind. *Nat Immunol*. 2010;11(11):989-96.





36. Paus D, Phan TG, Chan TD, Gardam S, Basten A, Brink R. Antigen recognition strength regulates the choice between extrafollicular plasma cell and germinal center B cell differentiation. *J Exp Med*. 2006;203(4):1081-91.
37. Chan TD, Gardam S, Gatto D, Turner VM, Silke J, Brink R. In vivo control of B-cell survival and antigen-specific B-cell responses. *Immunol Rev*. 2010;237(1):90-103.
38. Schwickert TA, Victora GD, Fooksman DR, Kamphorst AO, Mugnier MR, Gitlin AD, et al. A dynamic T cell-limited checkpoint regulates affinity-dependent B cell entry into the germinal center. *J Exp Med*. 2011;208(6):1243-52.
39. Kerfoot SM, Yaari G, Patel JR, Johnson KL, Gonzalez DG, Kleinstein SH, et al. Germinal center B cell and T follicular helper cell development initiates in the interfollicular zone. *Immunity*. 2011;34(6):947-60.
40. Kitano M, Moriyama S, Ando Y, Hikida M, Mori Y, Kurosaki T, et al. Bcl6 protein expression shapes pre-germinal center B cell dynamics and follicular helper T cell heterogeneity. *Immunity*. 2011;34(6):961-72.
41. Victora GD, Dominguez-Sola D, Holmes AB, Deroubaix S, Dalla-Favera R, Nussenzweig MC. Identification of human germinal center light and dark zone cells and their relationship to human B-cell lymphomas. *Blood*. 2012;120(11):2240-8.
42. Allen CD, 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;5(9):943-52.
43. Caron G, Le Gallou S, Lamy T, Tarte K, Fest T. CXCR4 expression functionally discriminates centroblasts versus centrocytes within human germinal center B cells. *J Immunol*. 2009;182(12):7595-602.
44. Vinuesa CG, Linterman MA, Goodnow CC, Randall KL. T cells and follicular dendritic cells in germinal center B-cell formation and selection. *Immunol Rev*. 2010;237(1):72-89.
45. Xu Z, Zan H, Pone EJ, Mai T, Casali P. Immunoglobulin class-switch DNA recombination: induction, targeting and beyond. *Nat Rev Immunol*. 2012;12(7):517-31.
46. Kurosaki T, Kometani K, Ise W. Memory B cells. *Nat Rev Immunol*. 2015;15(3):149-59.
47. Shapiro-Shelef M, Calame K. Regulation of plasma-cell development. *Nat Rev Immunol*. 2005;5(3):230-42.
48. Fairfax KA, Kallies A, Nutt SL, Tarlinton DM. Plasma cell development: from B-cell subsets to long-term survival niches. *Semin Immunol*. 2008;20(1):49-58.
49. MacLennan IC, Toellner KM, Cunningham AF, Serre K, Sze DM, Zuniga E, et al. Extrafollicular antibody responses. *Immunol Rev*. 2003;194:8-18.
50. Shlomchik MJ, Weisel F. Germinal center selection and the development of memory B and plasma cells. *Immunol Rev*. 2012;247(1):52-63.
51. Nutt SL, Tarlinton DM. Germinal center B and follicular helper T cells: siblings, cousins or just good friends? *Nat Immunol*. 2011;12(6):472-7.
52. Radbruch A, Muehlinghaus G, Luger EO, Inamine A, Smith KG, Dorner T, et al. Competence and competition: the challenge of becoming a long-lived plasma cell. *Nat Rev Immunol*. 2006;6(10):741-50.
53. Kometani K, Nakagawa R, Shinnakasu R, Kaji T, Rybouchkin A, Moriyama S, et al. Repression of the transcription factor Bach2 contributes to predisposition of IgG1 memory B cells toward plasma cell differentiation. *Immunity*. 2013;39(1):136-47.
54. Nutt SL, Taubenheim N, Hasbold J, Corcoran LM, Hodgkin PD. The genetic network controlling plasma cell differentiation. *Semin Immunol*. 2011;23(5):341-9.
55. 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;127(20):2375-90.
56. 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;403(6769):503-11.
57. Rosenwald A, Alizadeh AA, Widhopf G, Simon R, Davis RE, Yu X, et al. Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *J Exp Med*. 2001;194(11):1639-47.
58. Klein U, Tu Y, Stolovitzky GA, Mattioli M, Cattoretti G, Husson H, et al. Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J Exp Med*. 2001;194(11):1625-38.
59. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A*. 2003;100(17):9991-6.
60. Dave SS, Fu K, Wright GW, Lam LT, Kluijn P, Boerma EJ, et al. Molecular diagnosis of Burkitt's lymphoma. *N Engl J Med*. 2006;354(23):2431-42.
61. Isaacson PG, Du MQ. MALT lymphoma: from morphology to molecules. *Nat Rev Cancer*. 2004;4(8):644-53.
62. Basso K, Liso A, Tiacci E, Benedetti R, Pulsoni A, Foa R, et al. Gene expression profiling of hairy cell leukemia reveals a phenotype related to memory B cells with altered expression of chemokine and adhesion receptors. *J Exp Med*. 2004;199(1):59-68.
63. Isaacson PG, Norton AJ, Addis BJ. The human thymus contains a novel population of B lymphocytes. *Lancet*. 1987;2(8574):1488-91.
64. 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;104(10):3318-25.
65. Ramiro AR, Jankovic M, Eisenreich T, Difilippantonio S, Chen-Kiang S, Muramatsu M, et al. AID is required for c-myc/IgH chromosome translocations in vivo. *Cell*. 2004;118(4):431-8.
66. Kuppers R. Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer*. 2005;5(4):251-62.
67. 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;48(11):1281-6.
68. Dal Porto JM, Gauld SB, Merrell KT, Mills D, Pugh-Bernard AE, Cambier J. B cell antigen receptor signaling 101. *Mol Immunol*. 2004;41(6-7):599-613.
69. 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;2(9):a003004.





70. Kurosaki T, Hikida M. Tyrosine kinases and their substrates in B lymphocytes. *Immunol Rev.* 2009;228(1):132-48.
71. 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;463(7277):88-92.
72. Schmitz R, Young RM, Ceribelli M, Jhavar S, Xiao W, Zhang M, et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature.* 2012;490(7418):116-20.
73. 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;459(7247):717-21.
74. Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med.* 2002;346(25):1937-47.
75. Lenz G, Davis RE, Ngo VN, Lam L, George TC, Wright GW, et al. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science.* 2008;319(5870):1676-9.
76. Gauld SB, Dal Porto JM, Cambier JC. B cell antigen receptor signaling: roles in cell development and disease. *Science.* 2002;296(5573):1641-2.
77. Lam KP, Kuhn R, Rajewsky K. In vivo ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death. *Cell.* 1997;90(6):1073-83.
78. Kraus M, Alimzhanov MB, Rajewsky N, Rajewsky K. Survival of resting mature B lymphocytes depends on BCR signaling via the Igalphabeta heterodimer. *Cell.* 2004;117(6):787-800.
79. Thome M. CARMA1, BCL-10 and MALT1 in lymphocyte development and activation. *Nat Rev Immunol.* 2004;4(5):348-59.
80. Young RM, Wu T, Schmitz R, Dawood M, Xiao W, Phelan JD, et al. Survival of human lymphoma cells requires B-cell receptor engagement by self-antigens. *Proc Natl Acad Sci U S A.* 2015;112(44):13447-54.
81. 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;21(8):922-6.
82. Kee BL. E and ID proteins branch out. *Nat Rev Immunol.* 2009;9(3):175-84.
83. Richter J, Schlesner M, Hoffmann S, Kreuz M, Leich E, Burkhardt B, et al. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. *Nat Genet.* 2012;44(12):1316-20.
84. Mackay F, Schneider P. Cracking the BAFF code. *Nat Rev Immunol.* 2009;9(7):491-502.
85. Warnatz K, Salzer U, Rizzi M, Fischer B, Gutenberger S, Bohm 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;106(33):13945-50.
86. 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-kappaB family members RelB and cRel. *Blood.* 2016;127(10):1276-86.
87. He B, Santamaria R, Xu W, Cols M, Chen K, Puga I, et al. The transmembrane activator TAC1 triggers immunoglobulin class switching by activating B cells through the adaptor MyD88. *Nat Immunol.* 2010;11(9):836-45.
88. Novak AJ, Grote DM, Stenson M, Ziesmer SC, Witzig TE, Habermann TM, et al. Expression of BlyS and its receptors in B-cell non-Hodgkin lymphoma: correlation with disease activity and patient outcome. *Blood.* 2004;104(8):2247-53.
89. Hildebrand JM, Luo Z, Manske MK, Price-Troska T, Ziesmer SC, Lin W, et al. A BAFF-R mutation associated with non-Hodgkin lymphoma alters TRAF recruitment and reveals new insights into BAFF-R signaling. *J Exp Med.* 2010;207(12):2569-79.
90. 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;12(4):282-94.
91. Chiron D, Bekeredjian-Ding I, Pellat-Deceunynck C, Bataille R, Jego G. Toll-like receptors: lessons to learn from normal and malignant human B cells. *Blood.* 2008;112(6):2205-13.
92. Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature.* 2011;470(7332):115-9.
93. Rossi D, Trifonov V, Fangazio M, Brusca A, Rasi S, Spina V, et al. The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. *J Exp Med.* 2012;209(9):1537-51.
94. 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(9):826-33.
95. Xu L, Hunter ZR, Yang G, Zhou Y, Cao Y, Liu X, et al. MYD88 L265P in Waldenstrom macroglobulinemia, immunoglobulin M monoclonal gammopathy, and other B-cell lymphoproliferative disorders using conventional and quantitative allele-specific polymerase chain reaction. *Blood.* 2013;121(11):2051-8.
96. Iwasaki A, Medzhitov R. Regulation of adaptive immunity by the innate immune system. *Science.* 2010;327(5963):291-5.
97. Warner N, Nunez G. MyD88: a critical adaptor protein in innate immunity signal transduction. *J Immunol.* 2013;190(1):3-4.
98. Sen R, Baltimore D. Inducibility of kappa immunoglobulin enhancer-binding protein Nf-kappa B by a posttranslational mechanism. *Cell.* 1986;47(6):921-8.
99. Staudt LM. Oncogenic activation of NF-kappaB. *Cold Spring Harb Perspect Biol.* 2010;2(6):a000109.
100. 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;9(12):1371-8.



101. 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;10(11):1309-17.
102. Zhang Q, Lenardo MJ, Baltimore D. 30 Years of NF-kappaB: A Blossoming of Relevance to Human Pathobiology. *Cell.* 2017;168(1-2):37-57.
103. Kaileh M, Sen R. NF-kappaB function in B lymphocytes. *Immunol Rev.* 2012;246(1):254-71.
104. Shaffer AL, Rosenwald A, Hurt EM, Giltneane JM, Lam LT, Pickeral OK, et al. Signatures of the immune response. *Immunity.* 2001;15(3):375-85.
105. 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-kappaB transcription factor subunits. *J Exp Med.* 2014;211(10):2103-18.
106. 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;12(3):280-92.
107. 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;33(2):181-91.
108. 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;7(7):773-82.
109. 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;25(2):225-36.
110. Kuppers R. Molecular biology of Hodgkin lymphoma. *Hematology Am Soc Hematol Educ Program.* 2009:491-6.
111. Schmitz R, Hansmann ML, Bohle V, Martin-Subero JI, Hartmann S, Mechttersheimer G, et al. TNFAIP3 (A20) is a tumor suppressor gene in Hodgkin lymphoma and primary mediastinal B cell lymphoma. *J Exp Med.* 2009;206(5):981-9.
112. Kato M, Sanada M, Kato I, Sato Y, Takita J, Takeuchi K, et al. Frequent inactivation of A20 in B-cell lymphomas. *Nature.* 2009;459(7247):712-6.
113. 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;114(12):2467-75.
114. 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.
115. 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(9):1088-101.
116. Mansouri L, Sutton LA, Ljungstrom V, Bondza S, Arngarden L, Bhoi S, et al. Functional loss of IkappaBepsilon leads to NF-kappaB deregulation in aggressive chronic lymphocytic leukemia. *J Exp Med.* 2015;212(6):833-43.
117. Mansouri L, Noerenberg D, Young E, Mylonas E, Abdulla M, Frick M, et al. Frequent NFKBIE deletions are associated with poor outcome in primary mediastinal B-cell lymphoma. *Blood.* 2016;128(23):2666-70.
118. Rossi D, Deaglio S, Dominguez-Sola D, Rasi S, Vaisitti T, Agostinelli C, et al. Alteration of BIRC3 and multiple other NF-kappaB pathway genes in splenic marginal zone lymphoma. *Blood.* 2011;118(18):4930-4.
119. Rossi D, Gaidano G. The clinical implications of gene mutations in chronic lymphocytic leukaemia. *Br J Cancer.* 2016;114(8):849-54.
120. Rossi D, Ciardullo C, Gaidano G. Genetic aberrations of signaling pathways in lymphomagenesis: revelations from next generation sequencing studies. *Semin Cancer Biol.* 2013;23(6):422-30.
121. Hu S, Du MQ, Park SM, Alcivar A, Qu L, Gupta S, et al. cIAP2 is a ubiquitin protein ligase for BCL10 and is dysregulated in mucosa-associated lymphoid tissue lymphomas. *J Clin Invest.* 2006;116(1):174-81.
122. Ying CY, Dominguez-Sola D, Fabi M, Lorenz IC, Hussein S, Bansal M, et al. MEF2B mutations lead to deregulated expression of the oncogene BCL6 in diffuse large B cell lymphoma. *Nat Immunol.* 2013;14(10):1084-92.
123. 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;476(7360):298-303.
124. Brescia P, Schneider C, Holmes AB, Shen Q, Hussein S, Pasqualucci L, et al. MEF2B Instructs Germinal Center Development and Acts as an Oncogene in B Cell Lymphomagenesis. *Cancer Cell.* 2018;34(3):453-65 e9.
125. Dalla-Favera R, Bregni M, Erikson J, Patterson D, Gallo RC, Croce CM. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc Natl Acad Sci U S A.* 1982;79(24):7824-7.
126. Taub R, Kirsch I, Morton C, Lenoir G, Swan D, Tronick S, et al. Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. *Proc Natl Acad Sci U S A.* 1982;79(24):7837-41.
127. Calado DP, Sasaki Y, Godinho SA, Pellerin A, Kochert K, Sleckman BP, et al. The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat Immunol.* 2012;13(11):1092-100.
128. Dominguez-Sola D, Victoria 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;13(11):1083-91.
129. Klein U, Tu Y, Stolovitzky GA, Keller JL, Haddad J, Jr., Miljkovic V, et al. Transcriptional analysis of the B cell germinal center reaction. *Proc Natl Acad Sci U S A.* 2003;100(5):2639-44.
130. Lin CY, Loven J, Rahl PB, Paranal RM, Burge CB, Bradner JE, et al. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell.* 2012;151(1):56-67.
131. 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;151(1):68-79.
132. Pasqualucci L, Dalla-Favera R. Genetics of diffuse large B-cell lymphoma. *Blood.* 2018;131(21):2307-19.



133. Saito M, Novak U, Piovan E, Basso K, Sumazin P, Schneider C, et al. BCL6 suppression of BCL2 via Miz1 and its disruption in diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A*. 2009;106(27):11294-9.
134. Pasqualucci L, Neumeister P, Goossens T, Nanjangud G, Chaganti RS, Kuppers R, et al. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature*. 2001;412(6844):341-6.
135. 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.
136. Cobaleda C, Jochum W, Busslinger M. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature*. 2007;449(7161):473-7.
137. Nutt SL, Heavey B, Rolink AG, Busslinger M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature*. 1999;401(6753):556-62.
138. Cobaleda C, Schebesta A, Delogu A, Busslinger M. Pax5: the guardian of B cell identity and function. *Nat Immunol*. 2007;8(5):463-70.
139. Mikkola I, Heavey B, Horcher M, Busslinger M. Reversion of B cell commitment upon loss of Pax5 expression. *Science*. 2002;297(5578):110-3.
140. Revilla IDR, Bilic I, Vilagos B, Tagoh H, Ebert A, Tamir IM, et al. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J*. 2012;31(14):3130-46.
141. Schebesta A, McManus S, Salvaggio G, Delogu A, Busslinger GA, Busslinger M. Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration, and immune function. *Immunity*. 2007;27(1):49-63.
142. Pridans C, Holmes ML, Polli M, Wettenhall JM, Dakic A, Corcoran LM, et al. Identification of Pax5 target genes in early B cell differentiation. *J Immunol*. 2008;180(3):1719-28.
143. 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;26(5):555-66.
144. Rinkenberger JL, Wallin JJ, Johnson KW, Koshland ME. An interleukin-2 signal relieves BSAP (Pax5)-mediated repression of the immunoglobulin J chain gene. *Immunity*. 1996;5(4):377-86.
145. Reimold AM, Ponath PD, Li YS, Hardy RR, David CS, Strominger JL, et al. Transcription factor B cell lineage-specific activator protein regulates the gene for human X-box binding protein 1. *J Exp Med*. 1996;183(2):393-401.
146. Nera KP, Kohonen P, Narvi E, Peippo A, Mustonen L, Terho P, et al. Loss of Pax5 promotes plasma cell differentiation. *Immunity*. 2006;24(3):283-93.
147. 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;24(3):269-81.
148. Usui T, Wakatsuki Y, Matsunaga Y, Kaneko S, Koseki H, Kita T. Overexpression of B cell-specific activator protein (BSAP/Pax-5) in a late B cell is sufficient to suppress differentiation to an Ig high producer cell with plasma cell phenotype. *J Immunol*. 1997;158(7):3197-204.
149. Lin KI, Angelin-Duclos C, Kuo TC, Calame K. Blimp-1-dependent repression of Pax-5 is required for differentiation of B cells to immunoglobulin M-secreting plasma cells. *Mol Cell Biol*. 2002;22(13):4771-80.
150. Cozma D, Yu D, Hodawadekar S, Azvolinsky A, Grande S, Tobias JW, et al. B cell activator PAX5 promotes lymphomagenesis through stimulation of B cell receptor signaling. *J Clin Invest*. 2007;117(9):2602-10.
151. Morrison AM, Jager U, Chott A, Schebesta M, Haas OA, Busslinger M. Deregulated PAX-5 transcription from a translocated IgH promoter in marginal zone lymphoma. *Blood*. 1998;92(10):3865-78.
152. Souabni A, Cobaleda C, Schebesta M, Busslinger M. Pax5 promotes B lymphopoiesis and blocks T cell development by repressing Notch1. *Immunity*. 2002;17(6):781-93.
153. Puente XS, Bea S, Valdes-Mas R, Villamor N, Gutierrez-Abril J, Martin-Subero JL, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2015;526(7574):519-24.
154. 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(2):481-94 e15.
155. Murre C. Helix-loop-helix proteins and lymphocyte development. *Nat Immunol*. 2005;6(11):1079-86.
156. Kwon K, Hutter C, Sun Q, Bilic I, Cobaleda C, Malin S, et al. Instructive role of the transcription factor E2A in early B lymphopoiesis and germinal center B cell development. *Immunity*. 2008;28(6):751-62.
157. Igarashi K, Kataoka K, Itoh K, Hayashi N, Nishizawa M, Yamamoto M. Regulation of transcription by dimerization of erythroid factor NF-E2 p45 with small Maf proteins. *Nature*. 1994;367(6463):568-72.
158. Muto A, Hoshino H, Madisen L, Yanai N, Obinata M, Karasuyama H, et al. Identification of Bach2 as a B-cell-specific partner for small maf proteins that negatively regulate the immunoglobulin heavy chain gene 3' enhancer. *EMBO J*. 1998;17(19):5734-43.
159. Muto A, Tashiro S, Nakajima O, Hoshino H, Takahashi S, Sakoda E, et al. The transcriptional programme of antibody class switching involves the repressor Bach2. *Nature*. 2004;429(6991):566-71.
160. Ochiai K, Katoh Y, Ikura T, Hoshikawa Y, Noda T, Karasuyama H, et al. Plasmacytic transcription factor Blimp-1 is repressed by Bach2 in B cells. *J Biol Chem*. 2006;281(50):38226-34.
161. 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;20(3):453-60.
162. 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;29(23):4048-61.
163. 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;123(7):1012-20.
164. Sasaki S, Ito E, Toki T, Maekawa T, Kanezaki R, Umenai T, et al. Cloning and expression of human B cell-specific transcription factor BACH2 mapped to chromosome 6q15. *Oncogene*. 2000;19(33):3739-49.
165. Kamio T, Toki T, Kanezaki R, Sasaki S, Tandai S, Terui K, et al. B-cell-specific transcription factor BACH2 modifies the cytotoxic effects of anticancer drugs. *Blood*. 2003;102(9):3317-22.





166. Sakane-Ishikawa E, Nakatsuka S, Tomita Y, Fujita S, Nakamichi I, Takakuwa T, et al. Prognostic significance of BACH2 expression in diffuse large B-cell lymphoma: a study of the Osaka Lymphoma Study Group. *J Clin Oncol*. 2005;23(31):8012-7.
167. Herbaux C, Bertrand E, Marot G, Roumier C, Poret N, Soenen V, et al. BACH2 promotes indolent clinical presentation in Waldenstrom macroglobulinemia. *Oncotarget*. 2017;8(34):57451-9.
168. Dent AL, Shaffer AL, Yu X, Allman D, Staudt LM. Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science*. 1997;276(5312):589-92.
169. Dalla-Favera R, Ye BH, Cattoretti G, Lo Coco F, Chang CC, Zhang J, et al. BCL-6 in diffuse large-cell lymphomas. *Important Adv Oncol*. 1996:139-48.
170. 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;95(20):11816-21.
171. Basso K, Saito M, Sumazin P, Margolin AA, Wang K, Lim WK, et al. Integrated biochemical and computational approach identifies BCL6 direct target genes controlling multiple pathways in normal germinal center B cells. *Blood*. 2010;115(5):975-84.
172. Ci W, Polo JM, Cerchiotti L, Shaknovich R, Wang L, Yang SN, et al. The BCL6 transcriptional program features repression of multiple oncogenes in primary B cells and is deregulated in DLBCL. *Blood*. 2009;113(22):5536-48.
173. Huang C, Hatzi K, Melnick A. Lineage-specific functions of Bcl-6 in immunity and inflammation are mediated by distinct biochemical mechanisms. *Nat Immunol*. 2013;14(4):380-8.
174. Hatzi K, Jiang Y, Huang C, Garrett-Bakelman F, Gearhart MD, Giannopoulou 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(3):578-88.
175. Huang C, Gonzalez DG, Cote CM, Jiang Y, Hatzi K, Teater M, et al. The BCL6 RD2 domain governs commitment of activated B cells to form germinal centers. *Cell Rep*. 2014;8(5):1497-508.
176. 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(5):445-55.
177. Tunyaplin C, Shaffer AL, Angelin-Duclos CD, Yu X, Staudt LM, Calame KL. Direct repression of prdm1 by Bcl-6 inhibits plasmacytic differentiation. *J Immunol*. 2004;173(2):1158-65.
178. Linterman MA, Beaton L, Yu D, Ramiscal RR, Srivastava M, Hogan JJ, et al. IL-21 acts directly on B cells to regulate Bcl-6 expression and germinal center responses. *J Exp Med*. 2010;207(2):353-63.
179. 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;207(2):365-78.
180. Lee CH, Melchers M, Wang H, Torrey TA, Slota R, Qi CF, et al. Regulation of the germinal center gene program by interferon (IFN) regulatory factor 8/IFN consensus sequence-binding protein. *J Exp Med*. 2006;203(1):63-72.
181. 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(11):2169-81.
182. Cimmino L, Martins GA, Liao J, Magnusdottir E, Grunig G, Perez RK, et al. Blimp-1 attenuates Th1 differentiation by repression of ifng, tbx21, and bcl6 gene expression. *J Immunol*. 2008;181(4):2338-47.
183. Pasqualucci L, Migliazza A, Basso K, Houldsworth J, Chaganti RS, Dalla-Favera R. Mutations of the BCL6 proto-oncogene disrupt its negative autoregulation in diffuse large B-cell lymphoma. *Blood*. 2003;101(8):2914-23.
184. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*. 2002;298(5595):1039-43.
185. Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell*. 2002;111(2):185-96.
186. Muller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, et al. Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell*. 2002;111(2):197-208.
187. Su IH, Basavaraj A, Krutchinsky AN, Hobert O, Ullrich A, Chait BT, et al. Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat Immunol*. 2003;4(2):124-31.
188. van Galen JC, Dukers DF, Giroth C, Sewalt RG, Otte AP, Meijer CJ, et al. Distinct expression patterns of polycomb oncoproteins and their binding partners during the germinal center reaction. *Eur J Immunol*. 2004;34(7):1870-81.
189. 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(5):677-92.
190. 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(12):5009-22.
191. Velichutina I, Shaknovich R, Geng H, Johnson NA, Gascoyne RD, Melnick AM, et al. EZH2-mediated epigenetic silencing in germinal center B cells contributes to proliferation and lymphomagenesis. *Blood*. 2010;116(24):5247-55.
192. Beguelin W, Teater M, Gearhart MD, Calvo Fernandez MT, Goldstein RL, Cardenas MG, et al. EZH2 and BCL6 Cooperate to Assemble CBX8-BCOR Complex to Repress Bivalent Promoters, Mediate Germinal Center Formation and Lymphomagenesis. *Cancer Cell*. 2016;30(2):197-213.
193. Basso K, Dalla-Favera R. Roles of BCL6 in normal and transformed germinal center B cells. *Immunol Rev*. 2012;247(1):172-83.
194. Ye BH, Chaganti S, Chang CC, Niu H, Corradini P, Chaganti RS, et al. Chromosomal translocations cause deregulated BCL6 expression by promoter substitution in B cell lymphoma. *EMBO J*. 1995;14(24):6209-17.
195. Chen W, Iida S, Louie DC, Dalla-Favera R, Chaganti RS. Heterologous promoters fused to BCL6 by chromosomal translocations affecting band 3q27 cause its deregulated expression during B-cell differentiation. *Blood*. 1998;91(2):603-7.



196. Duan S, Cermak L, Pagan JK, Rossi M, Martinengo C, di Celle PF, et al. FBXO11 targets BCL6 for degradation and is inactivated in diffuse large B-cell lymphomas. *Nature*. 2012;481(7379):90-3.
197. Pasqualucci L, Dominguez-Sola D, Chiarenza A, Fabbri G, Grunn A, Trifonov V, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature*. 2011;471(7337):189-95.
198. Bereshchenko OR, Gu W, Dalla-Favera R. Acetylation inactivates the transcriptional repressor BCL6. *Nat Genet*. 2002;32(4):606-13.
199. Basso K, Dalla-Favera R. Germinal centres and B cell lymphomagenesis. *Nat Rev Immunol*. 2015;15(3):172-84.
200. Schubart K, Massa S, Schubart D, Corcoran LM, Rolink AG, Matthias P. B cell development and immunoglobulin gene transcription in the absence of Oct-2 and OBF-1. *Nat Immunol*. 2001;2(1):69-74.
201. Corcoran LM, Karvelas M, Nossal GJ, Ye ZS, Jacks T, Baltimore D. Oct-2, although not required for early B-cell development, is critical for later B-cell maturation and for postnatal survival. *Genes Dev*. 1993;7(4):570-82.
202. Corcoran LM, Karvelas M. Oct-2 is required early in T cell-independent B cell activation for G1 progression and for proliferation. *Immunity*. 1994;1(8):635-45.
203. Humbert PO, Corcoran LM. oct-2 gene disruption eliminates the peritoneal B-1 lymphocyte lineage and attenuates B-2 cell maturation and function. *J Immunol*. 1997;159(11):5273-84.
204. Schubart DB, Rolink A, Kosco-Vilbois MH, Botteri F, Matthias P. B-cell-specific coactivator OBF-1/OCA-B/Bob1 required for immune response and germinal centre formation. *Nature*. 1996;383(6600):538-42.
205. Corcoran LM, Hasbold J, Dietrich W, Hawkins E, Kallies A, Nutt SL, et al. Differential requirement for OBF-1 during antibody-secreting cell differentiation. *J Exp Med*. 2005;201(9):1385-96.
206. Emslie D, D'Costa K, Hasbold J, Metcalf D, Takatsu K, Hodgkin PO, et al. Oct2 enhances antibody-secreting cell differentiation through regulation of IL-5 receptor alpha chain expression on activated B cells. *J Exp Med*. 2008;205(2):409-21.
207. Karnowski A, Chevrier S, Belz GT, Mount A, Emslie D, D'Costa K, et al. B and T cells collaborate in antiviral responses via IL-6, IL-21, and transcriptional activator and coactivator, Oct2 and OBF-1. *J Exp Med*. 2012;209(11):2049-64.
208. Kanno Y, Levi BZ, Tamura T, Ozato K. Immune cell-specific amplification of interferon signaling by the IRF-4/8-PU.1 complex. *J Interferon Cytokine Res*. 2005;25(12):770-9.
209. 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;201(2):221-31.
210. Wang H, Yan M, Sun J, Jain S, Yoshimi R, Abolfath SM, et al. A reporter mouse reveals lineage-specific and heterogeneous expression of IRF8 during lymphoid and myeloid cell differentiation. *J Immunol*. 2014;193(4):1766-77.
211. Su GH, Ip HS, Cobb BS, Lu MM, Chen HM, Simon MC. The Ets protein Spi-B is expressed exclusively in B cells and T cells during development. *J Exp Med*. 1996;184(1):203-14.
212. Bartholdy B, Du Roure C, Bordon A, Emslie D, Corcoran LM, Matthias P. The Ets factor Spi-B is a direct critical target of the coactivator OBF-1. *Proc Natl Acad Sci U S A*. 2006;103(31):11665-70.
213. Shaffer AL, Lin KI, Kuo TC, Yu X, Hurt EM, Rosenwald A, et al. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity*. 2002;17(1):51-62.
214. 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;106(6):2083-90.
215. 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;202(10):1411-22.
216. Feng J, Wang H, Shin DM, Masiuk M, Qi CF, Morse HC, 3rd. IFN regulatory factor 8 restricts the size of the marginal zone and follicular B cell pools. *J Immunol*. 2011;186(3):1458-66.
217. Su GH, Chen HM, Muthusamy N, Garrett-Sinha LA, Baunoch D, Tenen DG, et al. Defective B cell receptor-mediated responses in mice lacking the Ets protein, Spi-B. *EMBO J*. 1997;16(23):7118-29.
218. 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;38(4):576-89.
219. Shin DM, Lee CH, Morse HC, 3rd. IRF8 governs expression of genes involved in innate and adaptive immunity in human and mouse germinal center B cells. *PLoS One*. 2011;6(11):e27384.
220. Hasbold J, Corcoran LM, Tarlinton DM, Tangye SG, Hodgkin PD. Evidence from the generation of immunoglobulin G-secreting cells that stochastic mechanisms regulate lymphocyte differentiation. *Nat Immunol*. 2004;5(1):55-63.
221. Schmidlin H, Diehl SA, Nagasawa M, Scheeren FA, Schotte R, Uittenbogaart CH, et al. Spi-B inhibits human plasma cell differentiation by repressing BLIMP1 and XBP-1 expression. *Blood*. 2008;112(5):1804-12.
222. 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;27(6):847-59.
223. 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(11):2183-98.
224. 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;8(1):1426.
225. 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(2):951-62.
226. John S, Russell L, Chin SS, Luo W, Oshima R, Garrett-Sinha LA. Transcription factor Ets1, but not the closely related factor Ets2, inhibits antibody-secreting cell differentiation. *Mol Cell Biol*. 2014;34(3):522-32.
227. Ochiai K, Maischein-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;38(5):918-29.



228. Willis SN, Good-Jacobson KL, Curtis J, Light A, Tellier J, Shi W, et al. Transcription factor IRF4 regulates germinal center cell formation through a B cell-intrinsic mechanism. *J Immunol*. 2014;192(7):3200-6.
229. Falini B, Fizzotti M, Pucciarini A, Bigerna B, Marafioti T, Gambacorta M, et al. A monoclonal antibody (MUM1p) detects expression of the MUM1/IRF4 protein in a subset of germinal center B cells, plasma cells, and activated T cells. *Blood*. 2000;95(6):2084-92.
230. Ciofani M, Madar A, Galan C, Sellars M, Mace K, Pauli F, et al. A validated regulatory network for Th17 cell specification. *Cell*. 2012;151(2):289-303.
231. Glasmacher E, Agrawal S, Chang AB, Murphy TL, Zeng W, Vander Lugt B, et al. A genomic regulatory element that directs assembly and function of immune-specific AP-1-IRF complexes. *Science*. 2012;338(6109):975-80.
232. Li P, Spolski R, Liao W, Wang L, Murphy TL, Murphy KM, et al. BATF-JUN is critical for IRF4-mediated transcription in T cells. *Nature*. 2012;490(7421):543-6.
233. Rengarajan J, Mowen KA, McBride KD, Smith ED, Singh H, Glimcher LH. Interferon regulatory factor 4 (IRF4) interacts with NFATc2 to modulate interleukin 4 gene expression. *J Exp Med*. 2002;195(8):1003-12.
234. Ise W, Kohyama M, Schraml BU, Zhang T, Schwer B, Basu U, et al. The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. *Nat Immunol*. 2011;12(6):536-43.
235. Eisenbeis CF, Singh H, Storb U. PU.1 is a component of a multiprotein complex which binds an essential site in the murine immunoglobulin lambda 2-4 enhancer. *Mol Cell Biol*. 1993;13(10):6452-61.
236. Decker T, Pasca di Magliano M, McManus S, Sun Q, Bonifer C, Tagoh H, et al. Stepwise activation of enhancer and promoter regions of the B cell commitment gene Pax5 in early lymphopoiesis. *Immunity*. 2009;30(4):508-20.
237. Turner CA, Jr., Mack DH, Davis MM. Blimp-1, a novel zinc finger-containing protein that can drive the maturation of B lymphocytes into immunoglobulin-secreting cells. *Cell*. 1994;77(2):297-306.
238. 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;200(8):967-77.
239. Angelin-Duclos C, Cattoretti G, Lin KI, Calame K. Commitment of B lymphocytes to a plasma cell fate is associated with Blimp-1 expression in vivo. *J Immunol*. 2000;165(10):5462-71.
240. Soro PG, Morales AP, Martinez MJ, Morales AS, Copin 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*. 1999;163(2):611-7.
241. Vasanwala FH, Kusam S, Toney LM, Dent AL. Repression of AP-1 function: a mechanism for the regulation of Blimp-1 expression and B lymphocyte differentiation by the B cell lymphoma-6 protooncogene. *J Immunol*. 2002;169(4):1922-9.
242. Kwon H, Thierry-Mieg D, Thierry-Mieg J, Kim HP, Oh J, Tunyaplin C, et al. Analysis of interleukin-21-induced Prdm1 gene regulation reveals functional cooperation of STAT3 and IRF4 transcription factors. *Immunity*. 2009;31(6):941-52.
243. 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;20(23):8684-95.
244. Shapiro-Shelef M, Lin KI, Savitsky D, Liao J, Calame K. Blimp-1 is required for maintenance of long-lived plasma cells in the bone marrow. *J Exp Med*. 2005;202(11):1471-6.
245. Doody GM, Care MA, Burgoyne NJ, Bradford JR, Bota M, Bonifer C, et al. An extended set of PRDM1/BLIMP1 target genes links binding motif type to dynamic repression. *Nucleic Acids Res*. 2010;38(16):5336-50.
246. Doody GM, Stephenson S, McManamy C, Tooze RM. PRDM1/BLIMP-1 modulates IFN-gamma-dependent control of the MHC class I antigen-processing and peptide-loading pathway. *J Immunol*. 2007;179(11):7614-23.
247. Savage HP, Yenson VM, Sawhney SS, Mousseau BJ, Lund FE, Baumgarth N. Blimp-1-dependent and -independent natural antibody production by B-1 and B-1-derived plasma cells. *J Exp Med*. 2017;214(9):2777-94.
248. D'Costa K, Emslie D, Metcalf D, Smyth GK, Karnowski A, Kallies A, et al. Blimp1 is limiting for transformation in a mouse plasmacytoma model. *Blood*. 2009;113(23):5911-9.
249. 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;203(2):311-7.
250. Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, Auclair D, et al. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell*. 2014;25(1):91-101.
251. 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;18(6):568-79.
252. 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;194(12):1861-74.
253. Todd DJ, McHeyzer-Williams LJ, Kowal C, Lee AH, Volpe BT, Diamond B, et al. XBP1 governs late events in plasma cell differentiation and is not required for antigen-specific memory B cell development. *J Exp Med*. 2009;206(10):2151-9.
254. Taubenheim N, Tarlinton DM, Crawford S, Corcoran LM, Hodgkin PD, Nutt SL. High rate of antibody secretion is not integral to plasma cell differentiation as revealed by XBP-1 deficiency. *J Immunol*. 2012;189(7):3328-38.
255. Iwakoshi NN, Lee AH, Glimcher LH. The X-box binding protein-1 transcription factor is required for plasma cell differentiation and the unfolded protein response. *Immunol Rev*. 2003;194:29-38.
256. Tam W, Gomez M, Chadburn A, Lee JW, Chan WC, Knowles DM. Mutational analysis of PRDM1 indicates a tumor-suppressor role in diffuse large B-cell lymphomas. *Blood*. 2006;107(10):4090-100.
257. Iqbal J, Greiner TC, Patel K, Dave BJ, Smith L, Ji J, et al. Distinctive patterns of BCL6 molecular alterations and their functional consequences in different subgroups of diffuse large B-cell lymphoma. *Leukemia*. 2007;21(11):2332-43.
258. 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(3):633-43.





259. Lenz G, Wright GW, Emre NC, Kohlhammer H, Dave SS, Davis RE, et al. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc Natl Acad Sci U S A*. 2008;105(36):13520-5.
260. Care MA, Cocco M, Laye JP, Barnes N, Huang Y, Wang M, et al. SPIB and BATF provide alternate determinants of IRF4 occupancy in diffuse large B-cell lymphoma linked to disease heterogeneity. *Nucleic Acids Res*. 2014;42(12):7591-610.
261. 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(6):723-37.
262. Calado DP, Zhang B, Srinivasan L, Sasaki Y, Seagal J, Unitt C, et al. Constitutive canonical NF-kappaB activation cooperates with disruption of BLIMP1 in the pathogenesis of activated B cell-like diffuse large cell lymphoma. *Cancer Cell*. 2010;18(6):580-9.
263. Leung-Hagesteijn C, Erdmann N, Cheung G, Keats JJ, Stewart AK, Reece DE, et al. Xbp1s-negative tumor B cells and pre-plasmablasts mediate therapeutic proteasome inhibitor resistance in multiple myeloma. *Cancer Cell*. 2013;24(3):289-304.
264. Shaffer AL, Emre NC, Lamy L, Ngo VN, Wright G, Xiao W, et al. IRF4 addiction in multiple myeloma. *Nature*. 2008;454(7201):226-31.
265. Lin Y, Wong K, Calame K. Repression of c-myc transcription by Blimp-1, an inducer of terminal B cell differentiation. *Science*. 1997;276(5312):596-9.
266. Yoshida T, Mei H, Dorner T, Hiepe F, Radbruch A, Fillatreau S, et al. Memory B and memory plasma cells. *Immunol Rev*. 2010;237(1):117-39.
267. 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;203(12):2683-90.
268. 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;105(4):1614-21.
269. Hauser AE, Debes GF, Arce S, Cassese G, Hamann A, Radbruch A, et al. Chemotactic responsiveness toward ligands for CXCR3 and CXCR4 is regulated on plasma blasts during the time course of a memory immune response. *J Immunol*. 2002;169(3):1277-82.
270. Tangye SG. Staying alive: regulation of plasma cell survival. *Trends Immunol*. 2011;32(12):595-602.
271. Winkelmann R, Sandrock L, Porstner M, Roth E, Mathews M, Hobeika E, et al. B cell homeostasis and plasma cell homing controlled by Kruppel-like factor 2. *Proc Natl Acad Sci U S A*. 2011;108(2):710-5.
272. Wang Y, Bhattacharya D. Adjuvant-specific regulation of long-term antibody responses by ZBTB20. *J Exp Med*. 2014;211(5):841-56.
273. Tokoyoda K, Egawa T, Sugiyama T, Choi BI, Nagasawa T. Cellular niches controlling B lymphocyte behavior within bone marrow during development. *Immunity*. 2004;20(6):707-18.
274. Chu VT, Frohlich A, Steinhauser G, Scheel T, Roch T, Fillatreau S, et al. Eosinophils are required for the maintenance of plasma cells in the bone marrow. *Nat Immunol*. 2011;12(2):151-9.
275. Winter O, Mohr E, Manz RA. Alternative cell types form a Multi-Component-Plasma-Cell-Niche. *Immunol Lett*. 2011;141(1):145-6.
276. Kawano MM, Mihara K, Huang N, Tsujimoto T, Kuramoto A. Differentiation of early plasma cells on bone marrow stromal cells requires interleukin-6 for escaping from apoptosis. *Blood*. 1995;85(2):487-94.
277. Belnoue E, Pihlgren M, McGaha TL, Tougne C, Rochat AF, Bossen C, et al. APRIL is critical for plasmablast survival in the bone marrow and poorly expressed by early-life bone marrow stromal cells. *Blood*. 2008;111(5):2755-64.
278. Benson MJ, Dillon SR, Castigli E, Geha RS, Xu S, Lam KP, et al. Cutting edge: the dependence of plasma cells and independence of memory B cells on BAFF and APRIL. *J Immunol*. 2008;180(6):3655-9.
279. O'Connor BP, Raman VS, Erickson LD, Cook WJ, Weaver LK, Ahonen C, et al. BCMA is essential for the survival of long-lived bone marrow plasma cells. *J Exp Med*. 2004;199(1):91-8.
280. Peperzak V, Vikstrom I, Walker J, Glaser SP, LePage M, Coquery CM, et al. Mcl-1 is essential for the survival of plasma cells. *Nat Immunol*. 2013;14(3):290-7.
281. Saini AS, Shenoy GN, Rath S, Bal V, George A. Inducible nitric oxide synthase is a major intermediate in signaling pathways for the survival of plasma cells. *Nat Immunol*. 2014;15(3):275-82.
282. Xiang Z, Cutler AJ, Brownlie RJ, Fairfax K, Lawlor KE, Severinson E, et al. FcgammaRIIb controls bone marrow plasma cell persistence and apoptosis. *Nat Immunol*. 2007;8(4):419-29.
283. Herrinton LJ, Weiss NS. Incidence of Waldenstrom's macroglobulinemia. *Blood*. 1993;82(10):3148-50.
284. Dimopoulos MA, Panayiotidis P, Mouloupoulos LA, Sfrikakis P, Dalakas M. Waldenstrom's macroglobulinemia: clinical features, complications, and management. *J Clin Oncol*. 2000;18(1):214-26.
285. Wang H, Chen Y, Li F, Delasalle K, Wang J, Alexanian R, et al. Temporal and geographic variations of Waldenstrom macroglobulinemia incidence: a large population-based study. *Cancer*. 2012;118(15):3793-800.
286. Ailawadhi S, Kardosh A, Yang D, Cozen W, Patel G, Alamgir MA, et al. Outcome disparities among ethnic subgroups of Waldenstrom's macroglobulinemia: a population-based study. *Oncology*. 2014;86(5-6):253-62.
287. Iwanaga M, Chiang CJ, Soda M, Lai MS, Yang YW, Miyazaki Y, et al. Incidence of lymphoplasmacytic lymphoma/Waldenstrom's macroglobulinemia in Japan and Taiwan population-based cancer registries, 1996-2003. *Int J Cancer*. 2014;134(1):174-80.
288. Carreon JD, Morton LM, Devesa SS, Clarke CA, Gomez SL, Glaser SL, et al. Incidence of lymphoid neoplasms by subtype among six Asian ethnic groups in the United States, 1996-2004. *Cancer Causes Control*. 2008;19(10):1171-81.
289. Hanzis C, Ojha RP, Hunter Z, Manning R, Lewicki M, Brodsky P, et al. Associated malignancies in patients with Waldenstrom's macroglobulinemia and their kin. *Clin Lymphoma Myeloma Leuk*. 2011;11(1):88-92.





290. 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;102(10):3759-64.
291. Landgren O, Kyle RA, Pfeiffer RM, Katzmann JA, Caporaso NE, Hayes RB, et al. Monoclonal gammopathy of undetermined significance (MGUS) consistently precedes multiple myeloma: a prospective study. *Blood*. 2009;113(22):5412-7.
292. Aoki H, Takishita M, Kosaka M, Saito S. Frequent somatic mutations in D and/or JH segments of Ig gene in Waldenstrom's macroglobulinemia and chronic lymphocytic leukemia (CLL) with Richter's syndrome but not in common CLL. *Blood*. 1995;85(7):1913-9.
293. Wagner SD, Martinelli V, Luzzatto L. Similar patterns of V kappa gene usage but different degrees of somatic mutation in hairy cell leukemia, prolymphocytic leukemia, Waldenstrom's macroglobulinemia, and myeloma. *Blood*. 1994;83(12):3647-53.
294. Martin-Jimenez P, Garcia-Sanz R, Balanzategui A, Alcoceba M, Ocio E, Sanchez ML, et al. Molecular characterization of heavy chain immunoglobulin gene rearrangements in Waldenstrom's macroglobulinemia and IgM monoclonal gammopathy of undetermined significance. *Haematologica*. 2007;92(5):635-42.
295. Koshiol J, Gridley G, Engels EA, McMaster ML, Landgren O. Chronic immune stimulation and subsequent Waldenstrom macroglobulinemia. *Arch Intern Med*. 2008;168(17):1903-9.
296. Giordano TP, Henderson L, Landgren O, Chiao EY, Kramer JR, El-Serag H, et al. Risk of non-Hodgkin lymphoma and lymphoproliferative precursor diseases in US veterans with hepatitis C virus. *JAMA*. 2007;297(18):2010-7.
297. Smedby KE, Hjalgrim H, Askling J, Chang ET, Gregersen H, Porwit-MacDonald A, et al. Autoimmune and chronic inflammatory disorders and risk of non-Hodgkin lymphoma by subtype. *J Natl Cancer Inst*. 2006;98(1):51-60.
298. Vajdic CM, Landgren O, McMaster ML, Slager SL, Brooks-Wilson A, Smith A, et al. Medical history, lifestyle, family history, and occupational risk factors for lymphoplasmacytic lymphoma/Waldenstrom's macroglobulinemia: the InterLymph Non-Hodgkin Lymphoma Subtypes Project. *J Natl Cancer Inst Monogr*. 2014;2014(48):87-97.
299. Kristinsson SY, Bjorkholm M, Goldin LR, McMaster ML, Turesson I, Landgren O. Risk of lymphoproliferative disorders among first-degree relatives of lymphoplasmacytic lymphoma/Waldenstrom macroglobulinemia patients: a population-based study in Sweden. *Blood*. 2008;112(8):3052-6.
300. McMaster ML, Goldin LR, Bai Y, Ter-Minassian M, Boehringer S, Giambarresi TR, et al. Genomewide linkage screen for Waldenstrom macroglobulinemia susceptibility loci in high-risk families. *Am J Hum Genet*. 2006;79(4):695-701.
301. Roccaro AM, Sacco A, Shi J, Chiarini M, Perilla-Glen A, Manier S, et al. Exome sequencing reveals recurrent germ line variants in patients with familial Waldenstrom macroglobulinemia. *Blood*. 2016;127(21):2598-606.
302. McMaster ML, Berndt SI, Zhang J, Slager SL, Li SA, Vajdic CM, et al. Two high-risk susceptibility loci at 6p25.3 and 14q32.13 for Waldenstrom macroglobulinemia. *Nat Commun*. 2018;9(1):4182.
303. Grass S, Preuss KD, Wikowicz A, Terpos E, Ziepert M, Nikolaus D, et al. Hyperphosphorylated paratarg-7: a new molecularly defined risk factor for monoclonal gammopathy of undetermined significance of the IgM type and Waldenstrom macroglobulinemia. *Blood*. 2011;117(10):2918-23.
304. Simon L, Baron M, Leblond V. How we manage patients with Waldenstrom macroglobulinaemia. *Br J Haematol*. 2018;181(6):737-51.
305. Treon SP, Hunter ZR, Castillo JJ, Merlini G. Waldenstrom macroglobulinemia. *Hematol Oncol Clin North Am*. 2014;28(5):945-70.
306. Hunter ZR, Xu L, Yang G, Zhou Y, Liu X, Cao Y, et al. The genomic landscape of Waldenstrom macroglobulinemia is characterized by highly recurring MYD88 and WHIM-like CXCR4 mutations, and small somatic deletions associated with B-cell lymphomagenesis. *Blood*. 2014;123(11):1637-46.
307. 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;30(2):182-6.
308. San Miguel JF, Vidriales MB, Ocio E, Mateo G, Sanchez-Guijo F, Sanchez ML, et al. Immunophenotypic analysis of Waldenstrom's macroglobulinemia. *Semin Oncol*. 2003;30(2):187-95.
309. Paiva B, Montes MC, Garcia-Sanz R, Ocio EM, Alonso J, de Las Heras N, et al. Multiparameter flow cytometry for the identification of the Waldenstrom's clone in IgM-MGUS and Waldenstrom's Macroglobulinemia: new criteria for differential diagnosis and risk stratification. *Leukemia*. 2014;28(1):166-73.
310. Roberts MJ, Chadburn A, Ma S, Hyjek E, Peterson LC. Nuclear protein dysregulation in lymphoplasmacytic lymphoma/waldenstrom macroglobulinemia. *Am J Clin Pathol*. 2013;139(2):210-9.
311. Paiva B, Corchete LA, Vidriales MB, Garcia-Sanz R, Perez JJ, Aires-Mejia I, et al. The cellular origin and malignant transformation of Waldenstrom macroglobulinemia. *Blood*. 2015;125(15):2370-80.
312. Morel P, Duhamel A, Gobbi P, Dimopoulos MA, Dhodapkar MV, McCoy J, et al. International prognostic scoring system for Waldenstrom macroglobulinemia. *Blood*. 2009;113(18):4163-70.
313. Treon SP, Cao Y, Xu L, Yang G, Liu X, Hunter ZR. Somatic mutations in MYD88 and CXCR4 are determinants of clinical presentation and overall survival in Waldenstrom macroglobulinemia. *Blood*. 2014;123(18):2791-6.
314. Treon SP, Gustine J, Xu L, Manning RJ, Tsakmakis N, Demos M, et al. MYD88 wild-type Waldenstrom Macroglobulinaemia: differential diagnosis, risk of histological transformation, and overall survival. *Br J Haematol*. 2018;180(3):374-80.
315. Treon SP, Gustine J, Meid K, Yang G, Xu L, Liu X, et al. Ibrutinib Monotherapy in Symptomatic, Treatment-Naive Patients With Waldenstrom Macroglobulinemia. *J Clin Oncol*. 2018;JCO2018786426.



316. Dimopoulos MA, Zervas C, Zomas A, Hamilos G, Gika D, Efstathiou E, et al. Extended rituximab therapy for previously untreated patients with Waldenstrom's macroglobulinemia. *Clin Lymphoma*. 2002;3(3):163-6.
317. Gertz MA, Rue M, Blood E, Kaminer LS, Vesole DH, Greipp PR. Multicenter phase 2 trial of rituximab for Waldenstrom macroglobulinemia (WM): an Eastern Cooperative Oncology Group Study (E3A98). *Leuk Lymphoma*. 2004;45(10):2047-55.
318. Olszewski AJ, Chen C, Gutman R, Treon SP, Castillo JJ. Comparative outcomes of immunochemotherapy regimens in Waldenstrom macroglobulinaemia. *Br J Haematol*. 2017;179(1):106-15.
319. Leblond V, Kastritis E, Advani R, Ansell SM, Buske C, Castillo JJ, et al. Treatment recommendations from the Eighth International Workshop on Waldenstrom's Macroglobulinemia. *Blood*. 2016;128(10):1321-8.
320. Dimopoulos MA, Anagnostopoulos A, Kyrtsolis MC, Zervas K, Tsatalas C, Kokkinis G, et al. Primary treatment of Waldenstrom macroglobulinemia with dexamethasone, rituximab, and cyclophosphamide. *J Clin Oncol*. 2007;25(22):3344-9.
321. Rummel MJ, Niederle N, Maschmeyer G, Banat GA, von Grunhagen U, Losem C, et al. Bendamustine plus rituximab versus CHOP plus rituximab as first-line treatment for patients with indolent and mantle-cell lymphomas: an open-label, multicentre, randomised, phase 3 non-inferiority trial. *Lancet*. 2013;381(9873):1203-10.
322. Gavriatopoulou M, Kastritis E, Dimopoulos MA. Ibrutinib for rituximab-refractory Waldenstrom macroglobulinemia. *Oncotarget*. 2018;9(16):12536-7.
323. Treon SP, Branagan AR, Ioakimidis L, Soumerai JD, Patterson CJ, Turnbull B, et al. Long-term outcomes to fludarabine and rituximab in Waldenstrom macroglobulinemia. *Blood*. 2009;113(16):3673-8.
324. Leblond V, Johnson S, Chevret S, Copplestone A, Rule S, Tournilhac O, et al. Results of a randomized trial of chlorambucil versus fludarabine for patients with untreated Waldenstrom macroglobulinemia, marginal zone lymphoma, or lymphoplasmacytic lymphoma. *J Clin Oncol*. 2013;31(3):301-7.
325. Kuppers R, Klein U, Hansmann ML, Rajewsky K. Cellular origin of human B-cell lymphomas. *N Engl J Med*. 1999;341(20):1520-9.
326. Perez-Andres M, Paiva B, Nieto WG, Caraux A, Schmitz A, Almeida J, et al. Human peripheral blood B-cell compartments: a crossroad in B-cell traffic. *Cytometry B Clin Cytom*. 2010;78 Suppl 1:S47-60.
327. Kjeldsen MK, Perez-Andres M, Schmitz A, Johansen P, Boegsted M, Nyegaard M, et al. Multiparametric flow cytometry for identification and fluorescence activated cell sorting of five distinct B-cell subpopulations in normal tonsil tissue. *Am J Clin Pathol*. 2011;136(6):960-9.
328. Caraux A, Klein B, Paiva B, Bret C, Schmitz A, Fuhler GM, et al. Circulating human B and plasma cells. Age-associated changes in counts and detailed characterization of circulating normal CD138- and CD138+ plasma cells. *Haematologica*. 2010;95(6):1016-20.
329. 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;100(4):1505-7.
330. Kriangkum J, Taylor BJ, Reiman T, Belch AR, Pilarski LM. Origins of Waldenstrom's macroglobulinemia: does it arise from an unusual B-cell precursor? *Clin Lymphoma*. 2005;5(4):217-9.
331. 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 Waldenstrom macroglobulinemia/lymphoplasmacytic lymphomas. *Leukemia*. 2013;27(1):183-9.
332. 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;104(7):2134-42.
333. 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(4):649-54.
334. 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;69(8):3579-88.
335. Poulain S, Roumier C, Galiegue-Zouitina S, Daudignon A, Herbaux C, Aijjou R, et al. Genome wide SNP array identified multiple mechanisms of genetic changes in Waldenstrom macroglobulinemia. *Am J Hematol*. 2013;88(11):948-54.
336. Schop RF, Kuehl WM, Van Wier SA, Ahmann GJ, Price-Troska T, Bailey RJ, et al. Waldenstrom macroglobulinemia neoplastic cells lack immunoglobulin heavy chain locus translocations but have frequent 6q deletions. *Blood*. 2002;100(8):2996-3001.
337. Chang H, Qi C, Trieu Y, Jiang A, Young KH, Chesney A, et al. Prognostic relevance of 6q deletion in Waldenstrom's macroglobulinemia: a multicenter study. *Clin Lymphoma Myeloma*. 2009;9(1):36-8.
338. Ocio EM, Schop RF, Gonzalez B, Van Wier SA, Hernandez-Rivas JM, Gutierrez NC, et al. 6q deletion in Waldenstrom macroglobulinemia is associated with features of adverse prognosis. *Br J Haematol*. 2007;136(1):80-6.
339. Terre C, Nguyen-Khac F, Barin C, Mozziconacci MJ, Eclache V, Leonard C, et al. Trisomy 4, a new chromosomal abnormality in Waldenstrom's macroglobulinemia: a study of 39 cases. *Leukemia*. 2006;20(9):1634-6.
340. Varettoni M, Zibellini S, Defrancesco I, Ferretti VV, Rizzo E, Malcovati L, et al. Pattern of somatic mutations in patients with Waldenstrom macroglobulinemia or IgM monoclonal gammopathy of undetermined significance. *Haematologica*. 2017;102(12):2077-85.
341. Poulain S, Roumier C, Bertrand E, Renneville A, Caillault-Venet A, Doye E, et al. TP53 Mutation and Its Prognostic Significance in Waldenstrom's Macroglobulinemia. *Clin Cancer Res*. 2017;23(20):6325-35.
342. Cohen L, Henzel WJ, Baeuerle PA. IKAP is a scaffold protein of the I kappa B kinase complex. *Nature*. 1998;395(6699):292-6.



343. Loiarro M, Gallo G, Fanto N, De Santis R, Carminati P, Ruggiero V, et al. Identification of critical residues of the MyD88 death domain involved in the recruitment of downstream kinases. *J Biol Chem*. 2009;284(41):28093-103.
344. Lin SC, Lo YC, Wu H. Helical assembly in the MyD88-IRAK4-IRAK2 complex in TLR/IL-1R signalling. *Nature*. 2010;465(7300):885-90.
345. 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 Waldenstrom macroglobulinemia. *Blood*. 2013;122(7):1222-32.
346. Kawagoe T, Sato S, Matsushita K, Kato H, Matsui K, Kumagai Y, et al. Sequential control of Toll-like receptor-dependent responses by IRAK1 and IRAK2. *Nat Immunol*. 2008;9(6):684-91.
347. Leleu X, Eeckhoutte J, Jia X, Roccaro AM, Moreau AS, Farag M, et al. Targeting NF-kappaB in Waldenstrom macroglobulinemia. *Blood*. 2008;111(10):5068-77.
348. Yang G, Buhrlage SJ, Tan L, Liu X, Chen J, Xu L, et al. HCK is a survival determinant transactivated by mutated MYD88, and a direct target of ibrutinib. *Blood*. 2016;127(25):3237-52.
349. Jimenez C, Sebastian E, Chillon MC, Giraldo P, Mariano Hernandez J, Escalante F, et al. MYD88 L265P is a marker highly characteristic of, but not restricted to, Waldenstrom's macroglobulinemia. *Leukemia*. 2013;27(8):1722-8.
350. Varettoni M, Arcaini L, Zibellini S, Boveri E, Rattotti S, Riboni R, et al. Prevalence and clinical significance of the MYD88 (L265P) somatic mutation in Waldenstrom's macroglobulinemia and related lymphoid neoplasms. *Blood*. 2013;121(13):2522-8.
351. Poulain S, Roumier C, Decambron A, Renneville A, Herbaux C, Bertrand E, et al. MYD88 L265P mutation in Waldenstrom macroglobulinemia. *Blood*. 2013;121(22):4504-11.
352. 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;127(22):2732-41.
353. Haribabu B, Richardson RM, Fisher I, Sozzani S, Peiper SC, Horuk R, et al. Regulation of human chemokine receptors CXCR4. Role of phosphorylation in desensitization and internalization. *J Biol Chem*. 1997;272(45):28726-31.
354. Teicher BA, Fricker SP. CXCL12 (SDF-1)/CXCR4 pathway in cancer. *Clin Cancer Res*. 2010;16(11):2927-31.
355. Poulain S, Roumier C, Venet-Caillaud A, Figeac M, Herbaux C, Marot G, et al. Genomic Landscape of CXCR4 Mutations in Waldenstrom Macroglobulinemia. *Clin Cancer Res*. 2016;22(6):1480-8.
356. Xu L, Hunter ZR, Tsakmaklis N, Cao Y, Yang G, Chen J, et al. Clonal architecture of CXCR4 WHIM-like mutations in Waldenstrom Macroglobulinaemia. *Br J Haematol*. 2016;172(5):735-44.
357. Cao Y, Hunter ZR, Liu X, Xu L, Yang G, Chen J, et al. The WHIM-like CXCR4(S338X) somatic mutation activates AKT and ERK, and promotes resistance to ibrutinib and other agents used in the treatment of Waldenstrom's Macroglobulinemia. *Leukemia*. 2015;29(1):169-76.
358. Schmidt J, Federmann B, Schindler N, Steinhilber J, Bonzheim I, Fend F, et al. MYD88 L265P and CXCR4 mutations in lymphoplasmacytic lymphoma identify cases with high disease activity. *Br J Haematol*. 2015;169(6):795-803.
359. Hernandez PA, Gorlin RJ, Lukens JN, Taniuchi S, Bohinjec J, Francois F, et al. Mutations in the chemokine receptor gene CXCR4 are associated with WHIM syndrome, a combined immunodeficiency disease. *Nat Genet*. 2003;34(1):70-4.
360. Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science*. 2010;330(6007):1066-71.
361. Lagane B, Chow KY, Balabanian K, Levoye A, Harriague J, Planchenault T, et al. CXCR4 dimerization and beta-arrestin-mediated signaling account for the enhanced chemotaxis to CXCL12 in WHIM syndrome. *Blood*. 2008;112(1):34-44.
362. Orsini MJ, Parent JL, Mundell SJ, Marchese A, Benovic JL. Trafficking of the HIV coreceptor CXCR4. Role of arrestins and identification of residues in the c-terminal tail that mediate receptor internalization. *J Biol Chem*. 1999;274(43):31076-86.
363. Roccaro AM, Sacco A, Jimenez C, Maiso P, Moschetta M, Mishima Y, et al. C1013G/CXCR4 acts as a driver mutation of tumor progression and modulator of drug resistance in lymphoplasmacytic lymphoma. *Blood*. 2014;123(26):4120-31.
364. Bannard O, Horton RM, Allen CD, An J, Nagasawa T, Cyster JG. Germinal center centroblasts transition to a centrocyte phenotype according to a timed program and depend on the dark zone for effective selection. *Immunity*. 2013;39(5):912-24.
365. Ma Q, Jones D, Borghesani PR, Segal RA, Nagasawa T, Kishimoto T, et al. Impaired B-lymphopoiesis, myelopoiesis, and derailed cerebellar neuron migration in CXCR4- and SDF-1-deficient mice. *Proc Natl Acad Sci U S A*. 1998;95(16):9448-53.
366. Nie Y, Waite J, Brewer F, Sunshine MJ, Littman DR, Zou YR. The role of CXCR4 in maintaining peripheral B cell compartments and humoral immunity. *J Exp Med*. 2004;200(9):1145-56.
367. Tachibana K, Hirota S, Iizasa H, Yoshida H, Kawabata K, Kataoka Y, et al. The chemokine receptor CXCR4 is essential for vascularization of the gastrointestinal tract. *Nature*. 1998;393(6685):591-4.
368. Balabanian K, Brotin E, Biajoux V, Bouchet-Delbos L, Lainey E, Fenneteau O, et al. Proper desensitization of CXCR4 is required for lymphocyte development and peripheral compartmentalization in mice. *Blood*. 2012;119(24):5722-30.
369. Biajoux V, Natt J, Freitas C, Alouche N, Sacquin A, Hemon P, et al. Efficient Plasma Cell Differentiation and Trafficking Require Cxcr4 Desensitization. *Cell Rep*. 2016;17(1):193-205.
370. Hunter ZR, Yang G, Xu L, Liu X, Castillo JJ, Treon SP. Genomics, Signaling, and Treatment of Waldenstrom Macroglobulinemia. *J Clin Oncol*. 2017;35(9):994-1001.





371. Wu JN, Roberts CW. ARID1A mutations in cancer: another epigenetic tumor suppressor? *Cancer Discov.* 2013;3(1):35-43.
372. Wiegand KC, Shah SP, Al-Agha OM, Zhao Y, Tse K, Zeng T, et al. ARID1A mutations in endometriosis-associated ovarian carcinomas. *N Engl J Med.* 2010;363(16):1532-43.
373. Zang ZJ, Cutcutache I, Poon SL, Zhang SL, McPherson JR, Tao J, et al. Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes. *Nat Genet.* 2012;44(5):570-4.
374. Huang J, Deng Q, Wang Q, Li KY, Dai JH, Li N, et al. Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma. *Nat Genet.* 2012;44(10):1117-21.
375. Chng WJ, Schop RF, Price-Troska T, Ghobrial I, Kay N, Jelinek DF, et al. Gene-expression profiling of Waldenstrom macroglobulinemia reveals a phenotype more similar to chronic lymphocytic leukemia than multiple myeloma. *Blood.* 2006;108(8):2755-63.
376. Gutierrez NC, Ocio EM, de Las Rivas J, Maiso P, Delgado M, Ferminan E, et al. Gene expression profiling of B lymphocytes and plasma cells from Waldenstrom's macroglobulinemia: comparison with expression patterns of the same cell counterparts from chronic lymphocytic leukemia, multiple myeloma and normal individuals. *Leukemia.* 2007;21(3):541-9.
377. Hunter ZR, Xu L, Yang G, Tsakmaklis N, Vos JM, Liu X, et al. Transcriptome sequencing reveals a profile that corresponds to genomic variants in Waldenstrom macroglobulinemia. *Blood.* 2016;128(6):827-38.
378. Nunn MF, Seeburg PH, Moscovici C, Duesberg PH. Tripartite structure of the avian erythroblastosis virus E26 transforming gene. *Nature.* 1983;306(5941):391-5.
379. Leprince D, Gegonne A, Coll J, de Taisne C, Schneeberger A, Lagrou C, et al. A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature.* 1983;306(5941):395-7.
380. Hollenhorst PC, McIntosh LP, Graves BJ. Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu Rev Biochem.* 2011;80:437-71.
381. Sharrocks AD. The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol.* 2001;2(11):827-37.
382. Klambt C. The *Drosophila* gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells. *Development.* 1993;117(1):163-76.
383. Lacronique V, Boureux A, Valle VD, Poirel H, Quang CT, Mauchauffe M, et al. A TEL-JAK2 fusion protein with constitutive kinase activity in human leukemia. *Science.* 1997;278(5341):1309-12.
384. 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;20(15):4173-82.
385. Baker DA, Mille-Baker B, Wainwright SM, Ish-Horowicz D, Dibb NJ. Mae mediates MAP kinase phosphorylation of Ets transcription factors in *Drosophila*. *Nature.* 2001;411(6835):330-4.
386. Fenrick R, Amann JM, Lutterbach B, Wang L, Westendorf JJ, Downing JR, et al. Both TEL and AML-1 contribute repression domains to the t(12;21) fusion protein. *Mol Cell Biol.* 1999;19(10):6566-74.
387. Hassler M, Richmond TJ. The B-box dominates SAP-1-SRF interactions in the structure of the ternary complex. *EMBO J.* 2001;20(12):3018-28.
388. 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.
389. Szymczynska 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;275(37):28363-70.
390. Poon GM, Macgregor RB, Jr. Base coupling in sequence-specific site recognition by the ETS domain of murine PU.1. *J Mol Biol.* 2003;328(4):805-19.
391. Brown TA, McKnight SL. Specificities of protein-protein and protein-DNA interaction of GABP alpha and two newly defined ets-related proteins. *Genes Dev.* 1992;6(12B):2502-12.
392. Nye JA, Petersen JM, Gunther CV, Jonsen MD, Graves BJ. Interaction of murine ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes Dev.* 1992;6(6):975-90.
393. 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;11(2):303-13.
394. Shore P, Sharrocks AD. The ETS-domain transcription factors Elk-1 and SAP-1 exhibit differential DNA binding specificities. *Nucleic Acids Res.* 1995;23(22):4698-706.
395. John S, Marais R, Child R, Light Y, Leonard WJ. Importance of low affinity Elf-1 sites in the regulation of lymphoid-specific inducible gene expression. *J Exp Med.* 1996;183(3):743-50.
396. Choi YS, Sinha S. Determination of the consensus DNA-binding sequence and a transcriptional activation domain for ESE-2. *Biochem J.* 2006;398(3):497-507.
397. Mao X, Miesfeldt S, Yang H, Leiden JM, Thompson CB. The FLI-1 and chimeric EWS-FLI-1 oncoproteins display similar DNA binding specificities. *J Biol Chem.* 1994;269(27):18216-22.
398. 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(13):2147-60.
399. 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;275(2):1216-25.
400. Hollenhorst PC, Chandler KJ, Poulsen RL, Johnson WE, Speck NA, Graves BJ. DNA specificity determinants associate with distinct transcription factor functions. *PLoS Genet.* 2009;5(12):e1000778.
401. 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;378(2):213-7.



402. 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;277(47):45529-36.
403. Bojovic BB, Hassell JA. The PEA3 Ets transcription factor comprises multiple domains that regulate transactivation and DNA binding. *J Biol Chem.* 2001;276(6):4509-21.
404. 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;89(19):8889-93.
405. Wasylyk C, Kerckaert JP, Wasylyk B. A novel modulator domain of Ets transcription factors. *Genes Dev.* 1992;6(6):965-74.
406. Lim F, Kraut N, Framptom J, Graf T. DNA binding by c-Ets-1, but not v-Ets, is repressed by an intramolecular mechanism. *EMBO J.* 1992;11(2):643-52.
407. Lionneton F, Lelievre E, Baillat D, Stehelin D, Soncin F. Characterization and functional analysis of the p42Ets-1 variant of the mouse Ets-1 transcription factor. *Oncogene.* 2003;22(57):9156-64.
408. Jonsen MD, Petersen JM, Xu QP, Graves BJ. Characterization of the cooperative function of inhibitory sequences in Ets-1. *Mol Cell Biol.* 1996;16(5):2065-73.
409. Petersen JM, Skalicky JJ, Donaldson LW, McIntosh LP, Alber T, Graves BJ. Modulation of transcription factor Ets-1 DNA binding: DNA-induced unfolding of an alpha helix. *Science.* 1995;269(5232):1866-9.
410. 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;18(6):1609-20.
411. Goetz TL, Gu TL, Speck NA, Graves BJ. Auto-inhibition of Ets-1 is counteracted by DNA binding cooperativity with core-binding factor alpha2. *Mol Cell Biol.* 2000;20(1):81-90.
412. Fitzsimmons D, Hodsdon W, Wheat W, Maira SM, Wasylyk B, Hagman J. Pax-5 (BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. *Genes Dev.* 1996;10(17):2198-211.
413. Cowley DO, Graves BJ. Phosphorylation represses Ets-1 DNA binding by reinforcing autoinhibition. *Genes Dev.* 2000;14(3):366-76.
414. Yang SH, Shore P, Willingham N, Lakey JH, Sharrocks AD. The mechanism of phosphorylation-inducible activation of the ETS-domain transcription factor Elk-1. *EMBO J.* 1999;18(20):5666-74.
415. Hipskind RA, Rao VN, Mueller CG, Reddy ES, Nordheim A. Ets-related protein Elk-1 is homologous to the c-fos regulatory factor p62TCF. *Nature.* 1991;354(6354):531-4.
416. Dalton S, Treisman R. Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. *Cell.* 1992;68(3):597-612.
417. Yates PR, Atherton GT, Deed RW, Norton JD, Sharrocks AD. Id helix-loop-helix proteins inhibit nucleoprotein complex formation by the TCF ETS-domain transcription factors. *EMBO J.* 1999;18(4):968-76.
418. Li R, Pei H, Watson DK. Regulation of Ets function by protein - protein interactions. *Oncogene.* 2000;19(55):6514-23.
419. 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(4):977-91.
420. 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;7(4):292-7.
421. 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;18(7):1495-501.
422. Gille H, Sharrocks AD, Shaw PE. Phosphorylation of transcription factor p62TCF by MAP kinase stimulates ternary complex formation at c-fos promoter. *Nature.* 1992;358(6385):414-7.
423. Marais R, Wynne J, Treisman R. The SRF accessory protein Elk-1 contains a growth factor-regulated transcriptional activation domain. *Cell.* 1993;73(2):381-93.
424. Whitmarsh AJ, Shore P, Sharrocks AD, Davis RJ. Integration of MAP kinase signal transduction pathways at the serum response element. *Science.* 1995;269(5222):403-7.
425. Yang SH, Sharrocks AD, Whitmarsh AJ. Transcriptional regulation by the MAP kinase signaling cascades. *Gene.* 2003;320:3-21.
426. Eisenbeis CF, Singh H, Storb U. Pip, a novel IRF family member, is a lymphoid-specific, PU.1-dependent transcriptional activator. *Genes Dev.* 1995;9(11):1377-87.
427. Farnham PJ. Insights from genomic profiling of transcription factors. *Nat Rev Genet.* 2009;10(9):605-16.
428. 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;21(15):1882-94.
429. 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;19(11):1963-73.
430. Kar A, Gutierrez-Hartmann A. Molecular mechanisms of ETS transcription factor-mediated tumorigenesis. *Crit Rev Biochem Mol Biol.* 2013;48(6):522-43.
431. 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.
432. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature.* 1992;359(6391):162-5.
433. Sorensen PH, Lessnick SL, Lopez-Terrada D, Liu XF, Triche TJ, Denny CT. A second Ewing's sarcoma translocation, t(21;22), fuses the EWS gene to another ETS-family transcription factor, ERG. *Nat Genet.* 1994;6(2):146-51.



434. Tognon C, Knezevich SR, Huntsman D, Roskelley CD, Melnyk N, Mathers JA, et al. Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. *Cancer Cell*. 2002;2(5):367-76.
435. 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(5748):644-8.
436. Lin B, Ferguson C, White JT, Wang S, Vessella R, True LD, et al. Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. *Cancer Res*. 1999;59(17):4180-4.
437. Tomlins SA, Mehra R, Rhodes DR, Smith LR, Roulston D, Helgeson BE, et al. TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. *Cancer Res*. 2006;66(7):3396-400.
438. Helgeson BE, Tomlins SA, Shah N, Laxman B, Cao Q, Prensner JR, et al. Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. *Cancer Res*. 2008;68(1):73-80.
439. Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS, et al. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature*. 2007;448(7153):595-9.
440. Carver BS, Tran J, Gopalan A, Chen Z, Shaikh S, Carracedo A, et al. Aberrant ERG expression cooperates with loss of PTEN to promote cancer progression in the prostate. *Nat Genet*. 2009;41(5):619-24.
441. King JC, Xu J, Wongvipat J, Hieronymus H, Carver BS, Leung DH, et al. Cooperativity of TMPRSS2-ERG with PI3-kinase pathway activation in prostate oncogenesis. *Nat Genet*. 2009;41(5):524-6.
442. Chen Y, Chi P, Rockowitz S, Iaquinta PJ, Shamu T, Shukla S, et al. ETS factors reprogram the androgen receptor cistrome and prime prostate tumorigenesis in response to PTEN loss. *Nat Med*. 2013;19(8):1023-9.
443. Bastian BC, LeBoit PE, Hamm H, Brocker EB, Pinkel D. Chromosomal gains and losses in primary cutaneous melanomas detected by comparative genomic hybridization. *Cancer Res*. 1998;58(10):2170-5.
444. Jane-Valbuena J, Widlund HR, Perner S, Johnson LA, Dibner AC, Lin WM, et al. An oncogenic role for ETV1 in melanoma. *Cancer Res*. 2010;70(5):2075-84.
445. Mesquita B, Lopes P, Rodrigues A, Pereira D, Afonso M, Leal C, et al. Frequent copy number gains at 1q21 and 1q32 are associated with overexpression of the ETS transcription factors ETV3 and ELF3 in breast cancer irrespective of molecular subtypes. *Breast Cancer Res Treat*. 2013;138(1):37-45.
446. Romana SP, Poirel H, Leconiat M, Flexor MA, Mauchauffe M, Jonveaux P, et al. High frequency of t(12;21) in childhood B-lineage acute lymphoblastic leukemia. *Blood*. 1995;86(11):4263-9.
447. Romana SP, Mauchauffe M, Le Coniat M, Chumakov I, Le Paslier D, Berger R, et al. The t(12;21) of acute lymphoblastic leukemia results in a tel-AML1 gene fusion. *Blood*. 1995;85(12):3662-70.
448. Peeters P, Wlodarska I, Baens M, Criel A, Selleslag D, Hagemeijer A, et al. Fusion of ETV6 to MDS1/EVI1 as a result of t(3;12)(q26;p13) in myeloproliferative disorders. *Cancer Res*. 1997;57(4):564-9.
449. Golub TR, Barker GF, Lovett M, Gilliland DG. Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. *Cell*. 1994;77(2):307-16.
450. Eguchi M, Eguchi-Ishimae M, Tojo A, Morishita K, Suzuki K, Sato Y, et al. Fusion of ETV6 to neurotrophin-3 receptor TRKC in acute myeloid leukemia with t(12;15)(p13;q25). *Blood*. 1999;93(4):1355-63.
451. Golub TR, Goga A, Barker GF, Afar DE, McLaughlin J, Bohlander SK, et al. Oligomerization of the ABL tyrosine kinase by the Ets protein TEL in human leukemia. *Mol Cell Biol*. 1996;16(8):4107-16.
452. Schwaller J, Parganas E, Wang D, Cain D, Aster JC, Williams IR, et al. Stat5 is essential for the myelo- and lymphoproliferative disease induced by TEL/JAK2. *Mol Cell*. 2000;6(3):693-704.
453. 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(8):1274-81.
454. 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(13):2233-41.
455. 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;43(9):830-7.
456. 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.
457. Horn S, Figl A, Rachakonda PS, Fischer C, Sucker A, Gast A, et al. TERT promoter mutations in familial and sporadic melanoma. *Science*. 2013;339(6122):959-61.
458. Huang FW, Hodis E, Xu MJ, Kryukov GV, Chin L, Garraway LA. Highly recurrent TERT promoter mutations in human melanoma. *Science*. 2013;339(6122):957-9.
459. Muller PA, Vousden KH. Mutant p53 in cancer: new functions and therapeutic opportunities. *Cancer Cell*. 2014;25(3):304-17.
460. Do PM, Varanasi L, Fan S, Li C, Kubacka I, Newman V, et al. Mutant p53 cooperates with ETS2 to promote etoposide resistance. *Genes Dev*. 2012;26(8):830-45.
461. Zhu J, Sammons MA, Donahue G, Dou Z, Vedadi M, Getlik M, et al. Gain-of-function p53 mutants co-opt chromatin pathways to drive cancer growth. *Nature*. 2015;525(7568):206-11.
462. Chi P, Chen Y, Zhang L, Guo X, Wongvipat J, Shamu T, et al. ETV1 is a lineage survival factor that cooperates with KIT in gastrointestinal stromal tumours. *Nature*. 2010;467(7317):849-53.
463. Ran L, Sirota I, Cao Z, Murphy D, Chen Y, Shukla S, et al. Combined inhibition of MAP kinase and KIT signaling synergistically destabilizes ETV1 and suppresses GIST tumor growth. *Cancer Discov*. 2015;5(3):304-15.
464. Hayashi Y, Bardsley MR, Toyomasu Y, Milosavljevic S, Gajdos GB, Choi KM, et al. Platelet-Derived Growth Factor Receptor-alpha Regulates Proliferation of Gastrointestinal Stromal Tumor Cells With Mutations in KIT by Stabilizing ETV1. *Gastroenterology*. 2015;149(2):420-32 e16.
465. Vitari AC, Leong KG, Newton K, Yee C, O'Rourke K, Liu J, et al. COP1 is a tumour suppressor that causes degradation of ETS transcription factors. *Nature*. 2011;474(7351):403-6.
466. Lu G, Zhang Q, Huang Y, Song J, Tomaino R, Ehrenberger T, et al. Phosphorylation of ETS1 by Src family kinases prevents its recognition by the COP1 tumor suppressor. *Cancer Cell*. 2014;26(2):222-34.





467. Moreau-Gachelin F, Tavitian A, Tambourin P. Spi-1 is a putative oncogene in virally induced murine erythroleukaemias. *Nature*. 1988;331(6153):277-80.
468. Moreau-Gachelin F, Ray D, Mattei MG, Tambourin P, Tavitian A. The putative oncogene Spi-1: murine chromosomal localization and transcriptional activation in murine acute erythroleukemias. *Oncogene*. 1989;4(12):1449-56.
469. Klemsz MJ, McKercher SR, Celada A, Van Beveren C, Maki RA. The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. *Cell*. 1990;61(1):113-24.
470. Nguyen VC, Ray D, Gross MS, de Tand MF, Frezal J, Moreau-Gachelin F. Localization of the human oncogene SPI1 on chromosome 11, region p11.22. *Hum Genet*. 1990;84(6):542-6.
471. Behre G, Whitmarsh AJ, Coghlan MP, Hoang T, Carpenter CL, Zhang DE, et al. c-Jun is a JNK-independent coactivator of the PU.1 transcription factor. *J Biol Chem*. 1999;274(8):4939-46.
472. Rekhtman N, Radparvar F, Evans T, Skoultschi AI. Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. *Genes Dev*. 1999;13(11):1398-411.
473. Nerlov C, Querfurth E, Kulesha H, Graf T. GATA-1 interacts with the myeloid PU.1 transcription factor and represses PU.1-dependent transcription. *Blood*. 2000;95(8):2543-51.
474. Hagemeyer C, Bannister AJ, Cook A, Kouzarides T. The activation domain of transcription factor PU.1 binds the retinoblastoma (RB) protein and the transcription factor TFIID in vitro: RB shows sequence similarity to TFIID and TFIIB. *Proc Natl Acad Sci U S A*. 1993;90(4):1580-4.
475. Nagulapalli S, Pongubala JM, Atchison ML. Multiple proteins physically interact with PU.1. Transcriptional synergy with NF-IL6 beta (C/EBP delta, CRP3). *J Immunol*. 1995;155(9):4330-8.
476. Pongubala JM, Van Beveren C, Nagulapalli S, Klemsz MJ, McKercher SR, Maki RA, et al. Effect of PU.1 phosphorylation on interaction with NF-EM5 and transcriptional activation. *Science*. 1993;259(5101):1622-5.
477. Goebel MK. The PU.1 transcription factor is the product of the putative oncogene Spi-1. *Cell*. 1990;61(7):1165-6.
478. Hromas R, Orazi A, Neiman RS, Maki R, Van Beveren C, Moore J, et al. Hematopoietic lineage- and stage-restricted expression of the ETS oncogene family member PU.1. *Blood*. 1993;82(10):2998-3004.
479. Galson DL, Hensold JO, Bishop TR, Schalling M, D'Andrea AD, Jones C, et al. Mouse beta-globin DNA-binding protein B1 is identical to a proto-oncogene, the transcription factor Spi-1/PU.1, and is restricted in expression to hematopoietic cells and the testis. *Mol Cell Biol*. 1993;13(5):2929-41.
480. Ray D, Bosselut R, Ghysdael J, Mattei MG, Tavitian A, Moreau-Gachelin F. Characterization of Spi-B, a transcription factor related to the putative oncoprotein Spi-1/PU.1. *Mol Cell Biol*. 1992;12(10):4297-304.
481. 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;14(1):373-81.
482. Hohaus S, Petrovick MS, Voso MT, Sun Z, Zhang DE, Tenen DG. PU.1 (Spi-1) and C/EBP alpha regulate expression of the granulocyte-macrophage colony-stimulating factor receptor alpha gene. *Mol Cell Biol*. 1995;15(10):5830-45.
483. Perez C, Coeffier E, Moreau-Gachelin F, Wietzerbin J, Benech PD. Involvement of the transcription factor PU.1/Spi-1 in myeloid cell-restricted expression of an interferon-inducible gene encoding the human high-affinity Fc gamma receptor. *Mol Cell Biol*. 1994;14(8):5023-31.
484. Feinman R, Qiu WQ, Pearse RN, Nikolajczyk BS, Sen R, Sheffery M, et al. PU.1 and an HLH family member contribute to the myeloid-specific transcription of the Fc gamma RIIIA promoter. *EMBO J*. 1994;13(16):3852-60.
485. Pahl HL, Rosmarin AG, Tenen DG. Characterization of the myeloid-specific CD11b promoter. *Blood*. 1992;79(4):865-70.
486. Rosmarin AG, Levy R, Tenen DG. Cloning and analysis of the CD18 promoter. *Blood*. 1992;79(10):2598-604.
487. 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;106(5):1590-600.
488. 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(6):624-30.
489. 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;201(9):1487-502.
490. Fisher RC, Lovelock JD, Scott EW. A critical role for PU.1 in homing and long-term engraftment by hematopoietic stem cells in the bone marrow. *Blood*. 1999;94(4):1283-90.
491. Back J, Dierich A, Bronn C, Kastner P, Chan S. PU.1 determines the self-renewal capacity of erythroid progenitor cells. *Blood*. 2004;103(10):3615-23.
492. Back J, Allman D, Chan S, Kastner P. Visualizing PU.1 activity during hematopoiesis. *Exp Hematol*. 2005;33(4):395-402.
493. Arinobu Y, Mizuno S, Chong Y, Shigematsu H, Iino T, Iwasaki H, et al. Reciprocal activation of GATA-1 and PU.1 marks initial specification of hematopoietic stem cells into myeloerythroid and myelolymphoid lineages. *Cell Stem Cell*. 2007;1(4):416-27.
494. 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;31(4):576-86.
495. Kamath MB, Houston IB, Janovski AJ, Zhu X, Gowrisankar S, Jegga AG, et al. Dose-dependent repression of T-cell and natural killer cell genes by PU.1 enforces myeloid and B-cell identity. *Leukemia*. 2008;22(6):1214-25.
496. DeKoter RP, Singh H. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science*. 2000;288(5470):1439-41.
497. Chen H, Zhang P, Radomska HS, Hetherington CJ, Zhang DE, Tenen DG. Octamer binding factors and their coactivator can activate the murine PU.1 (spi-1) promoter. *J Biol Chem*. 1996;271(26):15743-52.



498. Eklund EA, Kakar R. Recruitment of CREB-binding protein by PU.1, IFN-regulatory factor-1, and the IFN consensus sequence-binding protein is necessary for IFN-gamma-induced p67phox and gp91phox expression. *J Immunol.* 1999;163(11):6095-105.
499. Li Y, Okuno Y, Zhang P, Radomska HS, Chen H, Iwasaki H, et al. Regulation of the PU.1 gene by distal elements. *Blood.* 2001;98(10):2958-65.
500. Steidl U, Rosenbauer F, Verhaak RG, 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;38(11):1269-77.
501. Bai Y, Srinivasan L, Perkins L, Atchison ML. Protein acetylation regulates both PU.1 transactivation and Ig kappa 3' enhancer activity. *J Immunol.* 2005;175(8):5160-9.
502. Kihara-Negishi F, Yamamoto H, Suzuki M, Yamada T, Sakurai T, Tamura T, et al. In vivo complex formation of PU.1 with HDAC1 associated with PU.1-mediated transcriptional repression. *Oncogene.* 2001;20(42):6039-47.
503. Suzuki M, Yamada T, Kihara-Negishi F, Sakurai T, Hara E, Tenen DG, et al. Site-specific DNA methylation by a complex of PU.1 and Dnmt3a/b. *Oncogene.* 2006;25(17):2477-88.
504. Imoto A, Okada M, Okazaki T, Kitasato H, Harigae H, Takahashi S. Metallothionein-1 isoforms and vimentin are direct PU.1 downstream target genes in leukemia cells. *J Biol Chem.* 2010;285(14):10300-9.
505. Lio CW, Zhang J, Gonzalez-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;5.
506. Stephens DC, Poon GM. Differential sensitivity to methylated DNA by ETS-family transcription factors is intrinsically encoded in their DNA-binding domains. *Nucleic Acids Res.* 2016;44(18):8671-81.
507. Li X, Vradii D, Gutierrez S, Lian JB, van Wijnen AJ, Stein JL, et al. Subnuclear targeting of Runx1 is required for synergistic activation of the myeloid specific M-CSF receptor promoter by PU.1. *J Cell Biochem.* 2005;96(4):795-809.
508. Huang G, Zhang P, Hirai H, Elf S, Yan X, Chen Z, et al. PU.1 is a major downstream target of AML1 (RUNX1) in adult mouse hematopoiesis. *Nat Genet.* 2008;40(1):51-60.
509. Leddin M, Perrod C, Hoogenkamp M, Ghani S, Assi S, Heinz S, et al. Two distinct auto-regulatory loops operate at the PU.1 locus in B cells and myeloid cells. *Blood.* 2011;117(10):2827-38.
510. Laiosa CV, Stadtfeld M, Xie H, de Andres-Aguayo L, Graf T. Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBP alpha and PU.1 transcription factors. *Immunity.* 2006;25(5):731-44.
511. Feng R, Desbordes SC, Xie H, Tillo ES, Pixley F, Stanley ER, et al. PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells. *Proc Natl Acad Sci U S A.* 2008;105(16):6057-62.
512. Traver D, Miyamoto T, Christensen J, Iwasaki-Arai J, Akashi K, Weissman IL. Fetal liver myelopoiesis occurs through distinct, prospectively isolatable progenitor subsets. *Blood.* 2001;98(3):627-35.
513. Zhang P, Iwasaki-Arai J, Iwasaki H, Fenyus ML, Dayaram T, Owens BM, et al. Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity.* 2004;21(6):853-63.
514. Huber R, Pietsch D, Panterodt T, Brand K. Regulation of C/EBPbeta and resulting functions in cells of the monocytic lineage. *Cell Signal.* 2012;24(6):1287-96.
515. Dahl R, Iyer SR, Owens KS, Cuylear DD, Simon MC. The transcriptional repressor GFI-1 antagonizes PU.1 activity through protein-protein interaction. *J Biol Chem.* 2007;282(9):6473-83.
516. Zhang P, Behre G, Pan J, Iwama A, Wara-Aswapati N, Radomska HS, et al. Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. *Proc Natl Acad Sci U S A.* 1999;96(15):8705-10.
517. Cantor AB, Orkin SH. Transcriptional regulation of erythropoiesis: an affair involving multiple partners. *Oncogene.* 2002;21(21):3368-76.
518. Burda P, Laslo P, Stopka T. The role of PU.1 and GATA-1 transcription factors during normal and leukemogenic hematopoiesis. *Leukemia.* 2010;24(7):1249-57.
519. Zhang P, Zhang X, Iwama A, Yu C, Smith KA, Mueller BU, et al. PU.1 inhibits GATA-1 function and erythroid differentiation by blocking GATA-1 DNA binding. *Blood.* 2000;96(8):2641-8.
520. 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;16(2):285-96.
521. 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;38(1):27-37.
522. Laslo P, Spooner CJ, Warmflash A, Lancki DW, Lee HJ, Sciammas R, et al. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell.* 2006;126(4):755-66.
523. 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;9(8):927-36.
524. Dias S, Mansson R, Gurbuxani S, Sigvardsson M, Kee BL. E2A proteins promote development of lymphoid-primed multipotent progenitors. *Immunity.* 2008;29(2):217-27.
525. Greig KT, de Graaf CA, Murphy JM, Carpinelli MR, Pang SH, Frampton J, et al. Critical roles for c-Myb in lymphoid priming and early B-cell development. *Blood.* 2010;115(14):2796-805.
526. 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;10(3):289-96.
527. Carotta S, Dakic A, D'Amico A, Pang SH, 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;32(5):628-41.
528. Scott EW, Simon MC, Anastasi J, Singh H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science.* 1994;265(5178):1573-7.
529. 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.



530. Sokalski KM, Li SK, Welch I, Cadieux-Pitre HA, 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;118(10):2801-8.
531. 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.
532. 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;10(4):399-408.
533. Pang SH, 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.
534. Lu R, Medina KL, Lancki DW, Singh H. IRF-4,8 orchestrate the pre-B-to-B transition in lymphocyte development. *Genes Dev*. 2003;17(14):1703-8.
535. Wang H, Lee CH, Qi C, Tailor P, Feng J, Abbasi S, et al. IRF8 regulates B-cell lineage specification, commitment, and differentiation. *Blood*. 2008;112(10):4028-38.
536. Minnich M, Tagoh H, Bonelt P, Axelsson E, Fischer M, Cebolla B, et al. Multifunctional role of the transcription factor Blimp-1 in coordinating plasma cell differentiation. *Nat Immunol*. 2016;17(3):331-43.
537. Tellier J, Shi W, Minnich M, Liao Y, Crawford S, Smyth GK, et al. Blimp-1 controls plasma cell function through the regulation of immunoglobulin secretion and the unfolded protein response. *Nat Immunol*. 2016;17(3):323-30.
538. Eis PS, Tam W, Sun L, Chadburn A, Li Z, Gomez MF, et al. Accumulation of miR-155 and BIC RNA in human B cell lymphomas. *Proc Natl Acad Sci U S A*. 2005;102(10):3627-32.
539. O'Connell RM, Rao DS, Chaudhuri AA, Boldin MP, Taganov KD, Nicoll J, et al. Sustained expression of microRNA-155 in hematopoietic stem cells causes a myeloproliferative disorder. *J Exp Med*. 2008;205(3):585-94.
540. 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(12):3437-44.
541. 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(3):998-1007.
542. 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;100(13):4680-1.
543. Dohner K, Tobis K, Bischof T, Hein S, Schlenk RF, Frohling 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;102(10):3850; author reply -1.
544. Lavalley 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(9):1030-7.
545. Recaladin T, Fear DJ. Transcription factors regulating B cell fate in the germinal centre. *Clinical and experimental immunology*. 2016;183(1):65-75.
546. Garrett-Sinha LA. Review of Ets1 structure, function, and roles in immunity. *Cell Mol Life Sci*. 2013;70(18):3375-90.
547. Bories JC, Willerford DM, Grevin 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;377(6550):635-8.
548. Nutt SL, Kee BL. The transcriptional regulation of B cell lineage commitment. *Immunity*. 2007;26(6):715-25.
549. 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(2):110-5.
550. 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(8):1513-35.
551. 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. *Immunity*. 2006;210:67-85.
552. 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(10):4220-6.
553. Calvo J, BenYousef A, Baijer J, Rouyez MC, Pflumio F. Assessment of human multi-potent hematopoietic stem/progenitor cell potential using a single in vitro screening system. *PLoS One*. 2012;7(11):e50495.
554. 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(1):25-38.
555. Ridinger-Saison M, Boeva V, Rimmel P, Kulakovskiy I, Gallais I, Levavasseur B, et al. Spi-1/PU.1 activates transcription through clustered DNA occupancy in erythroleukemia. *Nucleic Acids Res*. 2012;40(18):8927-41.
556. Ernst J, Kellis M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat Protoc*. 2017;12(12):2478-92.
557. David MD, Petit D, Bertoglio J. The RhoGAP ARHGAP19 controls cytokinesis and chromosome segregation in T lymphocytes. *J Cell Sci*. 2014;127(Pt 2):400-10.
558. Kasar S, Kim J, Improgo R, Tiao G, Polak P, Haradhvala N, et al. Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. *Nat Commun*. 2015;6:8866.
559. 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(4):228-35.
560. 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(3):413-26.



561. 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(31):11365-70.
562. Kodandapani R, Pio F, Ni CZ, Piccialli G, Klemsz M, McKercher S, et al. A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. *Nature*. 1996;380(6573):456-60.
563. 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.
564. Foulds CE, Nelson ML, Blaszczyk AG, Graves BJ. Ras/mitogen-activated protein kinase signaling activates Ets-1 and Ets-2 by CBP/p300 recruitment. *Mol Cell Biol*. 2004;24(24):10954-64.
565. 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(11):3729-43.
566. de la Rica L, Rodriguez-Ubreva J, Garcia M, Islam AB, 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.
567. 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(7574):525-30.
568. 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(1):108-25.
569. Zhang XK, Moussa O, LaRue A, Bradshaw S, Molano I, Spyropoulos DD, et al. The transcription factor Fli-1 modulates marginal zone and follicular B cell development in mice. *J Immunol*. 2008;181(3):1644-54.
570. Xue HH, Bollenbacher-Reilley 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(4):421-31.
571. 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(12):7374-84.
572. Walker BA, Boyle EM, Wardell CP, Murison A, Begum DB, Dahir NM, et al. Mutational Spectrum, Copy Number Changes, and Outcome: Results of a Sequencing Study of Patients With Newly Diagnosed Myeloma. *J Clin Oncol*. 2015;33(33):3911-20.
573. Lu G, Middleton RE, Sun H, Naniong M, Ott CJ, Mitsiades CS, et al. The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. *Science*. 2014;343(6168):305-9.
574. Hardy RR, Carmack CE, Shinton SA, Kemp JD, Hayakawa K. Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *J Exp Med*. 1991;173(5):1213-25.
575. Heng TS, Painter MW, Immunological Genome Project C. The Immunological Genome Project: networks of gene expression in immune cells. *Nat Immunol*. 2008;9(10):1091-4.
576. de Miranda NF, Georgiou K, Chen L, Wu C, Gao Z, Zaravinos A, et al. Exome sequencing reveals novel mutation targets in diffuse large B-cell lymphomas derived from Chinese patients. *Blood*. 2014;124(16):2544-53.
577. White BS, Lanc I, O'Neal J, Gupta H, Fulton RS, Schmidt H, et al. A multiple myeloma-specific capture sequencing platform discovers novel translocations and frequent, risk-associated point mutations in IGLL5. *Blood Cancer J*. 2018;8(3):35.
578. 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.
579. 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;49(5):934-46.





**Titre :** Anomalies moléculaires dans la macroglobulinémie de Waldenström : identification d'une mutation somatique récurrente dans le gène codant pour le facteur de transcription SPI1/PU.1 et description de ses conséquences fonctionnelles

**Mots clés :** Waldenström, NGS, SPI1, PU.1

**Résumé :** Les facteurs de transcription (FT) ETS sont divisés en sous-familles en fonction de leurs similitudes en matière de séquence protéique, de séquences de liaison à l'ADN et d'interactions avec différents cofacteurs. Ils sont régulés par des signaux extracellulaires et contribuent à divers processus cellulaires, dont la prolifération cellulaire et la transformation tumorale. Les gènes de la famille *ETS* sont fréquemment ciblés par des processus oncogéniques que ce soit des translocations chromosomiques ou des gains du nombre de leurs copies. Le gène *PU.1/SPI1* est également ciblé par des mutations ponctuelles inactivatrices dans les hémopathies myéloïdes humaines. Nous avons étudié une mutation somatique récurrente du gène *PU.1/SPI1* (c.676C>G, p.Q226E), identifiée chez environ 6% des patients atteints d'une macroglobulinémie de Waldenström (MW). La mutation modifie les caractéristiques de liaison à l'ADN de la protéine mutante, passant des séquences classiques reconnues par SPI1 à des séquences reconnues par d'autres protéines ETS comme ETS1,

et d'une liaison à des régions enhancer 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. Les conséquences fonctionnelles de cette mutation sont une augmentation de la prolifération cellulaire et une diminution de la différenciation lymphoïde B terminale dans une lignée cellulaire modèle et des échantillons primaires de MW. Nous décrivons ici un mécanisme de subversion oncogénique de la fonction d'un FT suite à la modification subtile de la spécificité de liaison à l'ADN de la protéine mutante, menant à un arrêt de différenciation. La démonstration qu'une mutation somatique ponctuelle peut modifier l'équilibre de liaison d'un facteur de transcription à l'échelle du génome fournit un paradigme mécanistique sur la façon dont les mutations faux sens dans les gènes codant pour des FT pourraient être oncogéniques dans les tumeurs humaines.

**Title :** A recurrent activating missense mutation in Waldenström macroglobulinemia affects the DNA binding sequence of the ETS transcription factor SPI1 and enhances cellular proliferation

**Keywords :** Waldenström, NGS, SPI1, PU.1

**Abstract:** The ETS-domain transcription factors are divided into subfamilies based on protein similarities, DNA binding sequences and interaction with cofactors. They are regulated by extracellular clues and contribute to a variety of cellular processes, including proliferation and transformation. ETS genes are targeted by oncogenic processes through chromosomal translocations and copy number gains. The *PU.1/SPI1* gene is also targeted by inactivating point mutations in human myeloid malignancies. We investigated a recurrent somatic missense mutation (Q226E) of the *PU.1/SPI1* gene in Waldenström macroglobulinemia, a human B-cell lymphoproliferative disorder. The mutation changes DNA binding of the mutant protein from classical

*SPI1* to ETS1-like sequences, shifting the balance from binding to promoter regions from enhancers. Increased binding by mutant *SPI1* at promoters activates gene expression of intracellular signaling pathways typically promoted by other ETS factor family members. The functional consequences are decreased terminal B-cell differentiation in a model cell line and primary samples. In summary, we describe oncogenic subversion of transcription factor function through subtle alteration DNA binding specificity leading to differentiation arrest. The demonstration that a somatic point mutation subtly changes the balance of genome binding provides a mechanistic paradigm for how missense mutations in transcription factor genes may be oncogenic in human tumors.

