



**HAL**  
open science

# Impact des altérations génétiques de PAX5 sur le développement de la lignée lymphoïde B et dans la leucémogénèse des LAL-B

Laura Jamrog

► **To cite this version:**

Laura Jamrog. Impact des altérations génétiques de PAX5 sur le développement de la lignée lymphoïde B et dans la leucémogénèse des LAL-B. Cancer. Université Paul Sabatier - Toulouse III, 2021. Français. NNT : 2021TOU30306 . tel-04202700

**HAL Id: tel-04202700**

**<https://theses.hal.science/tel-04202700>**

Submitted on 11 Sep 2023

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

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



# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

---

**Présentée et soutenue par :**

**Laura JAMROG**

**le** Mercredi 31 Mars 2021

**Titre :**

Impact des altérations génétiques de PAX5 sur le développement de la lignée lymphoïde B et dans la leucémogénèse des LAL-B

---

**École doctorale et discipline ou spécialité :**

ED BSB : Cancérologie

**Unité de recherche :**

Centre de Recherche en Cancérologie de Toulouse (CRCT- U1037)

**Directeur/trice(s) de Thèse :**

Cyril BROCCARDO & Marlène PASQUET

**Jury :**

M. Eric DELABESSE, Président du jury  
Mme Michèle GOODHARDT, Rapporteur  
M. Michel AURRAND-LIONS, Rapporteur  
Mme Françoise PFLUMIO, Examinatrice  
M. Thomas MERCHER, Examineur  
M. Cyril BROCCARDO, Directeur de thèse  
Mme Marlène PASQUET, Co-directrice de thèse









## Remerciements

Je remercie la Région Occitanie et la Société Française de lutte contre les Cancers et les leucémies de l'Enfant et de l'adolescent (SFCE) pour avoir financé ces années de thèse.

J'adresse mes sincères remerciements aux membres de mon jury d'avoir accepté d'évaluer mes travaux de thèse : Michèle Goodhardt et Michel Aurrand-Lions en qualité de rapporteurs, Françoise Pflumio et Thomas Mercher en qualité d'examineurs. Merci au Professeur Eric Delabesse d'avoir accepté d'être président de mon jury de thèse.

Je remercie les membres de mon comité de suivi de thèse, Ahmed Amine Khamlichi et Pierre Lutz, pour m'avoir encouragée et aiguillée dans la bonne direction, et permis de prendre du recul pour évaluer mon travail.

Je tiens à exprimer toute ma reconnaissance à Eric Delabesse et Cyril Broccardo qui m'ont si bien accueillie dans leur équipe : merci de m'avoir donné ma chance, de m'avoir si bien accompagnée et valorisée tout au long de mon expérience doctorale, et au-delà.

Je remercie grandement mon directeur de thèse Cyril Broccardo de m'avoir encadrée et formée avec toute la bienveillance qui le caractérise tant. Merci pour ton partage, ta patience et ta compréhension face à mon besoin d'autonomie, merci pour ton respect, ton soutien permanent et surtout ta confiance sans faille durant toutes ces années ; tout cela m'a permis de progresser et de m'épanouir. Je suis fière de ce que nous avons réalisé ensemble.

Je remercie ma co-directrice de thèse Marlène Pasquet pour son entrain communicatif, ses encouragements et la confiance qu'elle m'a accordée pour reprendre le flambeau GATA2<sup>^^</sup>. Je remercie chaleureusement Bastien Gerby pour tous ses conseils avisés et son soutien remarquable, merci d'avoir upgradé ma formation doctorale.

Je remercie vivement tous mes collègues et collaborateurs qui m'ont permis de mener à bien tous nos projets de recherche. Travailler avec vous tous fut un plaisir. J'adresse toute ma gratitude à Sylvie Hébrard, une personne en or avec qui j'ai fait mes premiers pas dans cette équipe et qui n'a jamais cessé de m'épauler ; à Stéphanie et Naïs (les Supers Nanas) pour qui rien n'est impossible ; à Manon, Camille et Alicia qui m'ont toujours fait avancer dans la bonne humeur. Je souhaite également remercier nos stagiaires qui ont apporté leur pierre à l'édifice et à qui j'ai adoré transmettre autant que faire se peut.

J'adresse un immense merci à mes co-doctorants et amis Vincent et Laetitia qui m'ont boostée et ont propulsé mon expérience doctorale dans une autre dimension, épique !

Je remercie de tout mon cœur mes parents, ma sœur, ma grand-mère, toute ma famille et mes chers amis pour leur soutien d'exception. J'adresse mille mercis à tous mes amis pour tous les mémorables moments de partage et de détente, qui ont édulcoré mon parcours.

Petite dédicace à mon filleul qui est né pendant la rédaction laborieuse de cette thèse.

Merci à tous, je mesure ma chance d'avoir pu être si bien entourée et je vous suis infiniment reconnaissante d'avoir fait de mon parcours doctoral une expérience grandiose et tellement enrichissante.

# Table des matières

Table des matières .....	5
RÉSUMÉ .....	9
ABSTRACT .....	11
Liste des abréviations .....	13
Partie I : Introduction .....	21
Chapitre 1 de l'introduction .....	21
1.1. Les gènes de développement <i>Pax</i> .....	21
1.1.1. Évolution phylogénétique .....	21
1.1.2. Découverte des gènes <i>Pax</i> chez la drosophile.....	25
1.1.3. Classification des gènes <i>Pax</i> .....	25
1.2. Rôles majeurs des protéines Pax durant le développement .....	29
1.2.1. Groupe I : Pax1 & Pax9 .....	31
1.2.2. Groupe II : Pax2, Pax5 & Pax8.....	33
1.2.3. Groupe III : Pax3 & Pax7 .....	33
1.2.4. Groupe IV : Pax4 & Pax6 .....	35
1.3. Structure moléculaire et fonctions des domaines protéiques .....	37
1.3.1. Le signal de localisation nucléaire (NLS).....	37
1.3.2. Le domaine PAIRED, domaine de liaison à l'ADN .....	39
1.3.3. L'homéodomaine des homéoprotéines de la classe <i>paired</i> .....	43
1.3.4. Le motif octapeptide (OP).....	47
1.3.5. Domaines de transactivation (TAD) et d'inhibition (ID).....	49
1.3.6. Le motif OAR (Otp/Aristaless/Rax) .....	53
1.4. Organisation génomique et régulation génique de <i>PAX5</i> .....	55
1.4.1. Localisation et structure génique.....	55
1.4.2. Éléments de cis-régulation transcriptionnelle de <i>PAX5</i> .....	57
1.4.3. Variants issus de l'épissage alternatif .....	59
1.4.4. Éléments de trans-régulation transcriptionnelle de <i>PAX5</i> .....	61
Chapitre 2 de l'introduction .....	63
2.1. Généralités sur l'hématopoïèse.....	63
2.1.1. Hématopoïèse fœtale.....	63
2.1.2. Modèle actuel de l'hématopoïèse adulte.....	63
2.2. La lymphopoïèse B.....	67
2.2.1. Mécanismes de la lymphopoïèse B précoce dans la MO .....	67
2.2.1.1. La classification des étapes précoces du développement B .....	67
2.2.1.2. Mécanismes moléculaires de la lymphopoïèse B précoce .....	71

2.2.1.2.1.	La recombinaison V(D)J des loci des gènes d'immunoglobulines .....	71
2.2.1.2.2.	Formation et signalisation du pré-BCR.....	73
2.2.2.	Mécanismes de la lymphopoïèse B tardive hors de la MO.....	75
2.2.2.1.	Stades cellulaires tardifs.....	75
2.2.2.2.	Maturation et sélection des gènes des immunoglobulines.....	77
2.2.2.2.1.	Hypermutation somatique et commutation de classe des Ig.....	77
2.2.2.2.2.	Recombinaison-suicide du locus des immunoglobulines .....	79
2.2.3.	Voies de signalisation principales de la lymphopoïèse B précoce.....	79
2.2.3.1.	La signalisation de l'IL-7R .....	79
2.2.3.1.1.	L'axe de signalisation IL-7R-JAK-STAT .....	81
2.2.3.1.2.	La voie de signalisation PI3K.....	83
2.2.3.2.	La signalisation du pré-BCR .....	83
2.2.3.2.1.	Initiation de la signalisation du pré-BCR.....	83
2.2.3.2.2.	Transduction du signal et plateformes d'interaction .....	83
2.2.3.2.3.	La voie PI3K-AKT(PKB)-FOXO1 .....	85
2.2.3.2.4.	La voie RAS-RAF-MEK-ERK .....	87
2.2.3.3.	La signalisation du BCR .....	87
2.2.3.3.1.	Initiation de la signalisation.....	87
2.2.3.3.2.	Voies de signalisation du BCR.....	89
2.2.4.	Hierarchie transcriptionnelle orchestrant la différenciation B.....	89
2.2.4.1.	Réseaux transcriptionnels établissant l'identité cellulaire B.....	91
2.2.4.2.	Réseaux transcriptionnels de la différenciation terminale .....	97
Chapitre 3 de l'introduction .....		105
3.1.	Implication des protéines PAX dans la pathogenèse.....	107
3.1.1.	Groupe I : PAX1 & PAX9.....	111
3.1.2.	Groupe II : PAX2, PAX5 & PAX8 .....	111
3.1.3.	Implication de PAX5 dans l'oncogenèse.....	113
3.1.4.	Groupe III : PAX3 & PAX7.....	115
3.1.5.	Groupe IV : PAX4 & PAX6 .....	115
3.1.6.	Translocations chromosomiques des PAX associées aux cancers.....	117
3.2.	Les leucémies aiguës lymphoblastiques de la lignée B .....	121
3.2.1.	Généralités sur les LAL-B.....	121
3.2.1.1.	Incidences.....	121
3.2.1.2.	Caractéristiques biologiques et cliniques.....	123
3.2.1.3.	Prise en charge thérapeutique.....	123
3.2.1.4.	Taux de survie.....	127

3.2.2.	Bases moléculaires des LAL-B.....	129
3.2.3.	Altérations génétiques de <i>PAX5</i> dans les LAL-B.....	135
3.2.3.1.	Fréquence élevée et diversité des altérations de <i>PAX5</i> .....	135
3.2.3.1.1.	Diversité des mécanismes d'inactivation de <i>PAX5</i> .....	137
3.2.3.1.2.	Diversité des mutations de séquence de <i>PAX5</i> .....	139
3.2.3.1.3.	Impact fonctionnel des inactivations et mutations de <i>PAX5</i> .....	139
3.2.3.1.4.	Le hotspot mutationnel <i>PAX5</i> P80R.....	141
3.2.3.1.5.	Diversité des réarrangements chromosomiques de <i>PAX5</i> .....	141
3.2.3.2.	Différents degrés d'implication dans la leucémogénèse et exclusivité d'expression des altérations de <i>PAX5</i> .....	145
3.2.4.	Classification immunophénotypique des LAL-B.....	149
3.2.4.1.	Sous-types immunophénotypiques selon l'EGIL .....	149
3.2.4.2.	Corrélation immunophénotype-cytogénétique .....	151
3.2.5.	Pathogénèse des LAL-B.....	151
3.2.5.1.	Les cancers résultent de l'accumulation de mutations somatiques.....	151
3.2.5.2.	La dérégulation de la lymphopoïèse est la cause principale des LAL-B.....	153
3.2.5.3.	La leucémogénèse est un processus multi-étapes .....	153
3.2.5.4.	Les mécanismes de la niche favorisent la transformation maligne .....	159
3.2.5.4.1.	Des régulateurs extrinsèques de la survie des cellules leucémiques..	159
3.2.5.4.2.	Des mécanismes de chimiorésistance dépendant de la niche .....	159
3.2.5.4.3.	Le remodelage du microenvironnement immunitaire .....	159
3.2.6.	Mécanisme causal des LAL-B et facteurs de risque .....	161
3.2.6.1.	Susceptibilité génétique.....	161
3.2.6.2.	Rôle causal des infections courantes .....	163
3.2.6.3.	Le mode de vie.....	167
3.2.6.4.	Le facteur (mal)chance .....	167
Partie II : Objectifs et Résultats .....		169
Objectifs de la thèse.....		169
4.1.	Comparaison de l'expression et des fonctions des isoformes A et B de Pax5.....	169
4.2.	Caractérisation du modèle murin <i>in vivo</i> de LAL-B multi-étapes induite par l'oncoprotéine de fusion <i>PAX5-ELN</i> .....	170
4.3.	Étude fonctionnelle <i>in vitro</i> et <i>in vivo</i> de l'effet du mutant <i>PAX5</i> <sup>R38H</sup> .....	171
4.4.	Caractérisation du modèle murin <i>in vivo</i> exprimant <i>Gata2</i> <sup>R396Q/+</sup> .....	171
Publications originales des travaux de thèse collaboratifs.....		172
Cresson C, Péron S, Jamrog L <i>et al.</i> , <i>Oncotarget</i> , 2018.....		172
Jamrog L, Duchemin G <i>et al.</i> , <i>PNAS</i> , 2018.....		172
Duployez N, Jamrog LA <i>et al.</i> , <i>Blood</i> , 2021.....		172

Partie III : Conclusions & Discussion.....	255
Modèle murin de leucémogénèse B induite par l'oncoprotéine de fusion PAX5-ELN.....	255
Émergence de la composante héréditaire dans la prédisposition aux LAL-B.....	264
Références bibliographiques .....	271
Références supplémentaires (non répertoriées sur la base PubMed) :.....	324
Annexe : Attestation de suivi de l'enseignement doctoral.....	325

# RÉSUMÉ

Le gène *PAX5* (*Paired box 5*) code un facteur de transcription essentiel pour la différenciation lymphoïde B. Nous avons montré que les deux isoformes *PAX5A* et *PAX5B* étaient différentiellement régulées mais pouvaient exercer une fonction équivalente durant l'induction de la différenciation lymphoïde B et pourraient présenter des différences fonctionnelles à la suite de l'activation des lymphocytes B. Le contrôle précis de leur expression peut ainsi refléter un moyen d'ajuster finement le dosage de *PAX5* pendant le processus de différenciation des cellules de la lignée lymphoïde B.

*PAX5* est la cible principale d'une large diversité d'altérations somatiques dans les LAL-B de l'enfant et de l'adulte. Cependant, le rôle des protéines de fusion impliquant *PAX5* dans l'initiation et la transformation des LAL-B est encore méconnu. Nous avons précédemment décrit une nouvelle translocation chromosomique récurrente t(7;9)(q11;p13) dans les LAL-B qui juxtapose *PAX5* à la séquence codante du gène de l'élastine (*ELN*). Pour étudier la fonction de la protéine de fusion résultante, *PAX5-ELN*, au cours du développement leucémique, nous avons créé un modèle murin dans lequel le transgène *PAX5-ELN* est exprimé spécifiquement dans le compartiment B. Les souris exprimant *PAX5-ELN* développent un phénotype de LAL-B avec une pénétrance de 80%. Leur transformation leucémique est associée à l'acquisition de mutations secondaires récurrentes des gènes *Ptpn11*, *Kras*, *Pax5*, et *Jak3* affectant d'importantes voies de signalisation requises pour la prolifération cellulaire. Nos études fonctionnelles ont démontré que *PAX5-ELN* altérait *in vitro* et *in vivo* le développement lymphoïde B et pouvait induire une expansion aberrante du compartiment progéniteur B (pro-B) au stade préleucémique. Nos approches moléculaires et computationnelles ont identifié des gènes-candidats régulés par *PAX5-ELN* et pouvant être impliqués dans l'initiation leucémique. Nos données fournissent ainsi un nouveau modèle d'étude *in vivo* récapitulant la leucémogénèse multi-étapes des LAL-B décrites chez les patients et impliquent fortement les protéines de fusion engageant *PAX5* en tant que puissantes oncoprotéines dans le développement leucémique. Par ailleurs, il existe de plus en plus de preuves d'une base génétique héréditaire de prédisposition aux LAL-B pédiatriques. Dans ce contexte, quatre cas de familles non-apparentées affectées par des LAL-B et exprimant des mutations ponctuelles germinales et hétérozygotes de *PAX5* ont été récemment rapportés : la mutation *PAX5* G183S altérant le domaine octapeptide de *PAX5* a été décrite chez trois familles alors que la mutation *PAX5* R38H altérant le domaine de liaison à l'ADN de *PAX5* a été identifiée chez une autre. Nous avons renforcé l'hypothèse du caractère héréditaire des LAL-B familiales avec la description de trois nouveaux cas de LAL-B au sein d'une même famille exprimant la mutation germinale *PAX5* R38H. Pour étudier l'effet intrinsèque de la protéine mutée *PAX5* R38H dans le développement lymphoïde B, nous avons effectué des tests fonctionnels *in vitro* et *in vivo* combinés à une analyse d'expression génique, basés sur une approche de complémentation rétrovirale. Nos résultats ont indiqué que *PAX5* R38H agissait comme un fort variant hypomorphique qui échouait à induire la différenciation lymphoïde B et n'exerçait pas d'effet dominant-négatif sur la forme sauvage de *PAX5*. Des transplantations syngéniques de cellules exprimant *PAX5* R38H ont démontré qu'elles maintenaient une capacité de prise de greffe et menaient au développement leucémique chez la souris. Notre analyse transcriptomique a confirmé la perte de fonction de *PAX5* sur ses gènes cibles et a révélé une signature moléculaire spécifique au mutant. Nos données soulignent l'importance de la prise en compte de la composante héréditaire et mettent ainsi en évidence l'importance de la dérégulation transcriptionnelle dans la leucémogénèse des LAL-B familiales, en particulier des gènes impliqués dans la différenciation lymphoïde B.





# ABSTRACT

The *PAX5* (*Paired box 5*) gene encodes a key transcription factor crucial for B-cell differentiation. We showed that the two *PAX5* isoforms are differentially regulated but have equivalent function during early B-cell differentiation. Indeed, *PAX5A* and *PAX5B* isoforms can both induce B-cell program but may have functional differences after B-cell activation. The tight control of their expression may thus reflect a way to finely tune *PAX5* dosage during B-cell differentiation process.

*PAX5* is a well-known haploinsufficient tumor suppressor gene in human B-cell precursor acute lymphoblastic leukemia (BCP-ALL) and is the main target of a wide diversity of somatic alterations in childhood and adult BCP-ALL, occurring in one third of sporadic cases. However, the role of *PAX5* fusion proteins in BCP-ALL initiation and transformation is ill-known. We previously reported a new recurrent t(7;9)(q11;p13) chromosomal translocation in human BCP-ALL that juxtaposed *PAX5* to the coding sequence of elastin (*ELN*). To study the function of the resulting *PAX5-ELN* fusion protein in BCP-ALL development, we generated a mouse model in which the *PAX5-ELN* transgene is expressed specifically in B cells. *PAX5-ELN*-expressing mice efficiently developed BCP-ALL phenotype with a penetrance of 80%. Leukemic transformation was associated with clonal Immunoglobulin gene rearrangement and recurrent secondary mutations in *Ptpn11*, *Kras*, *Pax5*, and *Jak3* genes affecting key signaling pathways required for cell proliferation. Our functional studies demonstrated that *PAX5-ELN* impairs B-cell development *in vitro* and *in vivo* and induces an aberrant expansion of the pro-B cell compartment at the preleukemic stage. Our molecular and computational approaches identified *PAX5-ELN*-regulated candidate genes that establish the molecular bases of the preleukemic state to drive BCP-ALL initiation. In conclusion, our study provides a new *in vivo* model recapitulating the multistep leukemogenesis process of human BCP-ALL and strongly implicates *PAX5* fusion proteins as potent oncoproteins in leukemia development.

Furthermore, there is increasing evidence for an inherited genetic basis of susceptibility to childhood BCP-ALL. In this context, four unrelated families with childhood BCP-ALL expressing heterozygous *PAX5* germline point mutations were recently reported: the recurrent mutation *PAX5* G183S affecting the octapeptide domain of *PAX5* has been described in three families while *PAX5* R38H affecting its DNA-binding paired domain has been identified in another one. We strengthen the hypothesis of inherited character of familial BCP-ALL with the description of three novel familial BCP-ALL cases in related patients that express the germline *PAX5* R38H mutation. To uncover the intrinsic effect of *PAX5* R38H mutant in B-cell development, we performed *in vitro*, and *in vivo* functional assays combined with a gene expression analysis, based on a retroviral complementation approach. Our results indicated that *PAX5* R38H mutant acts as a strong hypomorphic variant that fails to drive B-cell differentiation and does not exert a dominant-negative effect on wild-type *PAX5*. Syngeneic transplantation of *PAX5* R38H-expressing cells demonstrated maintenance of engraftment capacity and led to development of BCP-ALL phenotype in mice. Our transcriptomic analysis of these *PAX5* R38H-expressing cells showed that *PAX5* R38H drastically alters the pattern of expression of *PAX5* target genes but also revealed a distinct molecular signature specific to *PAX5* R38H. Together with previous unrelated family study, our observations allow to establish the recurrence of the germline *PAX5* R38H mutation associated with BCP-ALL. Our data also highlight the importance of transcriptional dysregulation in leukemogenesis of familial BCP-ALL, particularly of genes involved in B-cell differentiation.



Liste des abréviations

#	désigne un nombre
ABL1	Abelson gene
ADN	Acide Déoxyribonucléique
ALPs	All Lymphoid Progenitors
AP-1	Activator Protein 1
ARN	Acide Ribonucléique
AJA	Adolescents et Jeunes Adultes
AYA	Adolescents and Young Adults
B220/CD45R	Receptor-type protein tyrosine phosphatase (PTP) C
BACH2	BTB and CNC homologue 2
BCAP	B-Cell PI3K Adaptator
BCL-xL	B-Cell Lymphoma-extra Large
BCL6	B cell lymphoma 6
BCL11B	B-cell lymphoma/leukemia 11B
Bcl2l1 gene	B-cell lymphoma 2-like 1 gene (encoding Bcl-xL)
BCR	B-Cell Receptor
BCR gene	Breakpoint Cluster Region gene
BLK	B Lymphoid Tyrosine Kinase
BLNK/SLP-65	B-Cell Linker protein (alias BASH, BCA)
BLIMP1	B lymphocyte-induced maturation protein 1
BLPs	B cell-biased lymphoid progenitors
BSAP	B-cell lineage Specific Activation Protein
BTK	Bruton Tyrosine Kinase
BTLA/CD272	B- and T-lymphocyte attenuator
CAR T-cell	Chimeric Antigen Receptor T cell
CCND2	Cyclin D2
CD	Cluster of Differentiation
CD4	lymphocyte T CD4+
CD8	lymphocyte T CD8+
CDKN2A/2B	Cyclin-dependent kinase inhibitor 2A/2B
CEBP $\alpha$	CCAAT/Enhancer-Binding Protein alpha
CEBP $\epsilon$	CCAAT/Enhancer-Binding Protein epsilon
Ciita	MHC class II transactivator
CLP	Common Lymphoid Progenitors
CMP	Common Myeloid Progenitors
c-Myc	c-Myelocytomatosis oncogene
CNV	Copy Number Variation
CREBBP	CREB-binding protein
CRLF2	Cytokine-Receptor Like Factor 2
CSF1R	Colony Stimulating Factor 1 (CSF1) Receptor (CD115 ou M-CSFR)
CSH	Cellule-s Souche-s Hématopoïétique-s
CSR	Class Switch Recombination (commutation isotypique des immunoglobulines)
(m)DC	(myeloid) Dendritic Cells
(p)DC	(plasmacytoid) Dendritic Cells
DN	Double-Negative
DNMT3A	DNA methyltransferase 3A

di	diencephalon
DP	Double-Positive
E2A/TCF3	Transcription factor E2-alpha
EBF1	Early B-Cell Factor 1
EGIL	European Group for Immunology of Leukemia
ELN	Elastin
ELP	Early Lymphoid Progenitor
EPOR	Erythropoietin (EPO) Receptor
ERG	ETS-Related Gene
ERK1/2	Extracellular signal-Regulated Kinases 1/2
ery	erythroid cell-s
ETP	Early T-cell Progenitor
ETV6	ETS translocation variant 6
Ezh2	EZH1/2 : Enhancer of Zeste Homolog 1/2
Flt3	Fms Like Tyrosine kinase 3
Flt3L	Flt3 Ligand
FOXO#	Forkhead box protein O#
FOXP1	Forkhead box protein P1
Fra1	FOS-related antigen 1
FT-s	Facteur-s de transcription
Fo B	Follicular B cells
GATA3	Trans-acting T-cell-specific transcription factor GATA-3. FT se liant à la séquence d'ADN [A/T] <u>GATA</u> [A/G].
GD	Gestational Day
GMP	Granulocyte-Monocyte/Macrophage lineage restricted Progenitor
GAB2	GRB2 associated-binding protein 2
GC	germinal centre
Gran	granulocyte-s
GRAP	GRB2-related adapter protein
Grb2	Growth factor receptor-bound protein 2
GSEA	Gene Set Enrichment Analysis
GWAS	Genome Wide Association Study
HAT	Histone Acetyltransferase
hb	hindbrain
HD	Homeodomain
HDAC	Histone Deacetylase
HR	High Risk
HSC	Hematopoietic Stem Cell
iAMP21	intrachromosomal amplification of chromosome 21
ID	Inhibition Domain
ID2/3	DNA-binding protein inhibitor ID2/3
IgH	Heavy Immunoglobulin
IgL	Light Immunoglobulin
IgM/IgD	Immunoglobulin de classe M/D
IKZF1/Ikaros	Ikaros family zinc finger protein 1
IKZF2/Helios	Ikaros family zinc finger protein 2
IKZF3/Aiolos	Ikaros family zinc finger protein 3
IL#	Interleukine #

IL-2R $\alpha$ /CD25	Interleukin-2 Receptor alpha chain
IL-7R	IL-7 Receptor
ILC2	type 2 Innate Lymphoid Cells
IRF4/8	Interferon-Regulatory Factors 4/8
ITAM	Immunoreceptor Tyrosine-based Activation Motif
ITK	Inhibiteurs de Tyrosine Kinase
JAK#	Janus Kinase #
Kit/SCFR	Mast/Stem Cell Growth Factor (SCF) Receptor
KMT2A	Histone-lysine N-methyltransferase 2A (MLL)
KRAS	<u>Kirsten</u> rat sarcoma viral oncogene homolog (GTPase)
LA	Leucémie Aiguë
LAL	Leucémie Aiguë Lymphoïde
LAL-B	Leucémie Aiguë Lymphoblastique B
LAL-T	Leucémie Aiguë Lymphoblastique T
LAM	Leucémie Aiguë Myéloïde
LEF1	Lymphoid Enhancer-binding Factor 1
Lin	Lineage
LSK	Lin- Sca1+ Kit+
LB	Lymphocyte B
LMPP	Lymphoid-primed Multipotent Progenitor
LT	Lymphocyte T
LT-HSC	Long-Term Hematopoietic Stem Cell
Ly6(D/G)	Lymphocyte antigen 6D/G
LYN	Lck/Yes Novel tyrosine kinase protein
Mac	Macrophages
MAPK	Mitogen-Activated Protein Kinase-s
Mcl1	Induced myeloid leukemia cell differentiation protein
M-CSFR	Monocyte Colony Stimulating Factor (M-CSF) Receptor (CD115 ou CSF1R)
MDS	Myelodysplastic Syndrome
MEF2	Myocyte-specific Enhancer Factor 2
MEK1/MEK2	<u>MAPK/ERK</u> kinases 1/2
MEP	Megakaryocyte-Erythrocyte lineage-restricted Progenitors
mes	mesencephalic
MHB	Midbrain–Hindbrain Boundary
Mk	Megakaryocyte
MLP	Multilineage Progenitor
MO	Moelle osseuse
mono	monocyte-s
mega	mégacaryocyte
MPAL	Mixed-Phenotype Acute Leukemia (leucémie aiguë biphénotypique)
MPP	Multipotent progenitors
MRD	Minimal Residual Disease
MYC	Myelocytomatosis oncogene
NCoR	Nuclear receptor Co-Repressor
Neu	Neutrophil-s
NF1	Neurofibromin 1 gene codant la Neurofibromine
NF B	Newly Formed B cells
NF- $\kappa$ B	Nuclear Factor Kappa B

NK	Natural Killer
NLS	Nuclear Localization Signal
NOL4L/C20orf112	Nucleolar Protein 4 Like/Chromosome 20 Open Reading Frame 112
Notch1/2	Neurogenic locus notch homolog protein 1/2
NRAS	<u>Neuroblastoma</u> rat sarcoma viral oncogene homolog (GTPase)
OAR	Otp/Aristaless/Rax
OBF1	Oct binding factor 1
OMS	Organisation Mondiale de la Santé
OC	optic cup
OP motif	Octapeptide motif
OP	otic placode
OPTP	optic placode
ov	optic vesicle or otic vesicle depending on the figure, cf legend.
P2RY8	P2Y purinoceptor 8
Pax5/PAX5	<u>Paired box</u> gene or protein 5
PBD	Paired Box Domain
PD	PAIRED (domain)
PDGFR $\beta$	Platelet-derived growth factor receptor beta
PDK1	Phosphoinositide-Dependent protein Kinase 1
PI3K	Phosphoinositide-3-Kinase
PIP2	Phosphatidylinositol-4,5-bisphosphate
PIP3	Phosphatidylinositol-3,4,5-trisphosphate
PKB/Akt	Protein Kinase B
PKC $\theta$	Protein Kinase C
PLC- $\gamma$ (2)	Phospholipase C- $\gamma$ (2)
Pou2af1	POU domain class 2 associating factor 1
PRC2	Polycomb-repressive complex 2
prd	paired
Prdm1	PR domain zinc finger protein 1
PtdIns(3,4,5)P3	Phosphatidylinositol 3,4,5 trisphosphate
PtdIns(4,5)P2,	Phosphatidylinositol 4,5 diphosphate
PTEN	Phosphatase and <u>T</u> ensin homolog
PTPN11/SHP2	Tyrosine-protein Phosphatase Non-receptor type 11
PU-1	FT codé par <i>SPI1</i> et se fixant sur une PU-box, séquence riche en purines
r#	rhombomere
Rag1/2	Recombination-activating genes 1/2
RB/RB1	Retinoblastoma gene or protein
rho	rhombencephalon
RORA	RAR Related Orphan Receptor A
RUNX1	Runt-related transcription factor 1
sc	spinal cord
Sca1	Stem Cell Antigen 1
SETD2	SET Domain-containing 2 (Histone Lysine Methyltransferase)
SFK	SRC-family protein tyrosine kinase
SH	SRC Homology
SHP2/PTPN11	Src Homology region 2 (SH2)-domain-containing protein Phosphatase-2
SLP-65/BLNK	Src Homology [SH2]-domain-containing leukocyte protein of 65 kDa
SNC	Système Nerveux Central

SPI1 gene	Spleen Focus Forming Virus (SFFV) Proviral Integration proto-oncogene
SR	Standard Risk
SRC	Proto-oncogene tyrosine-protein kinase (from <u>sar</u> coma)
STAT#	Signal Transducer and Activator of Transcription 1-alpha/beta #
ST-HSC	Short-Term Hematopoietic Stem Cell
SUZ12	<u>S</u> uppressor of <u>Z</u> este homolog 12
SYK	Spleen Tyrosine Kinase
TAD	Transactivation Domain
te	telencephalon
TCF3/E2A	Transcription Factor 3
Tdt	Terminal deoxynucleotidyl transferase
TKI	Tyrosine Kinase Inhibitors
TOX	Thymocyte Selection Associated High Mobility Group Box protein
TP53	Tumor Protein P53
TP63	Tumor Protein P63
UPR	Unfolded Protein Response
WHO	World Health Organization
WHSC1/NSD2	Wolf-Hirschhorn Syndrome Candidate 1/ nuclear receptor binding SET domain protein 2 (histone methyltransferase)
XBP1	X box-binding protein 1
ZAP70	Zeta( $\zeta$ )-chain-associated protein tyrosine kinase of 70 kDa
ZBTB16/20	Zinc finger and BTB Domain Containing 16/20







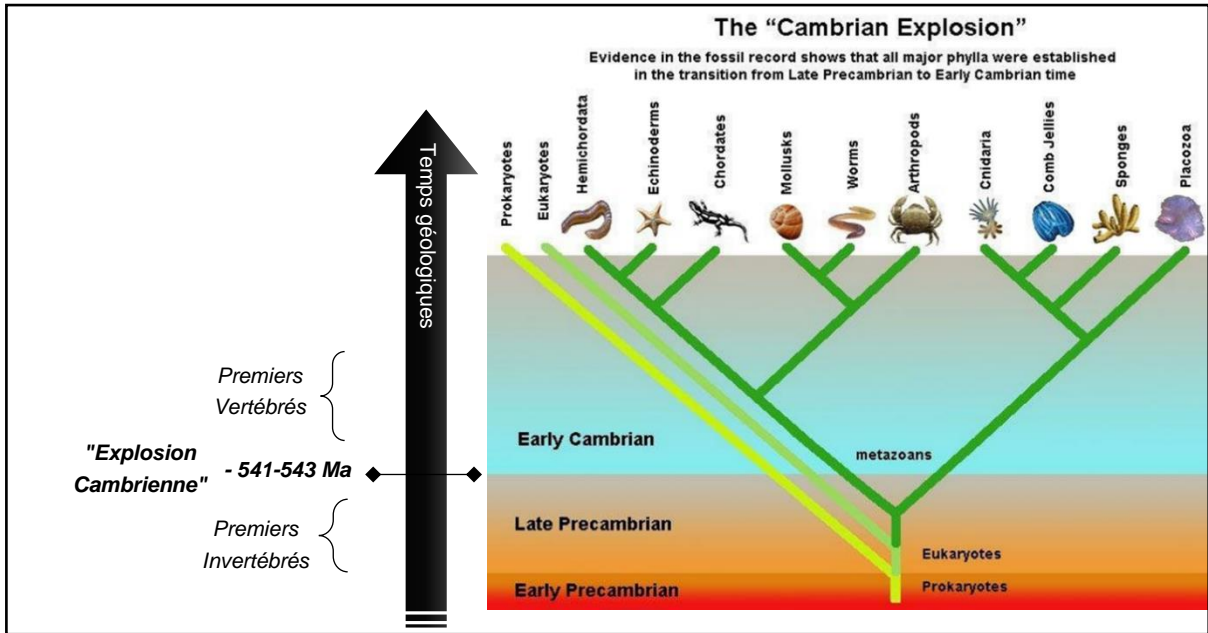


Figure 1.01 : L'explosion évolution du Cambrien

Figure de droite issue de Jackson, Graeme (2019). Evolution and Culture Through Subsistence Strategy. *Ma*: Millions d'années.

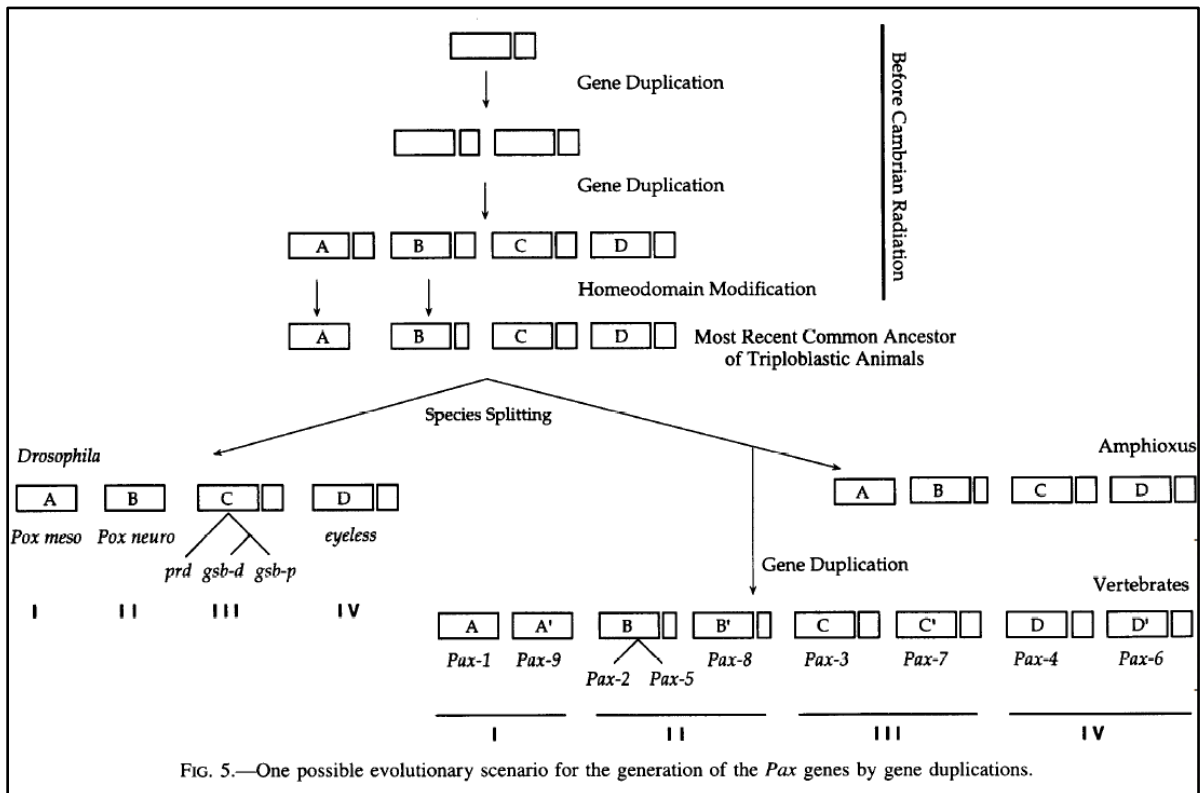


Figure 1.02 : Scénario d'évolution possible pour la génération des gènes Pax par duplication génomique. Figure issue de Balczarek et al., 1997.

# Partie I : Introduction

## Chapitre 1 de l'introduction

Mon travail de thèse a porté sur le gène de développement *PAX5* qui code un facteur de transcription (*PAX5*) fondamental pour la formation du système nerveux central et essentiel à l'établissement du programme de différenciation lymphoïde B.

### 1.1. Les gènes de développement *Pax*

#### 1.1.1. Évolution phylogénétique

*Pax5* est l'un des neuf membres de la famille de gènes de développement *Paired box* (*Pax1* à *Pax9*). Ces gènes codent des protéines caractérisées par la présence d'un domaine « Paired box » et généralement impliquées dans le contrôle des processus de différenciation tissulaire (Walther et al., 1991). Les gènes *Pax* sont fortement conservés au cours de l'évolution, étant présents dans tous les ordres des Métazoaires (désignant les organismes pluricellulaires) ; le gène ancestral *proto-pax* était déjà présent dans le génome des animaux à la fin du Précambrien, avant l'explosion évolutive Cambrienne qui a participé à la diversification des Métazoaires (Figure 1.01 et Tableau-repère des ères géologiques page suivante ; Paixão-Côrtes et al., 2015). Deux cycles de duplication du génome entier suivis de pertes génomiques divergentes entre les espèces ont généré les différents gènes *Pax* orthologues actuels (Figure 1.02 ; Balczarek et al., 1997 ; Paixão-Côrtes et al., 2015). Par exemple, les Spongiaires correspondant à l'embranchement basal des Métazoaires, donc considérés comme les organismes pluricellulaires les plus primitifs, possèdent un gène *Pax* nommé *PAXB*. La plupart des Cnidaires, apparus au Précambrien il y a 630 à 570 millions d'années, expriment déjà quatre gènes *Pax* : *PAXA*, *PAXB*, *PAXC* et *PAXD* (Miller et al., 2000). Les Cnidaires constituent un embranchement des Métazoaires, auquel appartiennent les premiers organismes à avoir développé une structure pluricellulaire organisée et dont font notamment partie les coraux, les méduses et les anémones de mer.

→ *PAXB* ou *PAXD* pourrait être le proto-*PAX* ancestral (Paixão-Côrtes et al., 2015).



Illustration d'espèces de Cnidaires (de gauche à droite : une gorgone éventail *Annella mollis* ; une méduse ; une anémone de mer).



Échelle de temps géologiques (Ma = Millions d'années), du plus récent au plus ancien			
Éon :	Ère :	Période :	
<b>Phanérozoïque</b> · 1/6 <sup>e</sup> des temps géologiques · durée moyenne des périodes : 50 Ma · Incluant SIX extinctions de masse	<b>Cénozoïque</b> <i>Depuis -66 Ma</i>	Quaternaire ( <i>extinction de l'Holocène en cours !</i> )	
		Tertiaire	
	<b>Mésozoïque</b> (anciennement ère secondaire ou ère des reptiles) <i>De -252,2 à -66,0 Ma</i>	Crétacé ( <i>suivi de l'extinction du Crétacé-Paléogène</i> )	
		Jurassique	
		Trias ( <i>suivi de l'extinction du Trias-Jurassique</i> )	
	<b>Paléozoïque</b> (anciennement ère primaire ou ère des poissons) <i>De -541 à -252,2 Ma</i>	Permien ( <i>suivi de l'extinction du Permien-Trias</i> )	
		Carbonifère	Pennsylvanien
			Mississippien
		Dévonien ( <i>suivi d'une extinction</i> )	
		Silurien	
		Ordovicien ( <i>suivi d'une extinction</i> )	
		Cambrien, de -541 à -485 Ma	C. Supérieur ('late')
			C. Moyen ('middle')
C. Inférieur ('early')			
<b>Protérozoïque</b> · 3 <sup>e</sup> éon du Précambrien ('late') · 5/6 <sup>e</sup> des temps géologiques · durée moyenne des périodes : 200 Ma	<b>Néoprotérozoïque</b> <i>De -1000 à -541 Ma</i>	Ediacarien	
		Cryogénien	
		Tonien	
	<b>Mésoprotérozoïque</b> <i>De -1600 à -1000 Ma</i>	Sténien	
		Ectasien	
		Calymmien	
	<b>Paléoprotérozoïque</b> <i>De -2500 à -1600 Ma</i>	Stathérien	
		Orosirien	
		Rhyacien	
		Sidérien	
<b>Archéen</b>	Néo-, Méso-, Paléo-, Eo-archéens.	<i>De -2500 à -4000 Ma</i>	
<b>Hadéen</b>	<i>Avant -4000 Ma</i>		

Tableau-repère : liste des ères géologiques en ordre chronologique inversé.

Super-familles	Sous-familles	Caractéristiques structurales	Paralogues au sein des sous-groupes	Localisation chromosomique chez la souris	Localisation chromosomique chez l'homme
PAXD-like	Groupe I	Motif octapeptide Pas d'homéodomaine	<b>Pax1</b> (Chalepakis et al., 1991)	<b>Chr 2</b> Walther et al., 1991	<b>20p11</b> Stapleton et al., 1993
			<b>Pax9</b> (Neubuser, Koseki, and Balling, 1995)	<b>Chr 12</b> Wallin et al., 1993	<b>14q12-q13</b> Stapleton et al., 1993
PAXB-like	Groupe II	Motif octapeptide Homéodomaine partiel	<b>Pax2</b> (Dressler et al., 1990)	<b>Chr 19</b> Walther et al., 1991	<b>10q25</b> Stapleton et al., 1993
			<b>Pax5</b> (Adams et al., 1992)	<b>Chr 4</b> Walther et al., 1991	<b>9p13</b> Stapleton et al., 1993
PAXD-like	Groupe III	Motif octapeptide Homéodomaine complet	<b>Pax8</b> (Plachov et al., 1990)	<b>Chr 2</b> Walther et al., 1991	<b>2q12-q14</b> Stapleton et al., 1993
			<b>Pax3</b> (Goulding et al., 1991)	<b>Chr 1</b> Walther et al., 1991	<b>2q35</b> Tsukamoto et al., 1992
PAXB-like	Groupe IV	Pas de motif octapeptide Homéodomaine complet	<b>Pax7</b> (Jostes, Walther, and Gruss 1991)	<b>Chr 4</b> Walther et al., 1991	<b>1p36.2-p36.12</b> Stapleton et al., 1993
			<b>Pax4</b> (Walther et al., 1991)	<b>Chr 6</b> Walther et al., 1991	<b>7q32</b> Stapleton et al., 1993
			<b>Pax6</b> (Walther and Gruss, 1991)	<b>Chr 2</b> Walther et al., 1991	<b>11p13</b> Ton et al., 1991

Tableau 1 : Classification des protéines Pax en fonction de leur structure protéique. Tableau créé d'après les données de Balczarek et al., 1997 et la revue de Paixão-Côrtes et al., 2015.

### 1.1.2. Découverte des gènes *Pax* chez la drosophile

Historiquement, la famille de gènes Paired box (*Pax*) est définie par la présence du domaine *paired box*, domaine initialement identifié dans des gènes de la drosophile (*Drosophila melanogaster*) : *paired* (*prd*) et *gooseberry* (*gsb*) (Bopp et al., 1986).

Le gène *paired* (*prd*) séquencé en 1986 chez la drosophile, isolé et décrit par Frigerio G et al., fait partie de la catégorie des gènes *pair-rule* requis pour la segmentation de l'embryon de la drosophile (*Drosophila melanogaster*). Il code une protéine initialement caractérisée par la présence d'une séquence spécifique (nommée PRD) enrichie en histidine et proline (*His-Pro repeat*) à l'extrémité carboxy-terminale et contenant plusieurs clusters de glutamine (*oligo-Gln M-repeat*) (Frigerio et al., 1986). Dès lors, il a été montré que cette séquence PRD était fréquemment associée à un homéodomaine. Deux motifs protéiques codés par le gène *paired* ont été identifiés : une séquence *homeobox* élargie, ainsi que le domaine *paired box* (de 128 acides aminés), deux fois plus grand que le motif *homeobox* (Bopp et al., 1986). A l'instar de l'homéodomaine, le domaine *paired box* bénéficie d'une activité de liaison à l'ADN spécifique (Treisman et al., 1991). Par recherche d'homologie de séquence, la séquence codant le domaine *paired box* a été identifiée dans d'autres gènes de segmentation et gènes tissu-spécifiques de la drosophile (Bopp et al., 1986).

Depuis, de nombreux gènes codant ce domaine Paired box ont été identifiés, conduisant par contraction du nom de ce domaine à la définition de la famille de gènes homologues Pax. Ces gènes ont été identifiés chez une myriade d'invertébrés et de vertébrés. Les gènes *Pax* sont fortement conservés entre les espèces (*Pax* orthologues) et la description de similitudes de structure entre certains *Pax* homologues a permis d'identifier des sous-familles de *Pax* au sein d'un même organisme (*Pax* paralogues).

### 1.1.3. Classification des gènes *Pax*

Les Vertébrés présentent le plus de gènes *Pax* homologues (Paixão-Côrtes et al., 2015). Chez les Mammifères, neufs gènes homologues ont jusqu'à présent été identifiés. Les neuf gènes homologues *Pax* sont répartis sur différents chromosomes chez la souris (Walther et al., 1991 pour *Pax1-8*, Wallin et al., 1993 pour *Pax9*) et chez l'homme (Ton et al., 1991 pour *PAX6*, Wilcox et al., 1992 et Tsukamoto et al., 1994 pour *PAX3*, Stapleton et al., 1993 pour *PAX1-2*, *PAX4-5*, *PAX7-9*) contrairement aux gènes de développement *Hox* qui sont regroupés en clusters (Tableau 1).



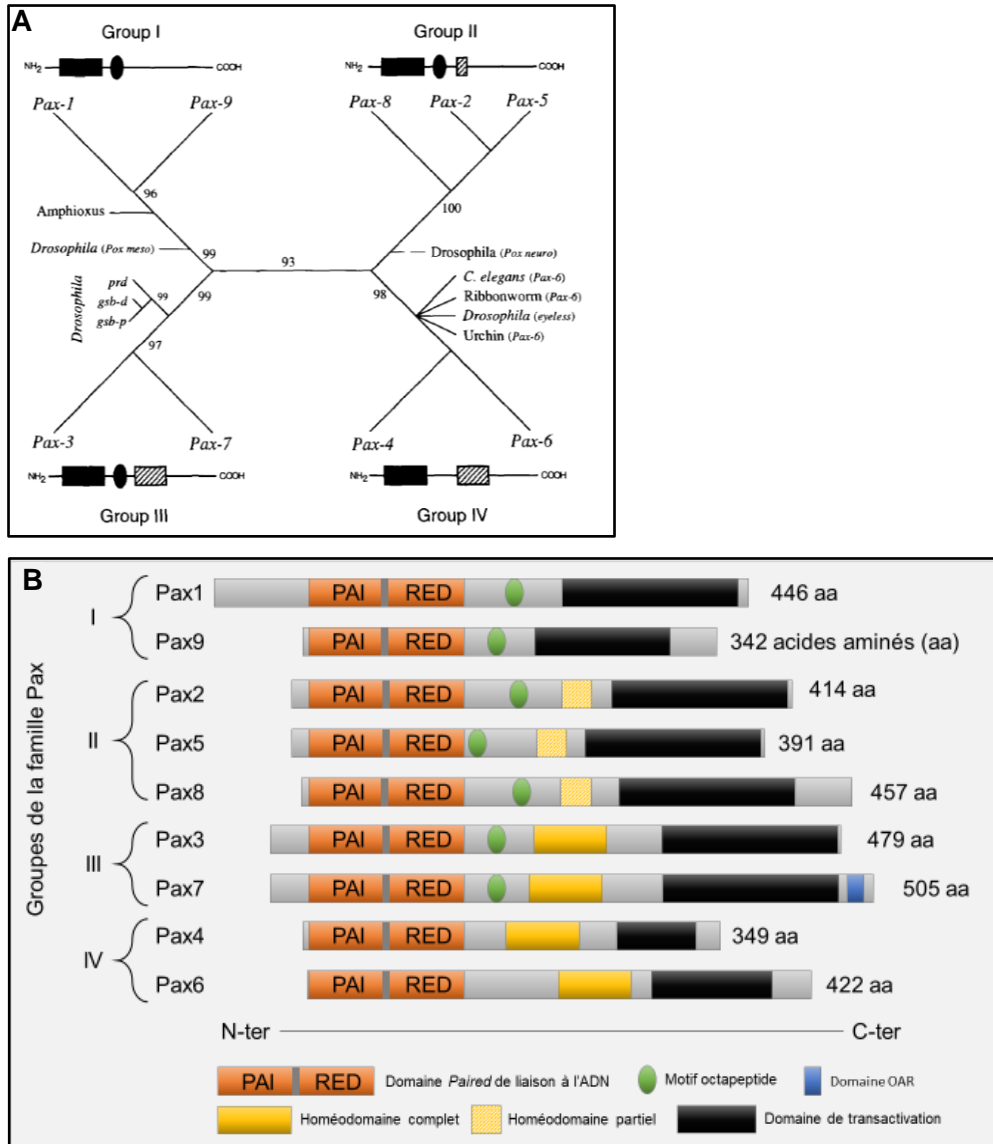


Figure 1.03 : Classification phylogénétique, structure protéique et domaines fonctionnels des facteurs de transcription de la famille Pax.

**A.** Schéma des relations phylogénétiques entre les gènes *Pax* d'après l'analyse des séquences du domaine Paired et de l'homéodomaine conduites en 1997 par Balczarek *et al.* Les *Bootstrap p-values* > 90% sont indiquées sur les branches correspondantes. Le domaine Paired est symbolisé par un rectangle noir, le motif octapeptide par une ellipse noire et l'homéodomaine par un rectangle hachuré. Les lignes pointillées indiquent une incertitude de positionnement phylogénétique. Figure issue de Balczarek *et al.*, 1997.

**B.** Illustration schématique des différences de structure des facteurs de transcription de la famille Pax chez les Mammifères, adaptée de la revue de Mayran *et al.*, 2015. La famille Pax est divisée en quatre groupes de membres Pax paralogues, définis en fonction de la présence de domaines fonctionnels tels qu'un motif octapeptide (en vert, mais absent dans Pax4/6) et/ou un homéodomaine de classe paired (en jaune) qui est soit complet (un motif hélice-coude-hélice à trois hélices alpha) dans Pax3/7/4/6, soit partiel (une seule hélice) dans Pax2/5/8 mais est absent chez Pax1/9. Pax7 est l'unique membre de la famille à contenir un domaine OAR (Otp/Aristaless/Rax) à son extrémité carboxy-terminale. Illustration B créée par Laura Jamrog.

Sur la base de leur évolution moléculaire, ces neuf gènes *Pax* sont classifiés en deux groupes (aussi nommés super-familles) : *PAXB-like* et *PAXD-like* (Paixão-Côrtés et al., 2015). Le terme gène *PAX-like* est un terme générique utilisé pour tous les gènes *PAX/Pax* homologues et orthologues. Le groupe *PAXB-like* contient le domaine *paired box* ainsi qu'un motif octapeptide et l'homéodomaine\* de la classe PAIRED des gènes à homéoboîtes, \*complet ou partiel. Le groupe *PAXD-like* se définit par la présence d'un quatrième domaine peptidique conservé au cours de l'évolution à l'extrémité caudale de l'homéodomaine (*paired type homeodomain tail*) (Paixão-Côrtés et al., 2015). Les neuf gènes *Pax* sont actuellement classés parmi quatre groupes de gènes paralogues en fonction de leur haut degré d'homologie de séquences primaires et de leurs similarités en termes d'organisation génomique, de structure des domaines protéiques qu'ils codent et de patron d'expression durant le développement (Noll, 1993 ; Balczarek et al., 1997 ; Mayran et al., 2015) ; Figure 1.03) :

- Les paralogues *Pax1* et *Pax9* appartiennent au groupe I : ils ne possèdent pas d'homéodomaine
- Les paralogues *Pax2*, *Pax5* et *Pax8* appartiennent au groupe II : ils ont la particularité de contenir un homéodomaine tronqué (perte de la dernière moitié de l'homéodomaine).
- Les paralogues *Pax3* et *Pax7* appartiennent au groupe III : ils sont dotés d'un motif octapeptide et d'un homéodomaine complet. *Pax7* est le seul à avoir conservé le motif OAR (Otp/Aristaless/Rax).
- Les paralogues *Pax4* et *Pax6* appartiennent au groupe IV : ils sont dénués de motif octapeptide.

Par ailleurs, il existe un *Pax10* chez certains Vertébrés (qui ferait partie du groupe III) qui a perdu le caractéristique domaine PAIRED au cours de l'évolution ; *Pax10* n'a pas été conservé par les Mammifères.

Bien que les groupes III et IV présentent tous deux un homéodomaine complet, ils sont moins apparentés entre eux que le groupe I avec le groupe III d'une part et le groupe II avec le groupe IV d'autre part, d'un point de vue phylogénétique (Balczarek et al., 1997) (Figure 1.03A).

Super-familles	Sous-familles	Caractéristiques structurales	Paralogues au sein des sous-groupes	Domaines d'expression (liste non-exhaustive)
PAXD-like	Groupe I	Motif octapeptide Pas d'homéodomaine	<i>Pax1</i>	Sclérotome, squelette, thymus
			<i>Pax9</i>	Sclérotome, squelette, dents, thymus
PAXB-like	Groupe II	Motif octapeptide Homéodomaine partiel	<i>Pax2</i>	Système nerveux central, reins, yeux, oreilles, glandes mammaires
			<i>Pax5</i>	Lignage lymphoïde B, testicules, système nerveux central
			<i>Pax8</i>	Système nerveux central, thyroïde, reins
PAXD-like	Groupe III	Motif octapeptide Homéodomaine complet	<i>Pax3</i>	Crête neurale, système nerveux central, somites/muscles
			<i>Pax7</i>	Crête neurale, système nerveux central, somites/muscles
PAXB-like	Groupe IV	Pas de motif octapeptide Homéodomaine complet	<i>Pax4</i>	Pancréas, intestins, glande pinéale, système nerveux central
			<i>Pax6</i>	Pancréas, intestins, système nerveux central, glande pinéale, yeux, nez

Tableau 2 : Diversité des territoires d'expression des gènes *Pax*. Tableau créé d'après les données de Chi & Epstein, 2002 et Dahl et al., 1997.

## 1.2. Rôles majeurs des protéines Pax durant le développement

Les neuf gènes de développement *Pax* des Mammifères codent une famille de facteurs de transcription (FTs) conservés depuis des centaines de millions d'années d'évolution. Les FTs *Pax* régulent de nombreuses fonctions clés, à de nombreuses étapes de l'embryogenèse (Dahl et al., 1997 ; Tableau 2). Ce sont des régulateurs majeurs du développement et de l'homéostasie d'une grande variété de tissus et d'organes, incluant ceux des yeux, du système nerveux central, des reins, du pancréas, du squelette, des muscles, des dents, du thymus et du tissu hématopoïétique.

Leur expression est strictement régulée et définie en termes de dosage, de territoires et de périodes d'expression. Ils peuvent partager des domaines d'expression tout en présentant une spécificité d'expression tissulaire, voire cellulaire. Ainsi des protéines *Pax* peuvent accomplir des fonctions redondantes, convergentes ou spécifiques (Wachtel & Schäfer, 2015).

D'une part, ces régulateurs transcriptionnels revêtent une grande importance dans la **spécification des lignées** cellulaires pendant le développement et ultérieurement : ils sont généralement impliqués dans **l'identité cellulaire** et le contrôle des processus de **différenciation tissulaire** (revue de Walther et al., 1991 ; revue de Blake & Ziman, 2014). Cependant, leur expression doit souvent être réprimée pour permettre la différenciation terminale. D'autre part, ils endossent des fonctions essentielles pour **préserver l'homéostasie tissulaire** dans l'organisme adulte. La persistance de l'expression de certaines protéines *Pax* garantit principalement la maintenance de divers progéniteurs et cellules souches permettant ainsi la régénération tissulaire, physiologique ou post-lésionnelle. À l'inverse, leur expression persiste parfois dans certaines cellules matures pour équilibrer une voie de régulation.

**La majorité des protéines Pax, à l'exception de Pax1 et Pax9, opèrent une forte coopération au cours de la formation du système nerveux central (SNC, composé du cerveau et de la moelle épinière), en s'exprimant transitoirement entre le huitième jour (E8.0-E8.5) et le dixième jour (E10.0-E10.5) de la période embryonnaire (murine) dans des zones spécifiques du tube neural qui est l'organe embryonnaire à l'origine du SNC. Les gènes Pax prennent part aux processus de régionalisation du cerveau ainsi qu'à la maintenance de l'identité cellulaire propre à chaque région ; c'est pourquoi leurs patrons d'expression sont strictement délimités dans le temps et l'espace (Figure 1.04 au dos).**

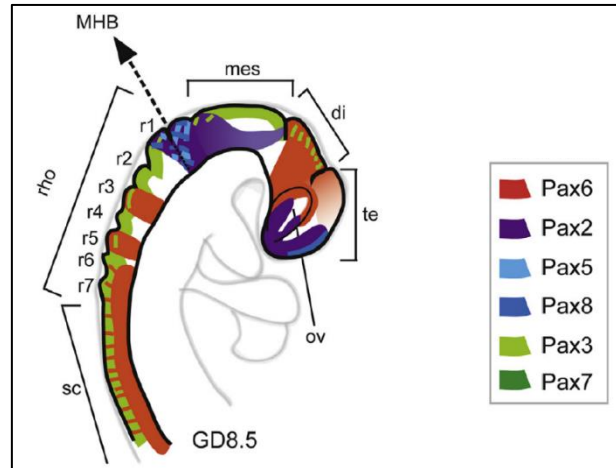


Figure 1.04 : Profils d'expression des protéines Pax au cours du développement du système nerveux central de l'embryon murin.

La plupart des cellules neurales proviennent des progéniteurs exprimant Pax. Embryon de souris au 8<sup>ème</sup> jour de gestation (GD8.5, *gestational day 8.5*) affichant les domaines d'expression des protéines Pax dans le système nerveux central. Les domaines d'expression de Pax2 incluent le télencéphale ventral (te), la vésicule optique (*optic vesicle, ov*), le mésencéphale (mes) et le métencéphale (r1), et le centre organisateur entre le mésencéphale et le rhombocéphale (*midbrain–hindbrain boundary, MHB*). Ses protéines paralogues Pax5 et Pax8 sont détectées dans ce centre organisateur. L'expression de Pax6 est complémentaire à celle de Pax2 dans le cerveau antérieur ou prosencéphale (*forebrain*), qui se subdivise en diencephale (di) et télencéphale (te). Les progéniteurs dans les rhombomères 3 (r3) et 5 (r5) et la moelle épinière (*spinal cord, sc*) expriment aussi Pax6. Pax3 est détecté dans les régions dorsales du diencephale (di), du mésencéphale (mes), du rhombocéphale (rho) et de la moelle épinière (sc). Figure issue de Curto et al., 2015.

D'une part, les gènes *Pax6* (groupe IV), *Pax3* et *Pax7* (groupe III) contribuent à la formation du tube neural et notamment de l'ébauche embryonnaire de la moelle épinière (Figure 1.04). D'autre part, le trio de gènes *Pax* du groupe II (*Pax2*, *Pax5* et *Pax8*) orchestre le centre organisateur, le centre de signalisation qui établit la séparation (*midbrain-hindbrain boundary*, *MHB*) entre le cerveau moyen (mésencéphale) et le cerveau postérieur (rhombencéphale). Les champs d'expression de *Pax2/Pax5/Pax8* se chevauchent spatialement et temporellement dans le cerveau et moelle épinière (*sc*, *spinal cord*) dès le stade embryonnaire [(Nornes et al., 1990) ; (Plachov et al., 1990) ; (Adams et al., 1992) ; revue de (Goode & Elgar, 2009) ; revue de (Curto et al., 2015) ; Figure 1.05]. De plus, *Pax2* coopère avec *Pax5* pour assurer le bon développement du mésencéphale et du cervelet (Urbánek et al., 1997). En résumé, la régionalisation des domaines d'expression des gènes *Pax2* à *Pax8* permet **la structuration de l'embryon**, et en particulier du SNC. Ainsi ils concourent tous à la diversité neuronale tandis que les gènes *Pax1*, *Pax3*, *Pax7* et *Pax9* coordonnent en plus le devenir des cellules de la crête neurale (*NCC*, notamment à l'origine des mélanocytes, des neurones, des cellules gliales, de cellules mésenchymateuses, du cartilage facial et des os) en termes de différenciation et de migration en périphérie. Par ailleurs, certains *Pax* sont exprimés dans les cellules des somites qui sont des organes embryonnaires transitoires positionnés latéralement au tube neural. En effet, *Pax3* et *Pax7* s'activent dans les cellules somitiques à l'origine du dermomyotome, participant ainsi à la genèse du derme du tronc et des muscles de la colonne vertébrale, alors que *Pax1* et *Pax9* s'activent dans les cellules somitiques à l'origine du sclérotome, participant à la formation des vertèbres (revue de Blake & Ziman, 2014).

### 1.2.1. Groupe I : *Pax1* & *Pax9*

*Pax1* et *Pax9* sont les seuls gènes *Pax* à ne pas s'exprimer dans le SNC. *Pax1* et *Pax9* partagent la plupart de leurs territoires d'expression et peuvent accomplir des fonctions compensatoires. Malgré cette redondance, ils n'ont pas la capacité de couvrir la fonction de l'autre dans les tissus où ils ne sont pas coexprimés. Ils sont notamment impliqués dans la spécification du sclérotome, un des tissus transitoires embryonnaires d'où proviennent entre autres la colonne vertébrale, le squelette facial, le sternum et de la ceinture osseuse pectorale/scapulaire qui permet de rattacher les membres antérieurs à la colonne vertébrale. *Pax1* et *Pax9* tiennent également un rôle important dans le développement de la glande parathyroïde ainsi que dans la thymopoïèse et la localisation correcte du thymus. *Pax9* intervient spécifiquement dans la régulation de l'odontogenèse et de la palatogenèse (c'est-à-dire la régulation des mésenchymes dentaire et palatin) alors que *Pax1* est exprimé dans un autre territoire du mésenchyme facial (Sivakamasundari et al., 2018).

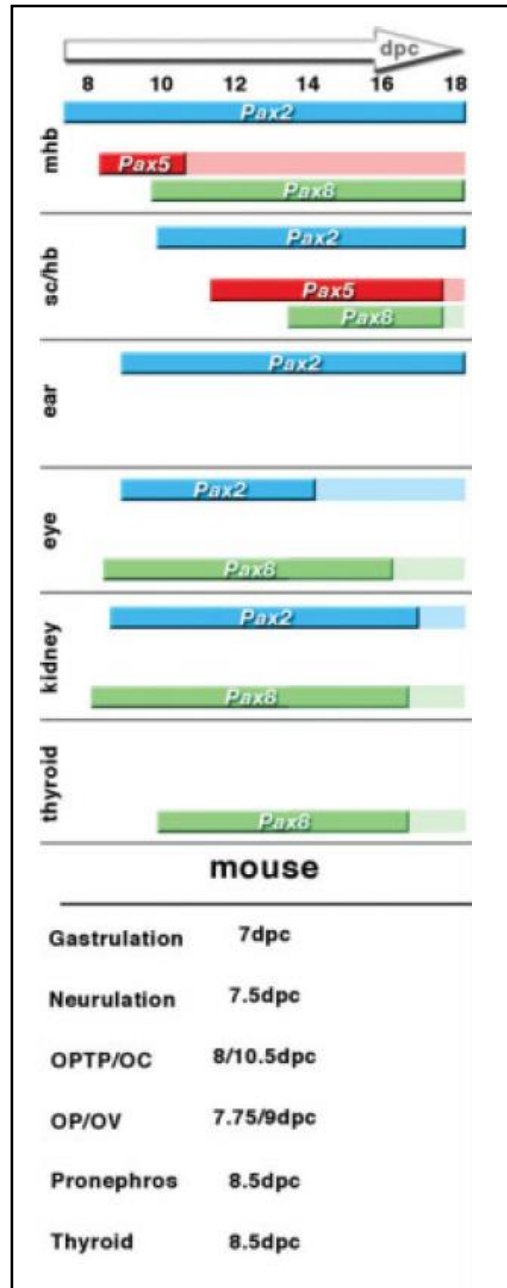


Figure 1.05 : Profils temporels d'expression des gènes paralogues *Pax2/Pax5/Pax8* dans des territoires communs, chez la souris.

Les barres colorées correspondent aux périodes d'expression d'après l'échelle de temps en jours de gestation (*dpc*, *days post-coit* : jours post-fertilisation). *Pax2* est représenté en bleu, *Pax5* en rouge et *Pax8* en vert. Les barres de couleur unie indiquent les preuves expérimentales d'expression alors que les stades plus tardifs représentés par les barres de couleurs transparentes n'ont pas été démontrés expérimentalement. Le panel d'informations inférieur indique les temps approximatifs quand débutent la gastrulation et la neurulation, et les temps approximatifs d'apparition des structures-clés durant le développement embryonnaire de la souris. Abréviations : mhb, *midbrain hindbrain boundary* ; OC, *optic cup* ; OP, *otic placode* ; OTP, *optic placode* ; OV, *otic vesicle* ; sc/hb, *spinal cord/hindbrain*. Figure issue de la revue de Goode & Elgar, 2009.

### 1.2.2. Groupe II : Pax2, Pax5 & Pax8

En dehors du SNC, le duo Pax2 et Pax8 exercent des fonctions déterminantes à plusieurs stades du développement des reins (Goode & Elgar, 2009). Nonobstant, Pax8 participe principalement au développement de la glande thyroïde (Plachov et al., 1990 ; Wachtel & Schäfer, 2015), alors que le gène Pax2 est davantage exprimé seul dans la plupart des tissus du système urogénital en développement. Pax2 fonctionne de concert avec Pax6 lors du développement des yeux, s'exprime également dans l'oreille interne et serait dans une moindre mesure impliqué dans la régulation de la formation du pancréas endocrine (Dressler et al., 1990 ; Nornes et al., 1990 ; Goode & Elgar, 2009).

**Pax5** est le seul des gènes *Pax* à s'exprimer dans les testicules et dans le système hématopoïétique, œuvrant ainsi en solo dans tous les tissus lymphoïdes B du stade embryonnaire au stade adulte (Adams et al., 1992). En effet, PAX5 assure des fonctions prépondérantes tout au long du développement de la lignée lymphoïde B (Nutt et al., 1999), dès le stade précoce de l'engagement des cellules hématopoïétiques dans le lignage B. Il intervient dans les réarrangements  $V_H$ - $DJ_H$  distaux du locus des immunoglobulines (Fuxa et al., 2004) et tient un rôle à un stade plus tardif de la différenciation, dans diverses cellules matures B hors de la moelle osseuse (Medvedovic et al., 2011 pour revue).

### 1.2.3. Groupe III : Pax3 & Pax7

Outre leur rôle dans la régulation du développement du tube neural et de la crête neurale, Pax3 et Pax7 remplissent des fonctions primordiales et distinctes durant la myogenèse du muscle strié squelettique en établissant l'engagement dans la différenciation du tissu musculaire strié. Pax3 est indispensable au développement musculaire embryonnaire : il est exprimé par les cellules qui vont migrer pour participer à la musculature squelettique des membres inférieurs et postérieurs et du diaphragme. Par ailleurs, Pax3 joue également un rôle dans la mélanogenèse. Quant à Pax7, il est responsable de l'auto-renouvellement et de la maintenance du pool hétérogène de cellules satellites, composé des cellules souches musculaire et de progéniteurs engagés. Ainsi Pax7 joue un rôle indispensable dans la régénération du muscle squelettique. La persistance de l'expression de *Pax3* empêche la différenciation des myoblastes alors que celle de *Pax7* empêche la différenciation en cellules satellites. Leur expression doit être inactivée pour permettre la différenciation terminale en myocytes (Wachtel & Schäfer, 2015).





#### 1.2.4. Groupe IV : Pax4 & Pax6

Outre la fonction primordiale de Pax6 dans le développement embryonnaire des yeux, Pax4 et Pax6 jouent aussi un rôle essentiel dans le SNC au-delà du développement embryonnaire, en agissant précisément dans la rétine (Pax4 est exprimé dans les cellules photoréceptrices) et la glande pinéale (ou épiphyse, une petite glande endocrine du cerveau sécrétant la mélatonine). *Pax6* est plus fortement exprimé chez le fœtus puisqu'il est requis pour le développement de la glande pinéale, alors que *Pax4* est plus fortement exprimé après la naissance et suit un rythme d'expression diurne dans la glande pinéale (Rath et al., 2009).

PAX4 et PAX6 régulent la différenciation des cellules pancréatiques endocrines et sont tous les deux impliqués dans la régulation du métabolisme du glucose. PAX4 contrôle le développement du pancréas dès le stade embryonnaire (Smith et al., 1999). PAX4 est effectivement un puissant régulateur de la spécification des cellules bêta pancréatiques qui composent majoritairement les Îlots de Langerhans et produisent l'insuline, hormone hypoglycémisante. PAX4 dicte l'identité des cellules bêta pancréatiques et l'expression constitutive de *PAX4* après la naissance est essentielle pour la production normale de l'insuline. PAX4 pourrait tenir un rôle dans la prolifération et la survie des cellules bêta dans le pancréas adulte. L'expression de *PAX6* dans le pancréas est requise pour la maturation des cellules alpha pancréatiques qui se localisent en périphérie des Îlots de Langerhans et produisent le glucagon, hormone hyperglycémisante. Pax6 est autant un élément clé de la fonctionnalité des cellules bêta matures que des cellules alpha chez l'adulte, puisqu'il participe à la fois à la régulation transcriptionnelle de la biosynthèse de l'insuline et à celle de la biosynthèse et de la sécrétion du glucagon (Gosmain, Katz, et al., 2012 ; Gosmain, Cheyssac, et al., 2012 ; Paixão-Côrtés et al., 2015).

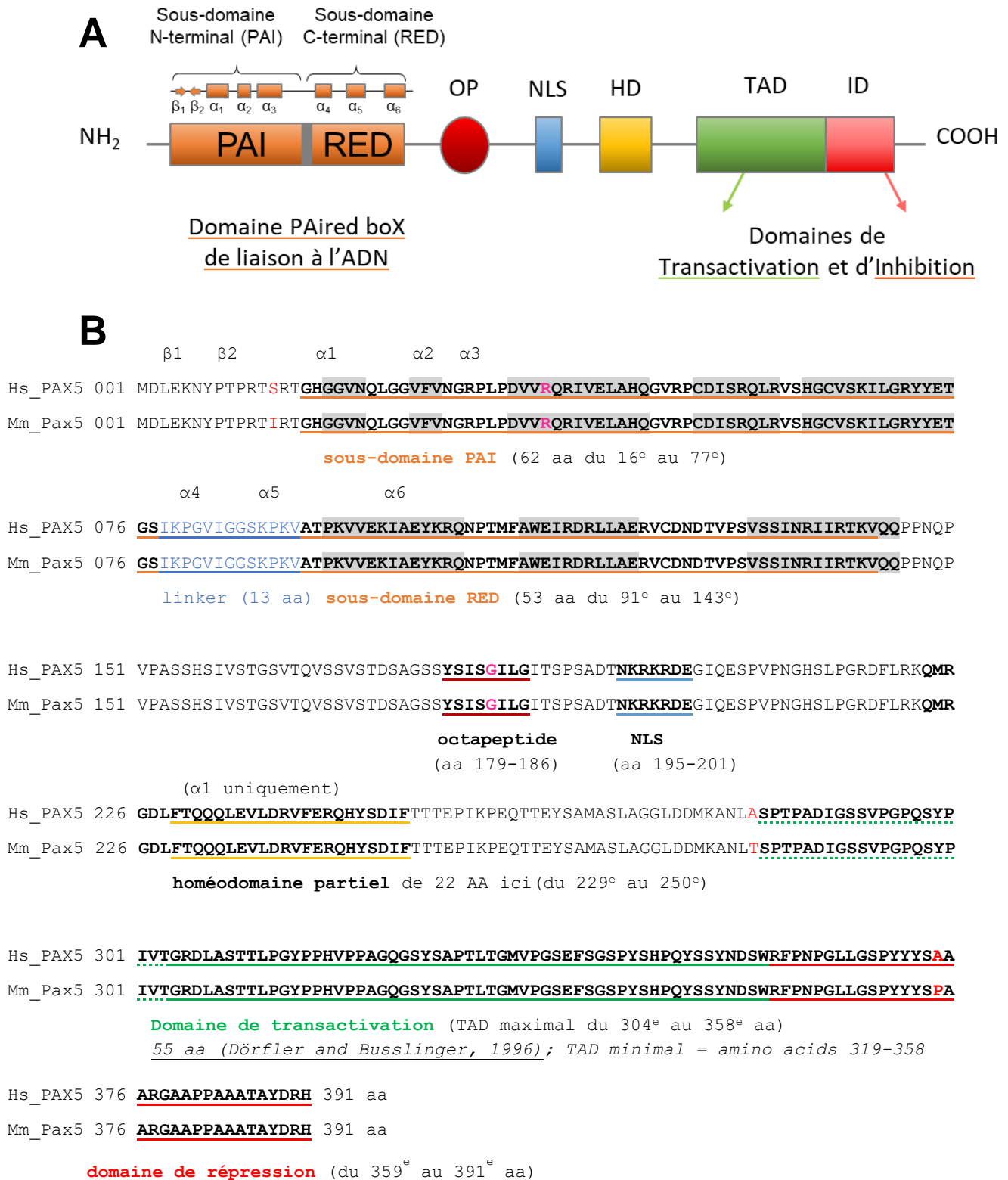


Figure 1.06 : Structure et séquence protéiques de Pax5/PAX5 chez la souris et l'homme.

**A.** Illustration schématique de l'organisation structurale de la protéine Pax5/PAX5.

**B.** Séquences d'acides aminés de la protéine PAX5 de l'homme (ligne supérieure) et Pax5 de la souris (ligne inférieure). Les deux résidus R38 et G183 qui sont affectés par des mutations germinales (R38H et G183S) sont indiqués en rose. Hs : *Homo sapiens*. Mm : *Mus musculus*. Illustrations créées par Laura Jamrog.

## 1.3. Structure moléculaire et fonctions des domaines protéiques

Les protéines PAX sont caractérisées par un domaine de liaison à l'ADN situé à l'extrémité N-terminale, appelé *PAired boX domain*. Elles possèdent un signal de localisation nucléaire (NLS) dans leur domaine central, ce qui leur permet d'être adressées au noyau pour y exercer leur fonction de facteurs de transcription. S'y ajoutent un domaine octapeptide (OP), un homéodomaine (HD) de la classe Paired, ainsi qu'une région accomplissant des fonctions de régulation transcriptionnelle des gènes cibles, comprenant un domaine activateur (TAD) et éventuellement un inhibiteur (ID) (Figure 1.06A). Tous ces domaines sont fortement conservés au cours de l'évolution et peuvent être impliqués dans des interactions avec les partenaires protéiques (Cobaleda et al., 2007 ; Medvedovic et al., 2011 pour revues).

Puisque j'ai étudié les fonctions de PAX5 au cours de ma thèse, une emphase est faite sur PAX5 dans les sections concernées. Comme mentionné précédemment, le gène *Pax5* fait partie du groupe II auquel appartient ses paralogues *Pax2* et *Pax8*. *Pax5* code une protéine composée de **391 acides aminés** (aa) chez la souris et l'homme. De l'extrémité N-terminale à l'extrémité C-terminale, *Pax5* contient le caractéristique domaine de liaison bipartite PAIRED (aa 16 à 143), du motif octapeptide (aa 179 à 186), de la séquence de localisation nucléaire (NLS), d'un homéodomaine partiel (aa 229 à 250) conservant la première hélice *alpha* sur trois, et enfin le module de régulation transcriptionnel C-terminal comportant le domaine de transactivation (aa 304 à 359) suivi du domaine d'inhibition (aa 359 à 391) (Figure 1.06B, créée d'après les alignements de séquences publiés de la revue de Bürglin & Affolter, 2016).

### 1.3.1. Le signal de localisation nucléaire (NLS)

Le signal de localisation nucléaire (*Nuclear Localization Signal*) de *Pax5* consiste en une courte séquence centrale de 7 acides en positions 195 à 201: **NKRKRDE** (Figure 1.05), initialement définie suite à l'identification de l'interaction entre *Pax5* et l'importine alpha1 (Rch1alpha), aussi appelée KPNA2 (karyopherin subunit alpha 2). La région C-terminale de l'importine alpha1 se lie à la séquence NLS de *Pax5*, permettant le transport de *Pax5* dans le noyau (Kovac et al., 2000).

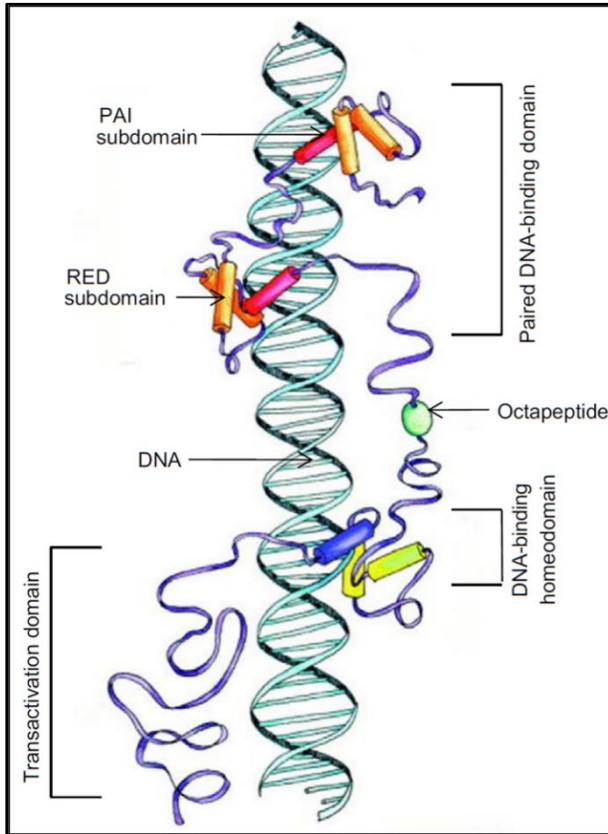


Figure 1.07 : Structure protéique et domaines fonctionnels des facteurs de transcription de la famille Pax chez la souris, en interaction avec l'ADN.

Toutes les protéines Pax comportent à leur extrémité amino(N)-terminale un domaine de liaison à l'ADN (schématisé en bleu turquoise), appelé domaine Paired (« Paired DNA-binding domain »). Le domaine *Paired* est composé de deux sous-domaines dénommés « PAI » en N-terminal et « RED » en C-terminal, qui sont chacun composés de trois hélices alpha réalisant un motif hélice-coude-hélice. La troisième hélice (C-terminale) de chaque sous-domaine Paired est montrée en contact avec le sillon majeur de l'ADN cible. À leur extrémité C-terminale, les protéines Pax contiennent également un domaine de transactivation. Certains membres de la famille Pax possèdent un motif octapeptide (en vert) et/ou un homéodomaine de classe paired (« DNA-binding homeodomain ») composé d'un motif

hélice-coude-hélice à trois hélices alpha lorsqu'il est complet. Cette figure a été créée par Nicole Boitos, publiée initialement par Chi and Epstein (Trends Genet., 2002) et adaptée, avec permission, par Blake and Ziman (Development, 2014).

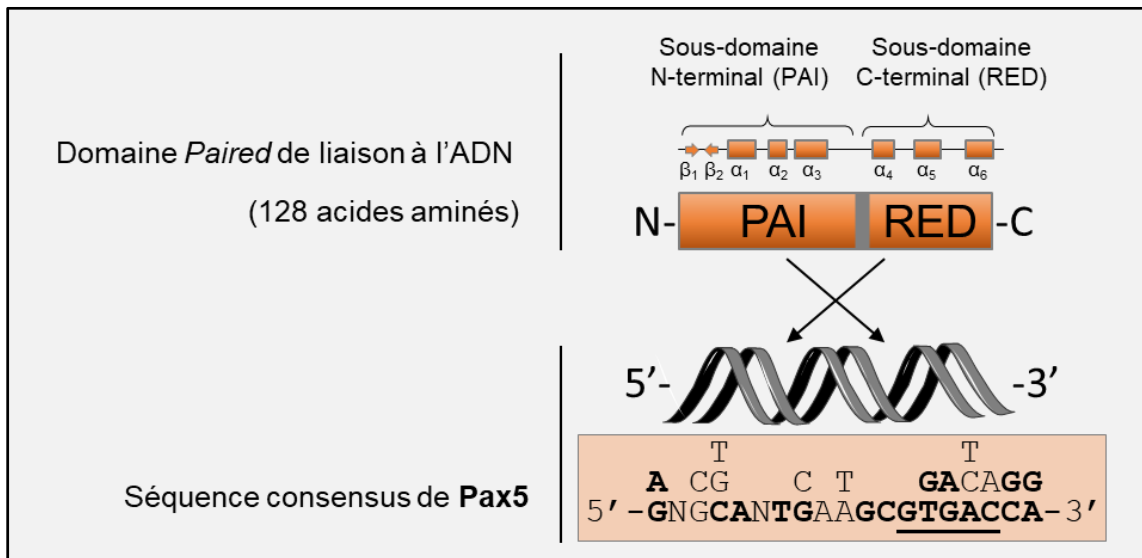


Figure 1.08 : Structure du domaine PAIRED et séquence de reconnaissance de Pax5.

La séquence consensus générale des protéines Pax est indiquée en gras. Le domaine PAIRED se compose des sous-domaines N-terminal PAI et C-terminal RED qui reconnaissent respectivement la moitié 3' et l'extrémité 5' de la séquence consensus. Figure adaptée par Laura Jamrog d'après les données de (Czerny et al., 1993) ; (Czerny & Busslinger, 1995) ; (W. Xu et al., 1995) ; (H. E. Xu et al., 1999). Le motif pentanucléotidique « GGAAC » souligné a été initialement décrit par Chalepakidis *et al.*, (1991), comme étant le motif central essentiel de la séquence consensus auquel se lie le domaine PAIRED.

### 1.3.2. Le domaine PAIRED, domaine de liaison à l'ADN

Le domaine PAIRED, composé de 128 acides aminés exerce une activité de liaison à l'ADN (Treisman et al., 1991 ; Czerny et al., 1993). Des analyses biochimiques *in vitro* (Czerny et al., 1993) et cristallographiques par rayons X (W. Xu et al., 1995) ont montré que ce domaine présentait une **structure bipartite** divisée en un sous-domaine amino(N)-terminal (PAI) et un sous-domaine carboxy(C)-terminal (RED) se liant à l'ADN aux niveaux de deux sillons majeurs adjacents (Figures 1.07 et 1.08). Les deux sous-domaines sont reliés par une région de liaison flexible (*linker*), composée de 13 acides aminés chez Pax5 (aa 78 à 90 : IKPGVIGGSKPKV) permettant d'établir des contacts étendus avec l'ADN via un sillon mineur de l'ADN (H. E. Xu et al., 1999). Chaque sous-domaine se présente sous forme d'une région tri-hélicoïdale dont la structure ressemble à celle d'un homéodomaine et chacun contient ainsi un motif hélice-coude-hélice (*Helix-Turn-Helix*, *HTH motif*) à trois hélices alpha (W. Xu et al., 1995 ; H. E. Xu et al., 1999). Le motif hélice-coude-hélice est une ancienne structure similaire à celle trouvée dans tous les homéodomaines. C'est la dernière hélice alpha, en position C-terminale de chaque sous-domaine, qui est responsable de tous les contacts avec les bases d'ADN des sillons majeurs (Czerny & Busslinger, 1995 ; Figure 1.07). Le sous-domaine N-terminal (PAI) est constitué de deux feuillets bêta antiparallèles suivis par trois hélices alpha alors que le sous-domaine C-terminal (RED) se compose seulement de trois hélices alpha (Figures 1.07 et 1.08). De plus, ces analyses d'interaction au sein du complexe ADN-PAIRED ont également montré que les sous-domaines PAI et RED reconnaissent une séquence d'ADN consensus elle-même divisée en deux parties distinctes (5' et 3') positionnées sur un même côté de l'hélice d'ADN. Ainsi le sous-domaine N-terminal PAI se lie à la moitié en 3' de la séquence de reconnaissance tandis que le sous-domaine C-terminal RED reconnaît le demi-site de liaison en 5' de la séquence de reconnaissance (W. Xu et al., 1995 ; Figure 1.08). Des études de liaison à l'ADN basées sur des analyses mutationnelles *in vitro* du domaine PAIRED et de sa séquence de reconnaissance ont permis l'identification d'une séquence consensus non-palindromique **TCACGC/G** (« core ») reconnue par toutes les protéines Pax, avec des variations mineures entre les différents Pax homologues (Czerny et al., 1993; Czerny & Busslinger, 1995; J. Epstein et al., 1994).

Par ailleurs, la plupart des protéines Pax (dont Pax5) ont la capacité de transactiver la transcription de gènes en se liant aussi bien à un promoteur qu'à un *enhancer* (Czerny & Busslinger, 1995).

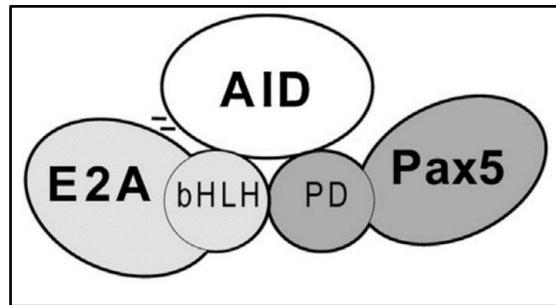


Figure 1.09 : Schéma du complexe en partie formé par AID, E2A et PAX5 recrutés au niveau du locus *IgH* dans les lymphocytes B.

bHLH, *basic helix-loop-helix domain*; PD, *PAIRED domain*. Figure issue de Hauser et al., 2016.

Le domaine N-terminal PAI est essentiel à l'activité de liaison aux gènes cibles et se lie à la moitié en 3' de la séquence de reconnaissance par des interactions extensives avec le squelette phosphate et les bases des sillons majeur et mineur de l'ADN (Czerny & Busslinger, 1995). Notamment, les acides aminés déterminant la spécificité de séquences cibles des protéines Pax5 se situent dans le sous-domaine N-terminal PAI : à la fin de la deuxième hélice alpha, dans le linker et au début de la troisième hélice (résidus en positions 42, 44 et 47, (Czerny & Busslinger, 1995). Les protéines Pax peuvent parfois se lier uniquement via un seul des sous-domaines. Par exemple, une forme tronquée du domaine PAIRED de Pax5 (comportant une délétion des 36 derniers acides aminés de RED) maintient la capacité de se lier à certaines séquences cibles, mais avec une affinité plus faible (Czerny et al., 1993). En effet, la liaison concomitante aux deux moitiés de la séquence consensus par les deux sous-domaines PAI et RED confèrent une affinité et/ou une spécificité plus importante-s (*in vitro*, Czerny et al., 1993).

Par ailleurs, il existe une spécificité de séquence de liaison pour chaque sous-domaine des protéines Pax : les sous-domaines PAI et RED se lient à des motifs de séquence d'ADN spécifiques et distincts et peuvent ainsi réguler des ensembles de gènes distincts. Les sites de liaison de Pax5 peuvent considérablement varier par rapport à la séquence consensus mentionnée précédemment, afin d'augmenter l'affinité de liaison d'un sous-domaine au détriment de l'habileté de fixation du second sous-domaine. Dans le cas de Pax5, il existe ainsi deux classes (I et II) de sites de reconnaissance (Czerny et al., 1993).

Le domaine PAIRED, outre ses propriétés de liaison à l'ADN, peut aussi intervenir dans des interactions protéiques qui modulent l'activité transcriptionnelle des protéines Pax. Ainsi, le domaine PAIRED de Pax5 recrute directement l'enzyme AID (*Activation-Induced-cytidine Deaminase*) et interagit avec le domaine bHLH (*basic helix-loop-helix*) du facteur de transcription TCF3 (E2A) dans le cadre du processus de réarrangements des chaînes des immunoglobulines (Figure 1.09 ; Hauser et al., 2016).





### 1.3.3. L'homéodomaine des homéoprotéines de la classe *paired*

Les protéines Pax possèdent un second domaine de liaison à l'ADN, le motif homéodomaine (HD), très conservé au cours de l'évolution et présent exclusivement chez les Eucaryotes (Chi & Epstein, 2002). Chez les animaux, il existe seize classes de protéines à HD (Bürglin & Affolter, 2016), dont les protéines codées par les gènes de développement *Hox* (Chi & Epstein, 2002).

L'HD est codé par la séquence homéoboîte (*homeobox*) composée d'environ 180 paires de bases. Il forme un domaine globulaire composé d'environ 60 acides aminés comportant un motif central (signature) d'acides aminés hydrophobes. L'HD complet adopte une structure tri-hélicale incluant le motif hélice-coude-hélice (*helix-turn-helix*) agrémenté d'une troisième hélice alpha (T. R. Bürglin, 1994 ; Gehring et al., 1990, 1994 ; Thomas R. Bürglin & Affolter, 2016). Les deux premières hélices (amino(N)-terminales) sont positionnées de manière antiparallèle tandis que l'axe de la troisième l'hélice (carboxy(C)-terminale) est perpendiculaire à l'axe des deux premières (Figure 1.03). L'hélice C-terminale est plus longue et riche en résidus arginine (Arg ; R) et lysine (Lys ; K), qui sont des acides aminés polaires chargés positivement établissant des liaisons hydrogènes avec l'ADN, chargé négativement. Les acides aminés hydrophobes conservés au centre de l'hélice C-terminale accroissent sa compaction. Les protéines Pax des groupes III (Pax3/7) et IV (Pax4/6) contiennent un HD complet (composé de 60 aa) tandis que Pax2, Pax5 et Pax8 (groupe II) n'ont qu'un HD tronqué, composé d'environ 22 à 27 aa (HD de Pax5 : **FTQQQLEVLDRVFERQHYSDF** ; Figure 1.05), correspondant seulement à la première hélice alpha. Pour rappel, Pax1 et Pax9 (groupe I) ne possèdent pas du tout d'HD (Tremblay & Gruss, 1994).

Un HD complet se fixe à l'ADN-cible principalement via son hélice C-terminale s'alignant dans un sillon majeur de l'ADN et son extrémité N-terminale non structurée qui s'aligne dans un sillon mineur (Stuart et al., 1994). Les contacts de l'hélice C-terminale au sein du sillon majeur de l'ADN apportent une contribution plus importante à la spécificité de l'interaction entre l'HD et l'ADN (d'où l'appellation de *recognition helix*) tandis que les contacts de l'hélice N-terminale au sein du sillon mineur de l'ADN aident à positionner le complexe, stabilisent et renforcent la fixation (Bürglin & Affolter, 2016).

La plupart des HD reconnaissent des sites de liaison composés de séquences répétées inversées du motif central **TAAT**, avec préférentiellement un espacement de 3 nucléotides (Czerny & Busslinger, 1995).

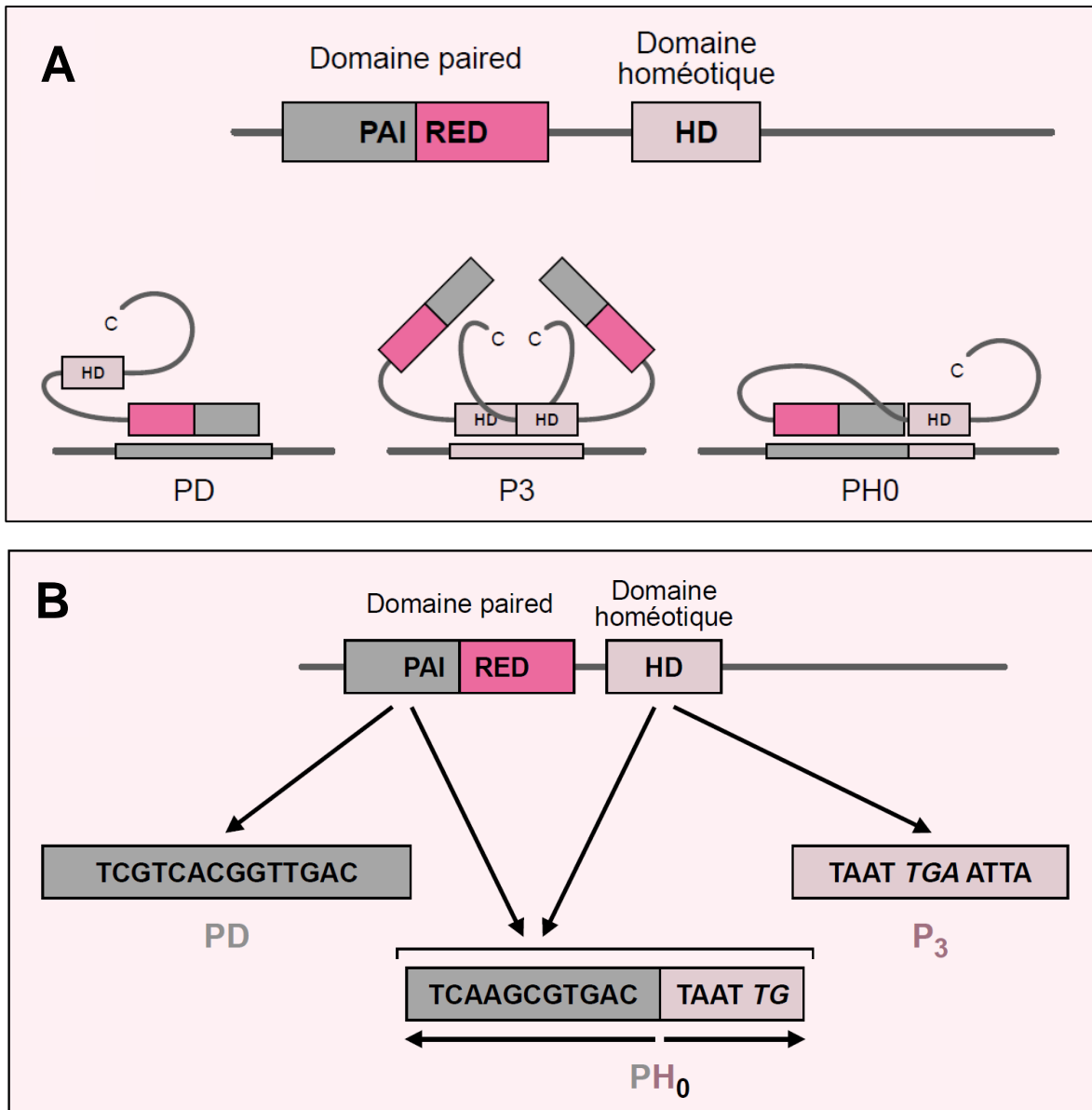


Figure 1.10 : Modalités de liaison à l'ADN de protéines à domaine PAIRED et homéodomaine de classe Paired.

**A.** Chez les mammifères, les protéines Pax contiennent toutes le domaine PAIRED, et éventuellement un homéodomaine (HD) complet ou partiel. Ces facteurs de transcription peuvent se fixer à l'ADN par le biais de leur sous-domaine PAI, par le biais d'un dimère d'homéodomaines ou par le biais d'une interaction entre le domaine PAIRED et l'homéodomaine. Des protéines, telles que Pax5, qui n'ont qu'un domaine PAIRED et un homéodomaine résiduel, utilisent à la fois leur sous-domaine PAI et leur sous-domaine RED. Des protéines Pax contenant un homéodomaine s'associent en dimères de manière coopérative pour se lier au site palindromique P3.

**B.** Le domaine PAIRED et l'homéodomaine des protéines Pax à homéodomaine ont la capacité de se lier à des sites très spécifiques (sites PD et P3). Ils peuvent aussi coopérer pour se lier aux sites PH0 qui représentent la juxtaposition sans espacement d'un site PD et d'un motif TAAT en orientation inversées.

Figures issues de Desplan, 1997.

- **Dimérisation des homéodomaines des protéines Pax**

Ne concernant pas Pax5, les HD complets des protéines Pax s'associent en homo- et hétérodimères de manière coopérative sur des sites palindromiques d'ADN de type **5'-TAAT(N)2-3ATTA-3'** (Czerny & Busslinger, 1995; Wilson et al., 1993). Ce type de dimérisation en coopération permet aux HD (des protéines Pax) de reconnaître des séquences qui ne le sont pas par d'autres HD (notamment d'autres homéoprotéines de la classe *paired*). De surcroît, la propriété d'homo- et d'hétérodimérisation des protéines Pax médiée par leurs HD permet aux différentes protéines Pax d'interagir entre elles ce qui peut correspondre à un mécanisme important pour la régulation de l'expression des gènes cibles des protéines Pax (Czerny & Busslinger, 1995; Wilson et al., 1993).

- **Coopération des deux domaines de liaison à l'ADN des protéines Pax**

Le domaine PAIRED et l'homéodomaine de la classe *paired* ont la capacité de se lier à l'ADN et peuvent ainsi tous les deux réguler des cibles transcriptionnelles communes ou spécifiques. Néanmoins, tous les domaines PAIRED des protéines homologues Pax reconnaissent des séquences d'ADN consensus fortement apparentées et la plupart des homéodomaines reconnaissent également des sites très similaires (séquences répétées TAAT). Afin de pouvoir cibler différentes séquences promotrices, les protéines Pax peuvent utiliser de multiples combinaisons de leurs (3) motifs hélice-coude-hélice (Jun & Desplan, 1996). Ainsi, le domaine PAIRED et l'HD peuvent fonctionner indépendamment l'un de l'autre ou bien peuvent coopérer, ce qui permet d'ajouter de la complexité fonctionnelle (entre les protéines Pax homologues) et d'augmenter la diversité de cibles transcriptionnelles (Figure 1.10). La séquence consensus résultante est **AATTA-GTCACGC** ('PHO'), reconnue par les protéines Pax à HD complet (Apuzzo & Gros, 2007; Jun & Desplan, 1996).

Puisque Pax5 (tout comme Pax2 et Pax8) possède seulement un HD résiduel, l'intégrité de son domaine PAIRED en tant que domaine de liaison à l'ADN est primordiale pour assurer sa fonction de facteur de transcription (Desplan, 1997; Jun & Desplan, 1996).

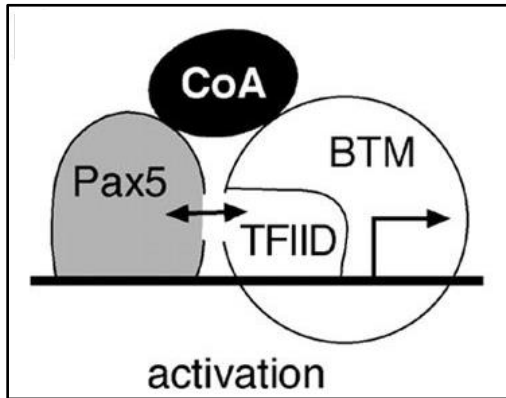


Figure 1.11 : Modèle de recrutement du complexe général de transcription TFIIID par Pax5 pour l'activation de ses gènes cibles.

TFIIID, *Transcription factor II D*; CoA, *coactivator protein-s*; BTM, *basal transcription machinery*. Figure issue de Eberhard et al., 2000.

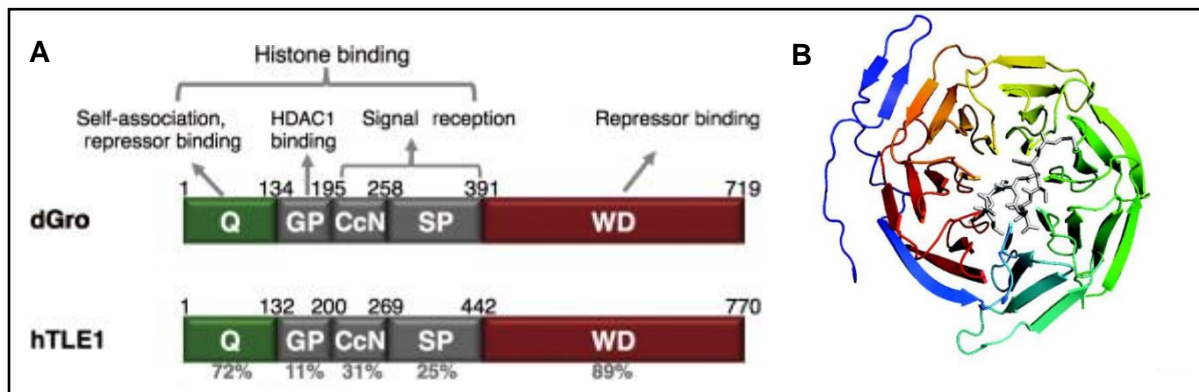


Figure 1.12 : Organisation structurale des corépresseurs de la famille Grg/TLE.

**A.** La comparaison de la séquence d'acides aminés de la protéine Groucho de la Drosophile (dGro) avec celle de la protéine humaine hTLE1 met en évidence une organisation structurale conservée, composée d'un domaine riche en glutamine (Q), suivi d'un domaine riche en glycine et proline (GP), d'un domaine CcN {contenant sa séquence de localisation nucléaire}, d'un domaine riche en sérine et proline (SP) et d'un domaine de séquences répétées se terminant par le dipeptide tryptophane–aspartate (WD). Les pourcentages d'identité de séquence des domaines d'hTLE1 sont indiqués par rapport à dGro. Figure issue de Turki-Judeh & Courey, 2012.

**B.** Représentation de la liaison de motifs peptidiques Eh1 (bâtonnets gris) à la région C-terminale du domaine WD (rubans multicolores). Le motif Eh1 se lie sur une surface à travers le pore central de l'hélice  $\beta$ . Figure issue de Jennings & Ish-Horowicz, 2008.

Outre cette activité de liaison à l'ADN, l'HD des protéines Pax peut aussi servir d'interface pour des interactions protéiques. Ainsi, l'HD de Pax3 se fixe aux corépresseurs transcriptionnels Daxx (*death domain associated protein*) et HIRA (*histone cell cycle regulator*) (Hollenbach et al., 1999; Magnaghi et al., 1998). Par ailleurs, en interagissant avec l'HD partiel de Pax5, Daxx a la capacité de jouer un double rôle, soit de corépresseur, soit de coactivateur transcriptionnel (Emelyanov et al., 2002). A l'instar de l'homéodomaine complet de Pax6, l'HD partiel de Pax5 interagit avec Rb1 (*RB transcriptional corepressor 1*) et TBP (*TATA-box binding protein*). La TBP est la sous-unité de liaison à l'ADN du complexe général de transcription TFIID (*Transcription factor II D*), ce qui peut ainsi faciliter le recrutement du complexe général de transcription TFIID au niveau de séquences promotrices ciblées par Pax5 (Cvekl et al., 1999; Eberhard & Busslinger, 1999; Figure 1.11).

#### 1.3.4. Le motif octapeptide (OP)

A l'exception de Pax4 et Pax6 (groupe IV), les sept autres protéines Pax des Mammifères ont conservé un motif octapeptide central (Burri et al., 1989), situé entre le domaine PAIRED et l'homéodomaine (Figure 1.05). Le motif octapeptide (OP) est apparenté au motif de répression Eh1 de la famille *engrailed homology 1* (*Eh1*) constitué d'un heptapeptide Eh1 (FSIDNIL), grâce auquel les facteurs de transcription de type Engrailed exercent leur activité de répression transcriptionnelle. Le motif OP des protéines Pax se caractérise par la présence de résidus hydrophobes espacés comme dans le motif Eh1 ; ainsi l'OP de Pax5 se compose de l'enchaînement **YSISGILG** (aa 179 à 186 ; Figure 1.05). Concernant les autres protéines Pax, il se compose de l'enchaînement consensus d'acides aminés suivants : (H/Y)-S-(I/V)-(N/S)-G/N-(I/L)-L-G (Bürglin & Affolter, 2016). Les motifs OP sont impliqués dans les interactions protéiques de type répressif. En effet, ils fonctionnent comme des motifs inhibiteurs de la transcription en recrutant des corépresseurs protéiques de la famille *Groucho-related genes/Transducin-Like Enhancer of split* (*Grg/TLE*) dont le membre fondateur est la protéine Groucho (Gro) de la Drosophile (Eberhard et al., 2000; Mailhos et al., 1998).

Chez les Mammifères, la famille Grg/TLE comprend jusqu'à sept membres (TLE1-7).

Ces protéines sont notamment caractérisées par un domaine C-terminal très conservé (Figure 1.12A) qui se compose de répétitions en tandem du motif WD40 (ou répétitions bêta(β)-transducine, de quarante acides aminés se terminant par le dipeptide tryptophane–aspartate (W-D), d'où elles tiennent leur nom). Ces répétitions WD40 se replient en hélice bêta à sept lames, constitués de quatre à huit feuillet β par lame, ainsi formant le domaine WD40 toroïdal avec un sillon central au niveau duquel se lie le motif Eh1 (Figure 1.12B ; Jennings et al., 2006). Les deux motifs (WD40 et EH1) se lient au sommet de l'embouchure du canal central et Eh1 forme une courte hélice alpha amphipathique qui se lie à ce renforcement hydrophobe.

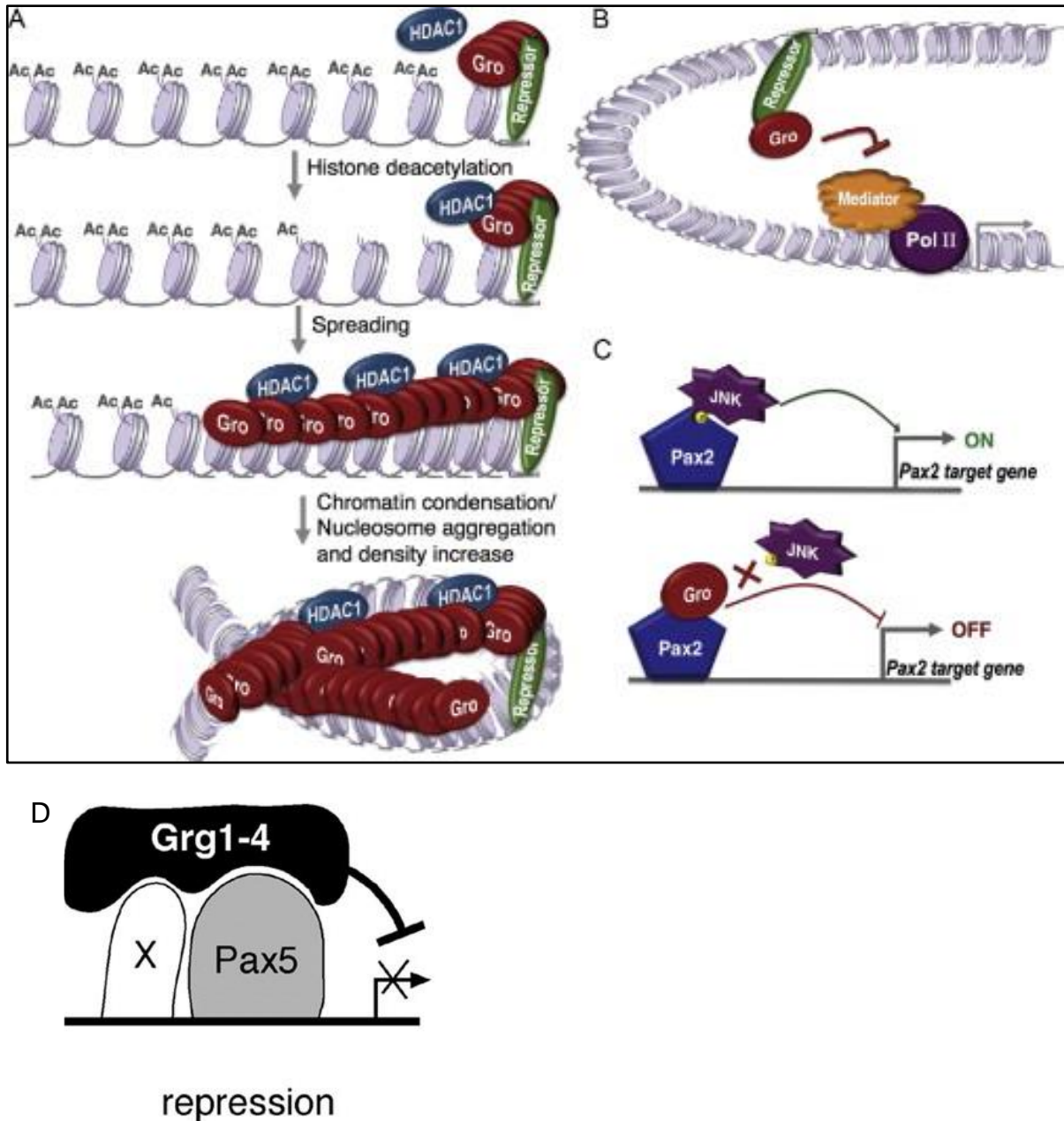


Figure 1.13 : Modèles de répression médiée par le recrutement des protéines Grg/TLE.

**A.** Les protéines Grg/TLE (Gro sur la figure) s'oligomérisent, recrutent des désacétylases d'histones (HDAC1), se lient à des histones désacétylés et induisent la condensation de la chromatine ciblée (ou augmentent la densité des nucléosomes). **B.** Les protéines Grg/TLE pourraient interagir avec un composant (Mediator) de la machinerie basale de transcription, ainsi interférant avec la formation ou la fonction du complexe de pré-initiation de la transcription. **C.** Les protéines Grg/TLE pourraient masquer les domaines d'activation des activateurs transcriptionnels. Ci-dessus en exemple, la liaison de Gro à Pax2 empêche la phosphorylation du domaine d'activation de Pax2, ainsi empêchant l'activation de Pax2. Figures A-C issues de Turki-Judeh & Courey, 2012. **D.** Modèle de répression opérée par Pax5 et médié par les protéines Grg. La collaboration d'un second FT (X) pourrait être requise pour spécifier l'interaction avec un des membres de la famille Grg/TLE et localiser le recrutement à un promoteur précis. Figure 1.13D (identique à Figure 1.14B) issue de Eberhard et al., 2000.

Les protéines Grg/TLE utilisent de multiples mécanismes de répression (Figure 1.13). Elles sont capables de s'oligomériser, pouvant renforcer leur effet répressif et/ou l'exercer à longue distance. Elles peuvent recruter des désacétylases d'histones (HDAC) et se lier aux histones désacétylés pour ensuite induire la condensation de la chromatine des gènes cibles à réprimer ou augmenter la densité des nucléosomes (Figure 1.13A). Elles peuvent interférer avec la machinerie basale de transcription (Figure 1.13B). En se fixant aux motifs OP/EH1, les protéines Grg/TLE pourraient également empêcher le recrutement de médiateurs par l'homéodomaine des protéines Pax, masquer les domaines d'activation des protéines Pax (Figure 1.13C) ou bloquer ces domaines d'activation dans une conformation inactive (Bürglin & Affolter, 2016; Turki-Judeh & Courey, 2012). En particulier, le corépresseur TLE4 (Grg4) interagit directement avec le motif OP de Pax1, Pax3, Pax6 et Pax2, Pax5 et Pax8. Le motif OP se fixe plutôt à la région centrale riche en sérine et proline (SP) de TLE4 pour médier l'activité transcriptionnelle répressive de Pax5 (Eberhard et al., 2000). Il est possible que Pax5 doive collaborer avec un second facteur transcription, comme c'est le cas avec PU.1, pour pouvoir recruter les corépresseurs Grg/TLE à un promoteur spécifique (Figure 1.13D ; Linderson et al., 2004).

### 1.3.5. Domaines de transactivation (TAD) et d'inhibition (ID)

L'activité transcriptionnelle des protéines Pax est déterminée par un module de régulation à l'extrémité C-terminale, comportant des séquences activatrices et inhibitrices (Figure 1.05). Cette région est riche en résidus proline (P), sérine (S) et thréonine (T), d'où la dénomination du domaine PST. Ce module de régulation de la transcription se compose d'un domaine de transactivation (TAD) associé dans certains cas à un domaine d'inhibition (ID). La région C-terminale des protéines Pax est la partie la plus divergente parmi les membres de la famille Pax. Le module de régulation des Pax du groupe II (Pax2/Pax5/Pax8 et leurs orthologues) est structurellement et fonctionnellement conservées au cours de l'évolution (Bürglin & Affolter, 2016; Dörfler & Busslinger, 1996) alors que les séquences de Pax1 et Pax9 du Groupe I divergent principalement de par leurs extrémités C-terminales (Sivakamasundari et al., 2018), à l'instar de la divergence de séquence C-terminale entre PAX4 et PAX6 pourtant appartenant au même groupe IV (Lorenzo et al., 2017).

Le module de régulation du groupe II (dont Pax5) se compose d'un domaine transactivation de 55 acides aminés (TAD de Pax5, aa 319 à 358 : **GRDLASTTLPGYPPHVPPAGQGSYSAPTLTGMVPGSEFSGSPYSH PQYSSYND SW**) et d'un domaine de répression de 33 acides aminés (ID de Pax5/PAX5, aa 359 à 391 : **RFPNPGLLGSPYYYS(P/A)AARGAAPPAAATAYDRH**) (Dörfler & Busslinger, 1996).





Le domaine TAD peut exercer sa fonction activatrice à partir de la position d'un promoteur ou celle d'un *enhancer*. La fonction de transactivation de Pax2/Pax5/Pax8 subit une forte cis-répression exercée par le domaine d'inhibition adjacent, à l'extrémité C-terminale (Dörfler & Busslinger, 1996). Similairement, les régions C-terminales de PAX3 et PAX7 sont riches en PST et leur confèrent une activité de transactivation. Notamment, le TAD de Pax3 comprend les 78 derniers acides aminés de l'extrémité C-terminale (Boudjadi et al., 2018; Chalepakis et al., 1994). À l'inverse des autres membres de la famille Pax, les TAD de PAX3 et PAX7 sont fortement inhibés par des domaines de cis-répression transcriptionnelle localisés dans la région N-terminale (les 90 premiers acides aminés), incluant la première moitié du domaine PAIRED (57 premiers acides aminés) (Bennicelli et al., 1999; Chalepakis et al., 1994). Ces domaines de cis-répression de PAX3 et PAX7 sont structurellement et fonctionnellement conservés alors que les régions C-terminales de PAX3 et PAX7 présentent une divergence structurelle qui expliquerait en partie leurs différences fonctionnelles (Bennicelli et al., 1999). Notamment, PAX7 a l'unique particularité de posséder un motif OAR à l'extrémité C-terminale (*cf section suivante*) pouvant potentialiser sa fonction de transactivation tout en étant plus susceptible à la cis-répression N-terminale (Bennicelli et al., 1999). Concernant PAX4 uniquement, les domaines de transactivation et d'inhibition se chevauchent au sein de la région C-terminale (Fujitani et al., 1999).

Les fonctionnalités transcriptionnelles de certaines protéines Pax sont impactées par leur concentration. En effet, certaines protéines Pax présentent une activité transcriptionnelle dose-dépendante dite « en cloche » : l'activation transcriptionnelle s'exécute à de faibles concentrations en protéines tandis qu'elle se réduit à de fortes concentrations protéiques. C'est par exemple le cas pour Pax3 (Chalepakis et al., 1994), PAX7 (Bennicelli et al., 1999) et PAX6 (Glaser et al., 1994). Ainsi, de faibles concentrations en Pax3 activent sa transactivation alors que de plus fortes concentrations inhibent son activité (Chalepakis et al., 1994). De surcroît, PAX3 et PAX7 présentent des différences d'activité de transactivation dose-dépendantes, probablement due à leur divergence structurelle dans la région C-terminale (Bennicelli et al., 1999).

Chez certains membres, tels que PAX6 et les membres du groupes II, PAX2/PAX5/PAX8, le module TAD/ID peut fonctionner indépendamment du type cellulaire où ils s'expriment (Czerny & Busslinger, 1995; Dörfler & Busslinger, 1996). En revanche, l'activité du TAD de PAX4 dépend du type cellulaire alors que son domaine de répression est actif indépendamment du type cellulaire, suggérant qu'il interagit directement avec la machinerie de transcription basale ou avec des corépresseurs ubiquitaires (Fujitani et al., 1999).

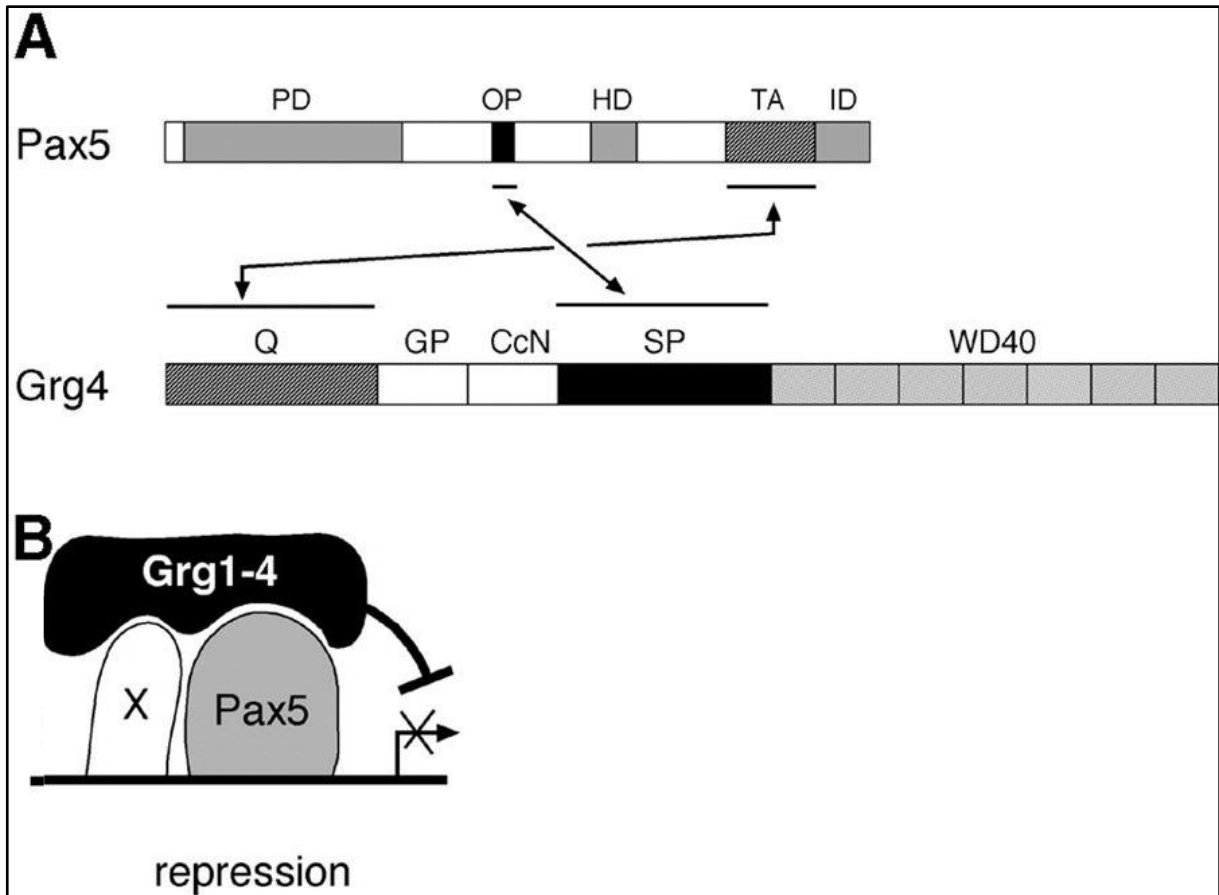


Figure 1.14 : Modèle de recrutement de corépresseurs Grg par Pax5.

**A.** Interactions identifiées entre les motifs structuraux des protéines Pax5 et Grg4. **B.** Rappel du modèle de répression opérée par Pax5 et médiée par les protéines Grg. La collaboration d'un second FT (X) pourrait être requise pour spécifier l'interaction avec un des membres de la famille Grg/TLE et localiser le recrutement à un promoteur précis. Figure issue de Eberhard et al., 2000.

Les facteurs de transcription tels que les protéines PAX n'agissent pas de manière isolée mais dépendent d'interactions protéiques pour réaliser leurs fonctions. Ces interactions fournissent plus de spécificité aux fonctions de régulation d'un FT donné. Les protéines Pax, comme Pax4 et Pax6, peuvent notamment employer leurs modules TAD/ID pour s'adjoindre de partenaires protéiques (Fujitani et al., 1999; Ritz-Laser et al., 2002). Pour exercer son activité de transactivation transcriptionnelle auprès de promoteurs ou d'*enhancers*, PAX6 module ses interactions protéiques en utilisant soit sa région C-terminale riche en PST en entier, soit uniquement son peptide C-terminal (Cooper & Hanson, 2005; Glaser et al., 1994). Pour assurer une régulation transcriptionnelle épigénétique, Pax5 recrute via son TAD des acétyltransférases d'histones (HAT) et des protéines adaptatrices d'HAT (Barlev et al., 2003; Emelyanov et al., 2002; He et al., 2011).

Comme mentionné précédemment, les corépresseurs de la famille Gro/TLE, recrutés par le biais du motif OP de Pax5, régulent négativement l'activité transcriptionnelle de Pax5. Le domaine TAD de Pax5 participe également au co-recrutement de ces corépresseurs (Figure 1.14A ; Eberhard et al., 2000). En effet, Pax5 et ces corépresseurs utilisent deux zones de contact pour interagir et cette coopération est indispensable à leur liaison : le motif OP de Pax5 interagit avec la région centrale de Tle4, riche en sérine et proline (domaine SP sur les Figure 1.12 et 1.14) tandis que son domaine TAD interagit avec la région N-terminale de Tle4, riche en glutamine (domaine Q sur les Figure 1.12 et 1.14). En conséquence de cette interaction, Pax5 va induire un changement d'état de phosphorylation de Tle4 (Eberhard et al., 2000).

### 1.3.6. Le motif OAR (Otp/Aristaless/Rax)

Pax7 est le seul membre de la famille Pax à avoir conservé un motif OAR (ou domaine aristaless) à l'extrémité C-terminale, correspondant à une séquence composée de 14 acides aminés (LSTQRRMKLGEHSA ; Bürglin & Affolter, 2016) initialement identifiée dans les facteurs de transcription à homéodomaine Otp (*Orthopedia*), Aristaless et Rax (*Retina and anterior neural fold homeobox*), d'où il tient son nom (Furukawa et al., 1997). Ce motif caudal est aussi nommé *paired-type homeodomain tail (PHT)*, *paired tail* ou *C-peptide domain* (Bürglin & Affolter, 2016; Vorobyov & Horst, 2006). Il est présent dans plus d'une trentaine de protéines contenant un homéodomaine de la classe *paired*, mais sa fonction est peu décrite. Il pourrait moduler la fixation à l'ADN et l'activité transcriptionnelle par un mécanisme impliquant une interaction intramoléculaire entre ce motif et l'homéodomaine se liant à l'ADN. Il pourrait aussi interagir avec des cofacteurs pour moduler la spécificité de cibles des facteurs de transcription à homéodomaine (Bürglin & Affolter, 2016 ; Furukawa et al., 1997). Ce motif OAR peut également établir des interactions fonctionnelles avec le motif OP. Il peut soit réprimer soit promouvoir la transcription en fonction du contexte cellulaire (Buescher & El-Hodiri, 2019).

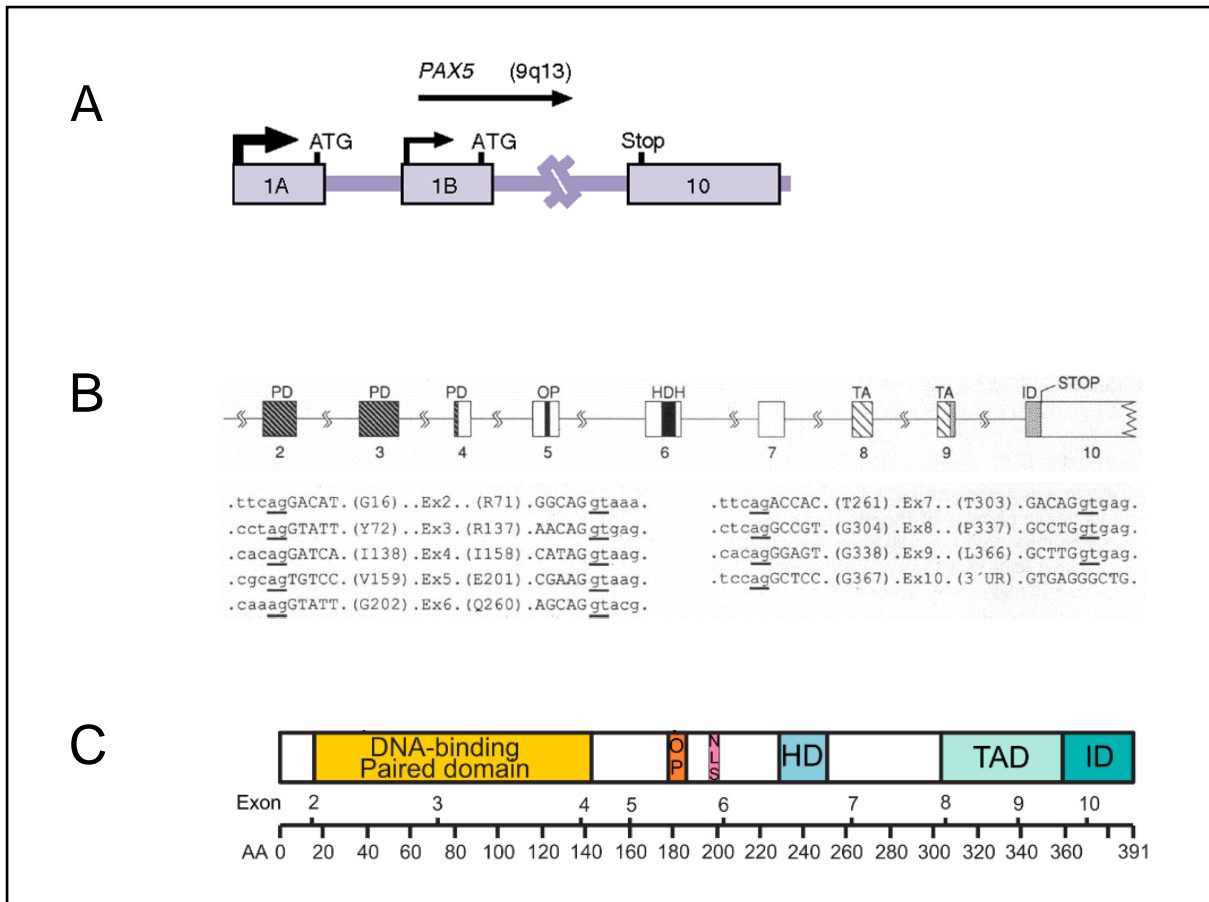


Figure 1.15 : Structures génomique et protéique de PAX5.

**A.** Le codon initiateur de *Pax5* se situe à la fin de l'exon 1 alternatif alors que le codon stop se situe au début de l'exon 10. **B.** Les séquences exoniques sont indiquées en lettres majuscules et les séquences introniques en lettres minuscules. Les dinucléotides invariants GT et AG au niveau des jonctions d'épissage 5' et 3' sont soulignés. Les acides aminés codés par le premier codon et le dernier codon de chaque exon sont indiqués suivis de leurs positions dans la séquence de PAX5A (BSAP). **C.** Organisation des domaines protéiques de Pax5 et correspondantes entre les exons de *Pax5* et la séquence d'acides aminés (AA) de Pax5. PD, *paired domain*; OP, *conserved octapeptide*; HD ou HDH *homeodomain homology region*; NLS, *nuclear localization sequence*; TA ou TAD, *transactivation domain*; ID, *inhibitory domain*; 3'UR, *3' untranslated region*. Figures A et B issue et adaptée de Busslinger et al., 1996.

En conclusion, selon le contexte cellulaire, la conformation protéique et les cofacteurs recrutés, les protéines Pax peuvent fonctionner sous forme soit d'activateurs, soit de répresseurs de la transcription. Les fonctions transcriptionnelles des protéines Pax sont finement régulées par des variations de combinaisons d'activités d'activation et de répression. Cette diversité moléculaire explique en partie comment les protéines Pax remplissent de multiples fonctions cellulaires et tissulaires.

## 1.4. Organisation génomique et régulation génique de *PAX5*

### 1.4.1. Localisation et structure génique

Le gène humain *PAX5*, d'une taille de **444 kb** (kilo-paires de base), est localisé sur la région p13.2 du bras court du **chromosome 9** (9p13.2 ; Stapleton et al., 1993) alors que le gène murin *Pax5*, d'une taille de **392 kb**, est situé sur le **chromosome 4** (Walther et al., 1991).

Le gène *PAX5* humain présente une structure génomique et une régulation transcriptionnelle identiques à son orthologue murin *Pax5*. De plus, les positions des introns sont strictement conservées entre *PAX5* et ses paralogues *PAX2* et *PAX8* (Busslinger et al., 1996). Le gène *PAX5* est composé de **11 exons** (Figure 1.15). La séquence du gène *PAX5* qui va coder les transcrits s'étend sur approximativement 201 kb et inclut 10 exons dont les deux premiers exons en 5' (exon 1A et exon 1B, alternatifs) sont mutuellement exclusifs lors de leur transcription (Busslinger et al. 1996 ; Figure 1.15). *Pax5* est transcrit similairement à partir de ses deux allèles, à tous les stades de la différenciation lymphoïde B (Fuxa & Busslinger, 2007).

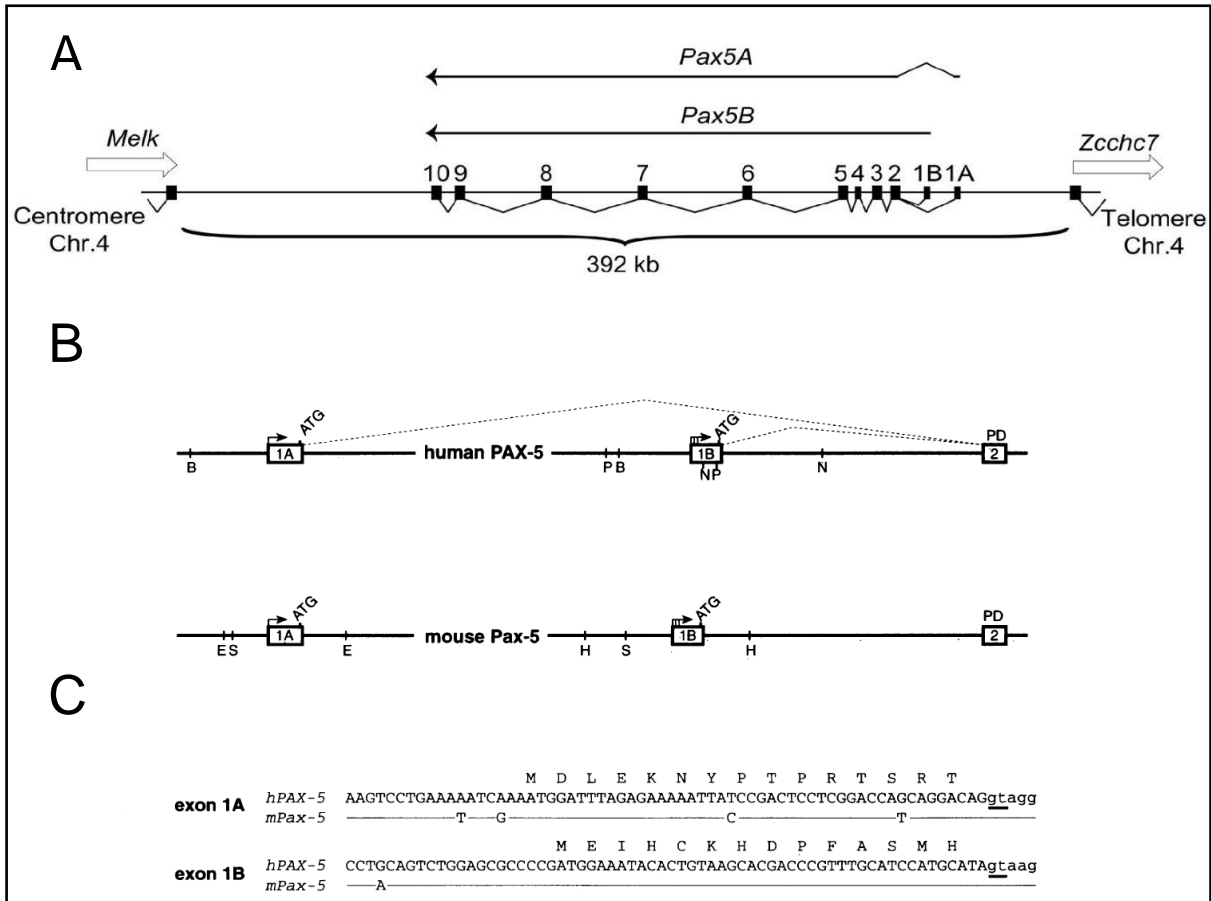


Figure 1.16 : Organisation génomique et transcriptionnelle de PAX5 : transcription de PAX5 de deux exons 1 alternatifs à partir de deux promoteurs distincts.

**A.** La structure génomique de *Pax5* murin est similaire à celle de *PAX5* humain. Le gène *PAX5* (de 392 kb chez la souris et 444 kb chez l'homme) est composé de 11 exons. Les deux premiers exons (exon 1A et exon 1B) sont transcrits alternativement pour générer les deux transcrits *Pax5A* ou *Pax5B* respectivement. Figure 1.16A issue de Cresson C et al. 2018.

**B.** Région 5' des gènes *PAX5* humain et murin. L'exon 1A et l'exon 1B sont éloignés par 6,7 kb chez la souris et par 7,1 kb chez l'homme. La distance entre l'exon 1B et l'exon 2 chez la souris est de 5,3 kb. **C.** Séquences de la partie 3' des deux exons 1 alternatifs. Seuls les acides aminés différents chez la souris par rapport à l'homme sont indiqués en lettres majuscules. Les séquences introniques sont indiquées en lettres minuscules. Le dinucléotide invariant au niveau du site d'épissage est souligné. *Nota Bene* : le transcrit *Pax5B* code une arginine à la place d'une glycine à la position du premier résidu du domaine PAIRED (codé par l'exon 2) du fait de la présence d'une adénosine à l'extrémité 3' de l'exon 1B à la place d'une guanine dans l'exon 1A. Les séquences exoniques sont indiquées en majuscules et les séquences introniques en minuscules. Figures 1.16B et C issue de Busslinger et al., 1996.

### 1.4.2. Éléments de cis-régulation transcriptionnelle de *PAX5*

La transcription du gène *PAX5* est contrôlée par deux promoteurs distincts : il s'agit d'un **promoteur distal** (fonctionnant à proximité d'une séquence TATA-box) et d'un **promoteur proximal** (sans TATA-box) qui vont respectivement initier la transcription de *PAX5* à partir de l'exon 1A ou de l'exon 1B suivi des neuf exons communs (exons 2 à 10) de *PAX5* (Figures 1.16A et B). Le gène *PAX5* peut ainsi coder deux transcrits *PAX5A* et *PAX5B* (d'approximativement 8,7kb) qui diffèrent seulement par leur premier exon alternatif en 5' (exon 1A ou exon 1B respectivement). Par conséquent, les deux isoformes protéiques *PAX5A* et *PAX5B* comportent une différence de séquences à l'extrémité N-terminale : l'exon 1A code 15 acides aminés qui diffèrent entièrement des 14 acides aminés codés par l'exon 1B (Figure 1.16C ; Busslinger et al., 1996).

Ces deux promoteurs alternatifs sont différentiellement régulés au cours du développement et ultérieurement. Le promoteur distal est principalement actif dans la lignée lymphoïde B – il subsiste une faible activité du promoteur distal dans le SNC – alors que le promoteur proximal est actif dans tous les domaines d'expression de *Pax5* : le SNC au cours du développement embryonnaire, la lignée lymphoïde B et les testicules adultes (Adams et al., 1992 ; Busslinger et al., 1996). Bien que les deux promoteurs soient actifs dans le compartiment lymphoïde B, les deux isoformes sont différentiellement exprimées durant la différenciation : la quantité de transcrit *Pax5A* est deux à huit fois plus élevée (selon le stade de différenciation) que celle du transcrit *Pax5B* dans la moelle osseuse de souris (Cresson et al., 2018), et six fois plus élevée dans la rate de souris (Busslinger et al., 1996). C'est donc l'isoforme *PAX5B* qui est essentiellement exprimée dans le SNC au cours de formation tandis que l'isoforme *PAX5A* est considérée comme strictement spécifique de la lignée lymphoïde B (Adams et al., 1992 ; Busslinger et al., 1996 ; Fuxa and Busslinger 2007).

Mis à part ses deux promoteurs alternatifs, *PAX5* comporte également d'autres séquences cis-régulatrices (*cis-regulatory modules, CRM*) : **des *enhancers*** (amplificateurs).

- D'une part, *Pax5* détient un *enhancer* de 435 pb qui est localisé à 5,6 kb en amont de l'exon 1A et qui pourrait suffire à activer la transcription de *Pax5(B)* dans le SNC en cours de développement (Pfeffer et al., 2000).
- D'autre part, un autre *enhancer*, se situe quant à lui à distance des promoteurs, dans l'intron 5-6 de *PAX5*, et est spécifique de l'activation de *PAX5* dans la lignée lymphoïde B (Decker et al., 2009). La coopération entre cet *enhancer* intronique et le promoteur distal de *PAX5* suffirait à activer correctement la transcription de *PAX5* dans la lignée lymphoïde B (Decker et al., 2009).



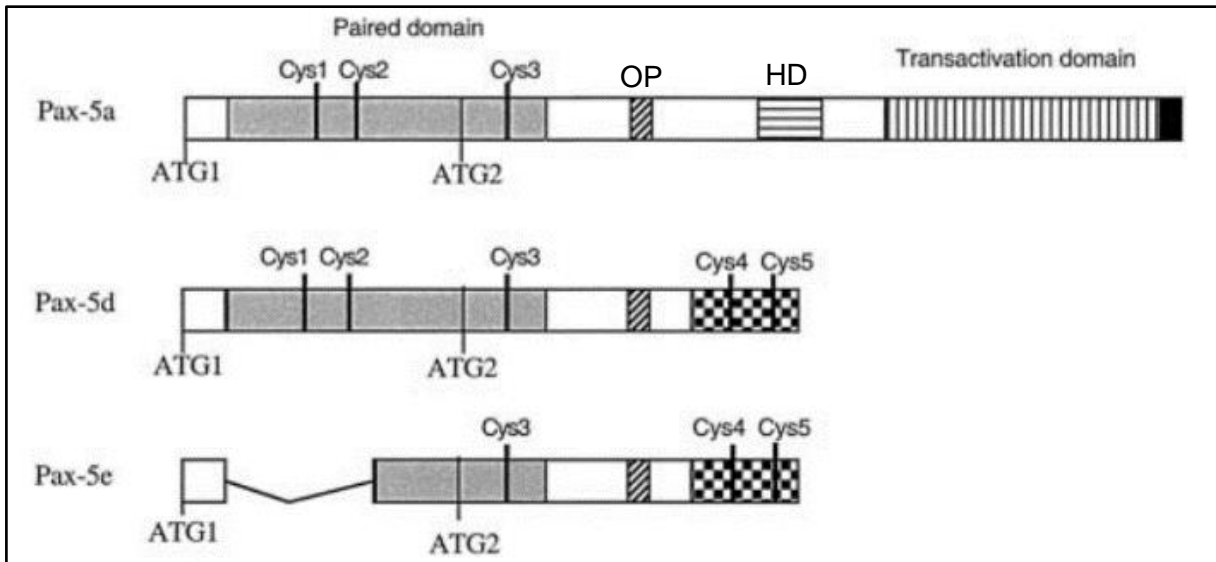


Figure 1.17 : Structure des isoformes Pax5a, Pax5d et Pax5e issues de l'épissage alternatif de *Pax5*.

L'isoforme *full-length* Pax5a contient 10 exons. L'isoforme Pax5d est lésé des exons 6 à 10. L'isoforme Pax5e ne possède ni l'exon 2 ni les exons 6 à 10. Cependant, Pax5d et Pax5e possèdent une séquence C-terminale unique (représentée en damier) de fonction inconnue. ATG1, ATG2 : codons initiateurs alternatifs. Figure issue de Lowen et al., 2001.

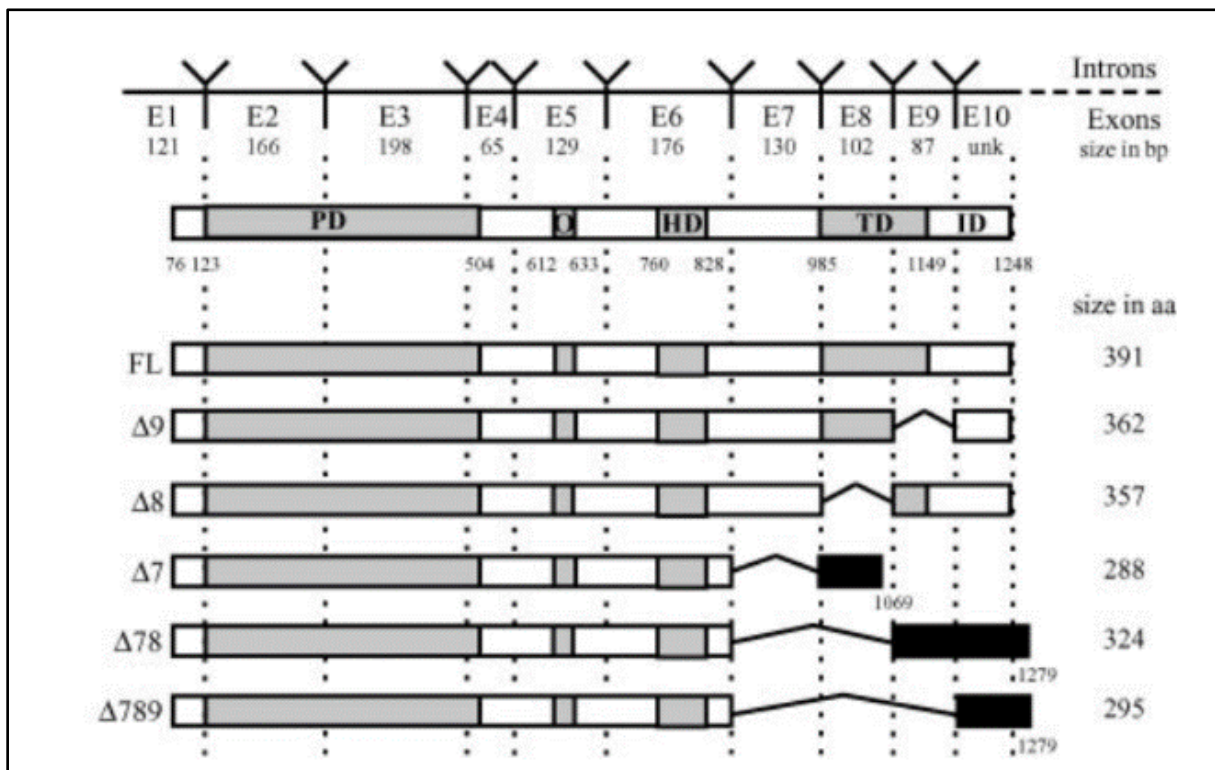


Figure 1.18 : Isoformes de *PAX5* chez l'homme.

FL, Full-length. Les domaines colorés en noir correspondent à de nouvelles séquences protéiques par rapport à PAX5A FL. Figure issue de Robichaud et al., 2004.

### 1.4.3. Variants issus de l'épissage alternatif

Treize à seize variants de transcrits résultant de l'épissage alternatif de *PAX5* sont décrits chez l'homme et la souris dans la littérature mais ne sont pas tous traduits en protéines (Oppezco et al., 2005; Robichaud et al., 2004). L'épissage de ces transcrits peut conduire à la troncature du domaine PAIRED ou à l'exclusion d'un ou plusieurs des exons 6 à 10 (Figures 1.17 et 1.18). Ils présentent en conséquence des modifications de leurs propriétés de liaison à l'ADN et de transactivation, et sont régulés par des voies de signalisation spécifiques (Lowen et al., 2001; Robichaud et al., 2004).

Les variants tronqués issus d'épissage alternatif de *PAX5(A)* pourrait jouer un rôle important dans le contrôle de sa propre transcription et ce en fonction des quantités de formes épissées, ainsi qu'un rôle d'ajustement de l'activité transcriptionnelle de Pax5a (Lowen et al., 2001; Oppezco et al., 2005). Notamment chez la souris, l'isoforme Pax5d, ayant conservé le domaine de liaison à l'ADN intact, est capable d'interagir avec les sites de reconnaissance de Pax5 et donc d'entrer en compétition avec Pax5a (*in vitro*), mais sans pouvoir recruter le complexe basal d'initiation de la transcription, remplissant une fonction de répresseur de Pax5a. Inversement, Pax5e peut amplifier l'activité de transactivation de Pax5a. Le ratio entre Pax5d et Pax5e est régulé au niveau protéique et est corrélé aux statuts d'activation et de prolifération des cellules B (Lowen et al., 2001). Quant aux transcrits *full-length Pax5A* et *Pax5B*, tous les deux exprimés dans le tissu lymphoïde, ils pourraient réguler à la fois leur propre transcription et interférer avec celle de l'autre isoforme de manière dépendante du niveau d'expression de chaque transcrit au cours de la différenciation (Kikuchi, Nakayama, Kuribayashi, Mimuro, Imajoh-Ohmi, Nishitoh, Takami, et al., 2015).

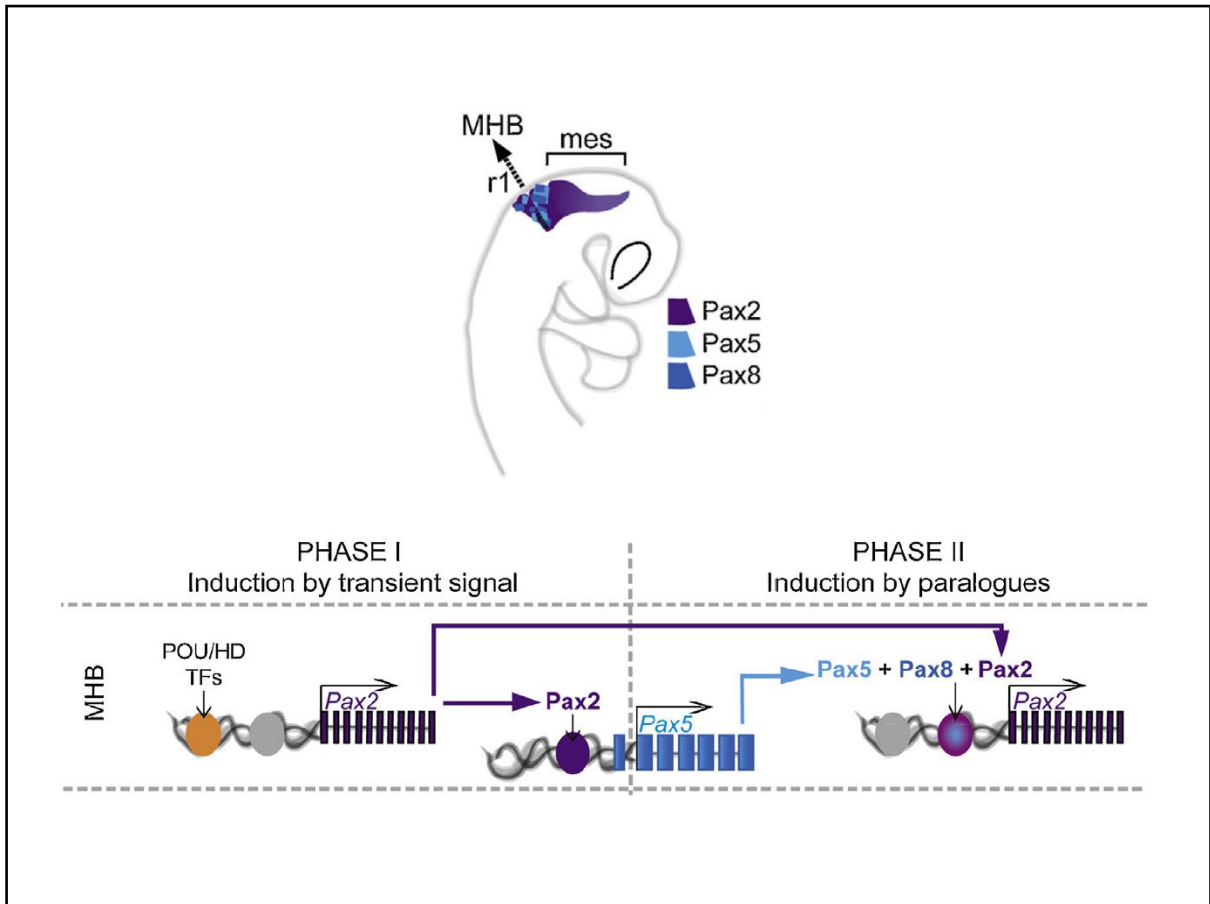


Figure 1.19 : Architecture génomique orchestrant l'induction et le maintien de l'expression de Pax2, Pax5 et Pax8 au sein de la limite (MHB) entre le cerveau moyen (mésencéphale, mes) et le cerveau postérieur (rhombencéphale, r1) Figure issue de la revue de Curto et al., 2015.

#### 1.4.4. Éléments de trans-régulation transcriptionnelle de *PAX5*

Les deux régions promotrices de *PAX5* sont associées à l'acétylation des histones H3, correspondant à une marque d'activation génique (Cresson et al., 2018), suggérant le recrutement d'acétyltransférases d'histones (HAT) au locus de *PAX5* (Kikuchi, Nakayama, Kuribayashi, Mimuro, Imajoh-Ohmi, Nishitoh, & Nakayama, 2015). En comparaison avec les promoteurs de *PAX5*, l'*enhancer* intronique de *PAX5A* présente une faible acétylation des histone H3 (Cresson et al., 2018).

Tout au long de la différenciation (précoce et tardive), les facteurs de transcription PU.1, IRF4, IRF8 et NF- $\kappa$ B contribuent à la régulation de l'*enhancer* intronique spécifique à la lignée lymphoïde B. Celui-ci est d'ores déjà activé dans les progéniteurs multipotents hématopoïétiques et les progéniteurs lymphoïdes : les histones H3 y sont déméthylées et acétylées, indiquant une chromatine active. Cependant, le promoteur ne s'active qu'ultérieurement, à un stade précoce de la différenciation B (stade pro-B) grâce au remodelage de la chromatine de la région promotrice entière de *Pax5* (les histones H3 acétylées et déméthylées). Cette modification épigénétique requiert l'expression du facteur de transcription EBF1 qui se fixe au promoteur proximal de *Pax5* (Decker et al., 2009). Au cours des stades précoces de la différenciation lymphoïde B, la transcription de *Pax5* peut être renforcée par STAT5 puisque le promoteur de *Pax5* possède un motif de fixation STAT (Hirokawa et al., 2003).

Plusieurs études suggèrent que *Pax5* est capable de finement réguler sa transcription par des boucles d'auto-régulation (notamment par le biais de ses isoformes) et d'inter-régulation comme dans le cas des gènes *Pax2*, *Pax5* et *Pax8* dans le SNC (Figure 1.19, Curto et al., 2015). Lors du développement embryonnaire du SNC, l'expression précoce de *Pax2* dans le centre organisateur de la frontière entre le mésencéphale et le diencéphale (MHB) permet de déclencher la transcription de *Pax5* en régulant l'*enhancer* situé en amont de l'exon 1A de *Pax5* (l'*enhancer* spécifique du MHB, Pfeffer et al., 2000). Pour rappel, *Pax5* s'exprime transitoirement dans le SNC au stade embryonnaire (Figure 1.05). Par la suite, les répressions transcriptionnelles croisées de *Pax2*, *Pax5* et *Pax8* induisent l'établissement d'une frontière nette entre le diencéphale et le mésencéphale (Curto et al., 2015).

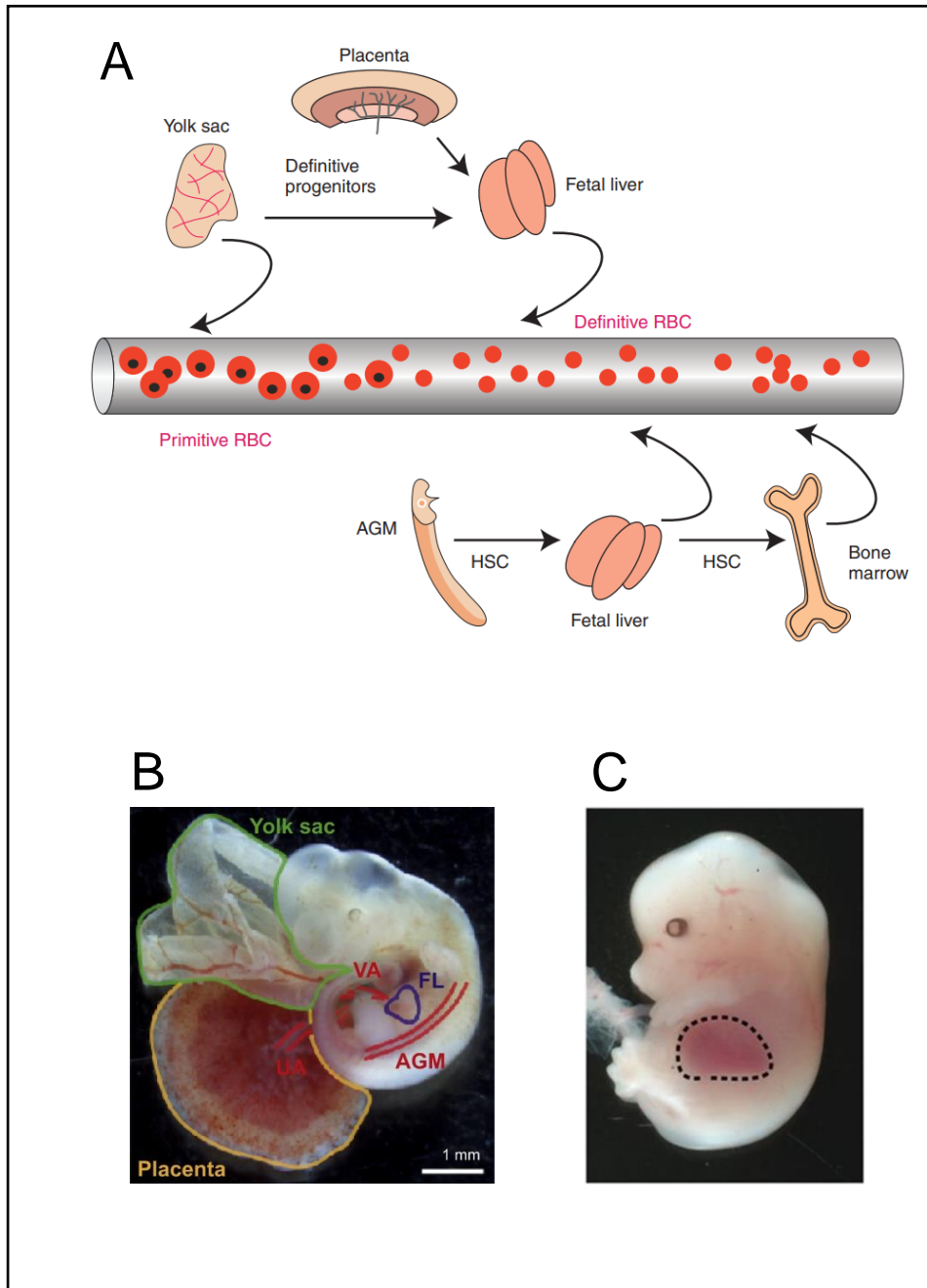


Figure 2.1 : Les sites principaux de production hématopoïétiques dans l'embryon de souris.

A. Ontogénie de la lignée érythroïde. *RBC* : *red blood cells* pour globules rouges. Figure issue de Dzierzak & Philipsen, 2013. B. L'embryon à E11 est connecté au sac vitellin (*yolk sac*) par des vaisseaux vitellins et au placenta par des vaisseaux ombilicaux. Les premières CSH sont présentes au niveau de l'AGM (*aorta-gonad-mesonephros*), de l'artère vitelline (VA), de l'artère ombilicale (UA) et de la tête. Plus tard, elles sont également présentes dans le sac vitellin, le placenta et le foie fœtal (FL pour *fetal liver*), Image issue de Klaus & Robin, 2017. C. Délimitation du foie fœtal à E14.5, principal siège de l'hématopoïèse embryonnaire. Image issue de Dartsch et al., 2014.

# Chapitre 2 de l'introduction

## 2.1. Généralités sur l'hématopoïèse

### 2.1.1. Hématopoïèse fœtale

**Chez la souris** (Figure 2.1), l'hématopoïèse primitive s'initie dans le sac vitellin à partir du septième jour de gestation (E7.5) et génère des cellules hématopoïétiques à potentiel restreint, notamment des érythrocytes embryonnaires primitifs qui sont retrouvés dans la circulation jusqu'à E11/E12 (Dzierzak & Philipsen, 2013). À E9, le sac vitellin et le placenta produisent des progéniteurs définitifs qui migrent dans le foie fœtal où ils se différencient en érythrocytes définitifs qui entrent alors dans la circulation fœtale (Figure 2.1A). Les premières cellules souches hématopoïétiques (CSH) définitives n'émergent qu'au dixième jour de développement embryonnaire (E10.5) dans le plancher de l'aorte dorsale de la région appelée Aorte-Gonades-Mésonephros (AGM, Figure 2.1A et B). Elles proviennent de précurseurs endothéliaux hémogéniques qui effectuent une transition d'une cellule endothéliale vers une cellule hématopoïétique (Boisset & Robin, 2010). Les CSH vont progressivement quitter l'AGM pour notamment rejoindre le foie fœtal qui devient la niche principale de l'hématopoïèse embryonnaire (Figure 2.1C ; Orkin & Zon, 2008; Swain et al., 2014). Les CSH s'y multiplient massivement et le foie fœtal atteint son pic d'activité au stade E13.5. Le stade E14.5 marque un changement sur le plan morphologique du foie fœtal qui va s'orienter vers sa fonction métabolique au détriment de l'hématopoïèse (Crawford et al., 2010). À partir d'E16.5, la proportion de cellules hématopoïétiques hépatiques va décliner alors que les CSH vont coloniser la moelle osseuse. L'hématopoïèse définitive va s'établir principalement dans la moelle osseuse (MO), notamment suppléée par la rate en tant que site secondaire.

### 2.1.2. Modèle actuel de l'hématopoïèse adulte

L'hématopoïèse comprend l'ensemble des processus par lesquels les cellules souches hématopoïétiques (CSH) génèrent continuellement les différents types de lignages de cellules sanguines et cellules immunitaires dans des proportions bien définies à l'équilibre. Pour ce faire, les CSH bénéficient d'une capacité d'auto-renouvellement et font preuve de multipotence ; ainsi par définition, chaque CSH a la capacité de reconstituer tous les lignages hématopoïétiques dans un organisme. L'hématopoïèse est fonctionnellement organisée en hiérarchie, avec des CSH siégeant au sommet de la pyramide.

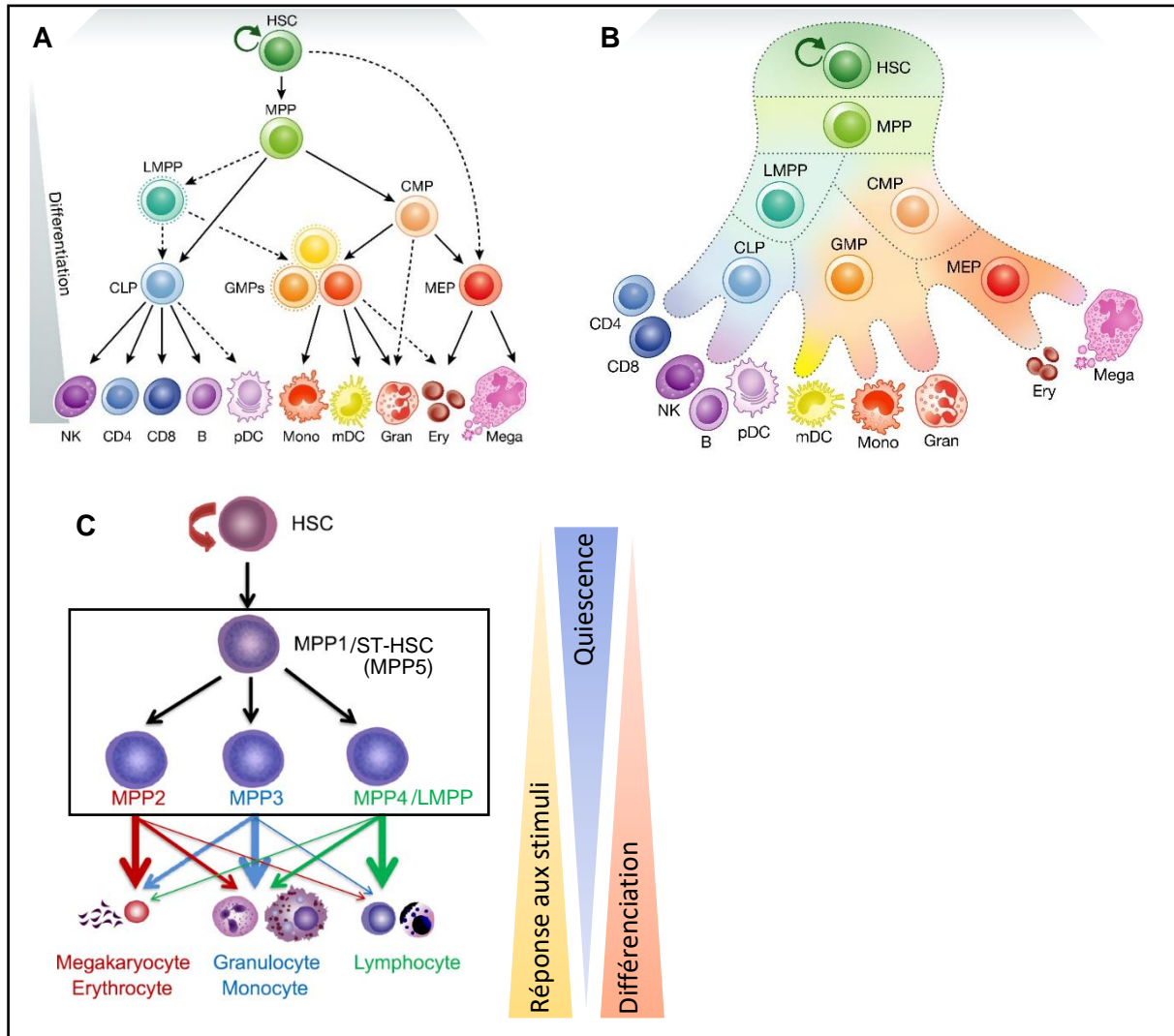


Figure 2.2 : Le modèle revisité de la hiérarchie clonale hématopoïétique (2019)

A. Schéma de la hiérarchie hématopoïétique chez l'Homme. Les lignes pointillées indiquent les chemins de différenciation découverts récemment. Figure issue de Corces et al., 2016, adaptée par Bao et al., 2019. B. Visualisation de la hiérarchie hématopoïétique dans laquelle l'engagement s'effectue en continuum plutôt que via des stades distincts, une perspective suggérée par de récentes études transcriptomiques sur cellule unique. Figure issue de Grootens et al., 2019 adaptée par Bao et al., 2019. C. Les cellules souches hématopoïétiques se différencient d'abord en MPP1/ST-HSC(MPP5) puis génèrent des sous-ensembles de progéniteurs multipotents (MPP2-4). Les MPP1 peuvent générer tous les lignages. Les MPP2/3 comportent un biais myéloïde et les MPP4 un biais lymphoïde. De plus, les MPP2 comportent un biais plaquettaire. Figure C adaptée à partir de Cheng et al., 2020.

Abréviations par ordre alphabétique : **B**, lymphocyte B ; **CD4**, lymphocyte T CD4+ ; **CD8**, lymphocyte T CD8+ ; **CLP**, progéniteur lymphoïde commun ; **CMP**, progéniteur myéloïde commun ; **ery**, cellule érythroïde ; **mono**, monocyte ; **gran**, granulocyte ; **GMP**, progéniteur granulocyte-macrophage ; **HSC**, cellule souche hématopoïétique ; **mDC**, cellule dendritique myéloïde ; **mega**, mégacaryocyte ; **NK**, cellule *natural killer* ; **MEP**, progéniteur mégacaryocyte-érythroïde ; **MPP**, progéniteur multipotent ; **LMPP**, progéniteur multipotent lymphoïde-primé ; **pDC**, cellule dendritique plasmacytoïde.



Le modèle traditionnel (*tree-like and balanced hierarchy model*, Figure 2.2A, flèches continues) fait état d'une cellule souche qui évolue vers un type cellulaire mature précis, en restreignant graduellement ses alternatives par le biais d'une succession de progéniteurs intermédiaires de plus en plus différenciés. Les CSH se différencient en progéniteurs multipotents (MPP) puis en progéniteurs communs myéloïdes (CMP) ou progéniteurs multipotents avec un biais lymphoïde (LMPP pour *lymphoid-primed* MPP). Les CMP génèrent majoritairement les lignées mono-granulocytaire, érythrocytaire et des mégacaryocytaire (Figure 2.2A) et les LMPP se différenciant en progéniteurs lymphoïdes communs (CLP) génèrent les cellules NK et les lymphocytes B et T. Néanmoins, les avancées technologiques des deux dernières décennies ont permis de démontrer la complexité de l'hématopoïèse et de proposer plusieurs modèles révisés de la hiérarchie hématopoïétique (Figure 2.2A-D) qui s'avère beaucoup plus dynamique et adaptative (Cheng et al., 2020 pour revue). Ces modèles illustrent l'hétérogénéité phénotypique et fonctionnelle des CSH et des progéniteurs (notamment MPP, CMP, MEP et CLP) au sein de leurs propres compartiments. Ils mettent en exergue le continuum d'options qui s'offrent aux CSH plutôt qu'un processus multi-étapes prédéterminé en amont et bimodal en aval. Ainsi, les CSH peuvent soit réduire progressivement les alternatives de différenciation, soit faire un choix de lignage immédiat à partir de ce continuum qui englobe initialement toutes les options de cellules matures. Ces modèles révisés encouragent à considérer la différenciation hématopoïétique comme un processus continu sans transition abrupte et impliquent que l'environnement joue un rôle important dans le choix de différenciation des CSH et la diversification des cellules matures. Un niveau de complexité supplémentaire a émergé avec la caractérisation plus précise du microenvironnement (Méndez-Ferrer et al., 2020 pour revue) : en effet, leurs localisations particulières dans la niche hématopoïétique et leurs interactions avec les composants du microenvironnement modulent également la destinée des cellules souches et progéniteurs hématopoïétiques.

L'hématopoïèse est un processus dynamique à d'autres égards : quel que soit son paradigme, elle n'en reste pas moins finement régulée par des facteurs extrinsèques amenés par le stroma (cytokines, facteurs de croissance, molécules d'adhésion, antigène) qui modifient l'activité et l'expression de facteurs intrinsèques (voies de signalisation cellulaire et facteurs de transcription) dont la balance de dosage importe beaucoup en termes de décision de destinée cellulaire. En particulier, les cytokines ont une fonction déterminante dans la régulation de l'expansion et la spécification des progéniteurs multipotents ; elles influencent certains facteurs de transcription exprimés de manière latente dans les progéniteurs multipotents.



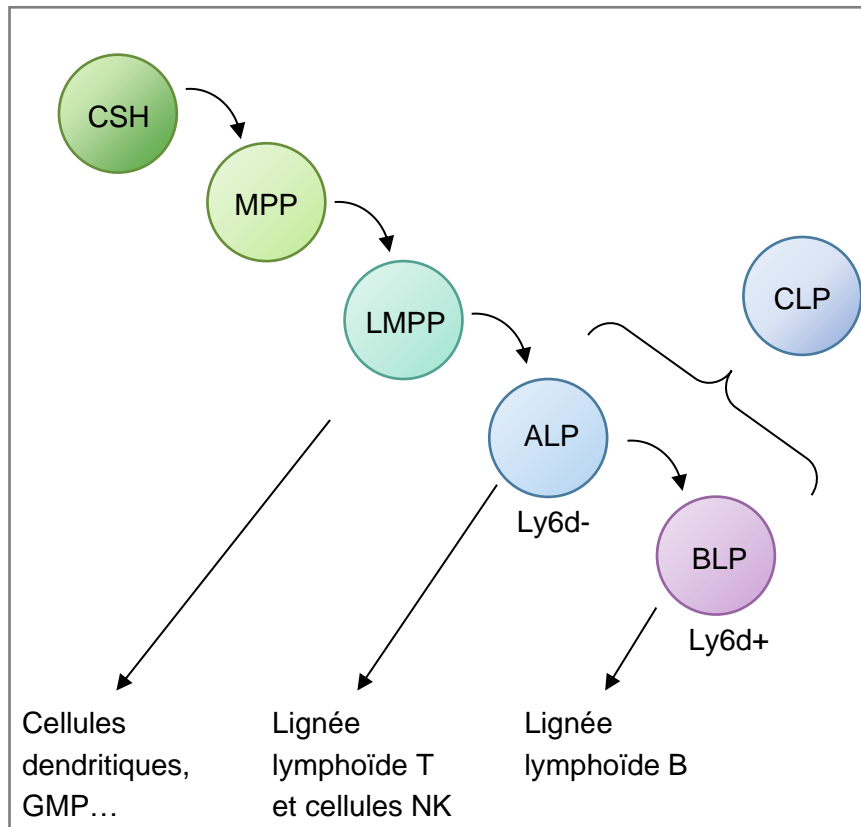


Figure 2.3 : Modèle de spécification de la lignée lymphoïde B.

Abréviations par ordre alphabétique : **ALP**, progéniteurs lymphoïdes totaux (*All Lymphoid Progenitors*) ; **BLP**, progéniteurs lymphoïdes B (*B cell-biased lymphoid progenitor*) ; **CLP**, progéniteurs lymphoïdes communs ; **CSH**, cellules souches hématopoïétiques ; **GMP**, progéniteurs granulocytes–macrophages ; **NK**, *natural killer* ; **MPP**, progéniteurs multipotents; **LMPP**, progéniteurs multipotents lymphoïde-*primed*. Figure créée d'après Inlay et al., 2009.

## 2.2. La lymphopoïèse B

La lymphopoïèse B récapitule spécifiquement les processus impliqués dans la production des lymphocytes B, responsables de l'immunité à médiation humorale via la production d'anticorps. L'hétérogénéité des progéniteurs et précurseurs hématopoïétiques tout comme la continuité dans le processus de différenciation s'appliquent également au compartiment lymphoïde. La spécification, l'engagement et la différenciation de la lignée lymphoïde B dépendent essentiellement de signaux extracellulaires et des facteurs de transcription.

### 2.2.1. Mécanismes de la lymphopoïèse B précoce dans la MO

#### 2.2.1.1. La classification des étapes précoces du développement B

En ce qui concerne le développement du lignage lymphoïde B (Figure 2.3), les CSH engendrent en premier lieu des progéniteurs multipotents possédant un biais lymphoïde tout en préservant un potentiel myéloïde (*MPP4, multipotent progenitor subset 4*, Pietras et al., 2015 ; *LMPP, lymphoid-primed multipotent progenitors*, Adolfsson et al., 2005). Résultant de la différenciation des LMPP, les progéniteurs lymphoïdes communs CLP (*common lymphoid progenitors*, Kondo et al., 1997) traditionnels (exprimant CD10) conservent une capacité limitée à générer des cellules myéloïdes (Rumfelt et al., 2006). Ils s'avèrent fonctionnellement hétérogènes et peuvent notamment être redéfinis en deux populations distinctes sur la base de l'expression du Ly6D (Inlay et al., 2009; Mansson et al., 2010): les ALP (*All Lymphoid Progenitors*) Ly6d-négatives, et les BLP (*B cell-biased lymphoid progenitor*, alias ELP, *Earliest B-cell specified lymphoid progenitor*) Ly6d-positives. Situées à l'embranchement entre développement B et T, les ALP génèrent principalement les cellules de la lignée lymphoïde T et peuvent produire toutes les cellules NK – des CLP exprimant l'intégrine  $\alpha 2\beta 7$  (LPAM1) ou CXCR6 (*C-X-C Motif Chemokine Receptor 6*, (Possot et al., 2011) – alors que les BLP génèrent presque exclusivement des cellules de la lignée lymphoïde B. La fraction de CLP exprimant la téthérine PDCA-1 (alias BST2, *bone marrow stromal antigen 2*) correspond aux précurseurs de cellules dendritiques (Medina et al., 2013). Les BLP entrent dans le lignage lymphoïde B en se différenciant en cellules pré-pro-B qui sont destinées à se restreindre au lignage B et expriment des niveaux d'expression variables du marqueur B220 (Hardy et al., 1991). Le stage le plus précoce de la spécification B correspondraient soit aux BLP soit aux cellules pré-pro-B, sans pour autant être irrévocablement engagées. Les cellules pré-pro-B sont abondamment présentes dans le foie fœtal et prolifèrent facilement *ex vivo* sur un stroma en présence d'interleukine-7 (IL-7) mais elles se font rares dans la moelle osseuse adulte (0.05-0.1% de la moelle osseuse totale, Rumfelt et al., 2006).

STADES	Précoce → Tardif						
Localisation	Moelle osseuse						Rate
Fractions de Hardy	Fraction A	Fraction B	Fraction C	Fraction C'	Fraction D	Fraction E	Fraction F
Nomenclature de Philadelphia (stade de différenciation)	Pre-pro-B	Early Pro-B	Late pro-B	Early (large) Pre-B	Late (small) Pre-B	Newly Formed B	Follicular B
Nomenclature de Basel (Rolink et Melchers)	Pro-B	Pre-B-I	Pre-B-I	Large Pre-B-II (cycling)	Small Pre-B-II (resting)	Immature B	Mature B
Développement des pré-BCR et BCR				Pré-BCR fonctionnel		BCR fonctionnel	BCR
Ig de surface				Ig $\mu$ et SLC	Ig $\mu$ intracellulaire	Ig $\mu$ (IgH), Ig $\kappa$ et Ig $\lambda$ (IgL)	Ig $\mu$ (IgH), Ig $\kappa$ et Ig $\lambda$ (IgL)
Statut de réarrangement du locus <i>IgH</i>	Germinal	Recombinaison $D_H \rightarrow J_H$ en cours	Recombinaison $V_H \rightarrow DJ_H$ en cours	VDJ (Recombinaison $V_H-DJ_H$ se termine)	VDJ	VDJ	VDJ
Statut de réarrangement du locus <i>IgL</i>	Germinal	Germinal	Germinal	Germinal	Recombinaison $V_L \rightarrow J_L$ en cours	VJ	VJ

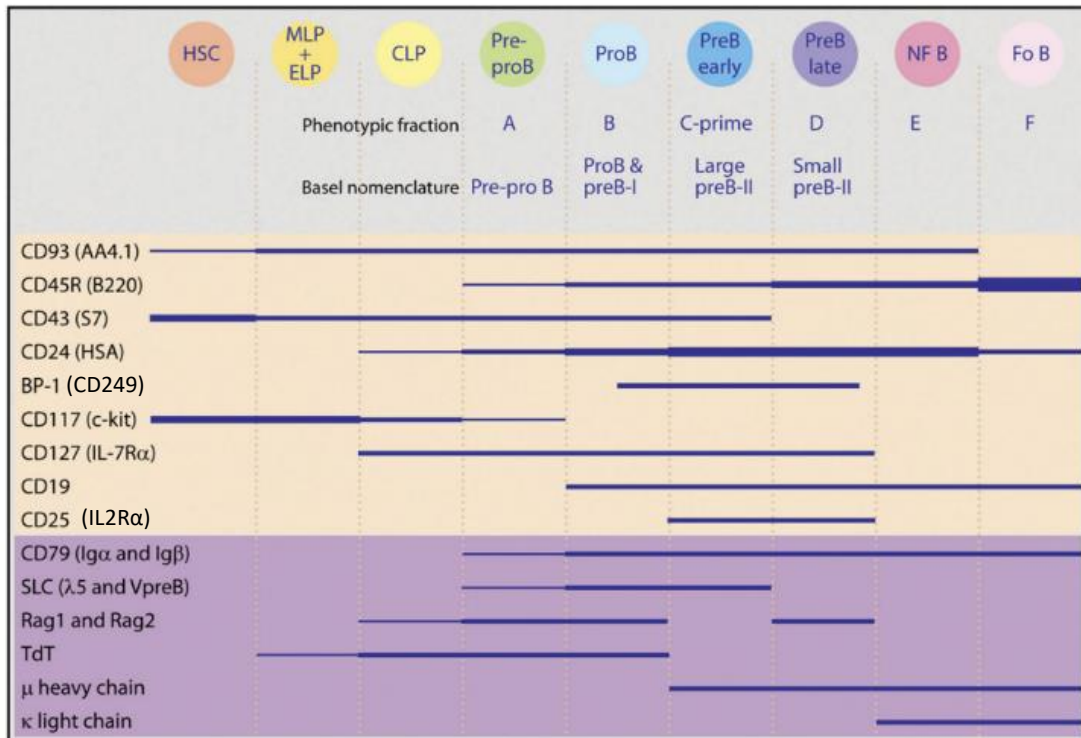


Figure 2.4 : Classification phénotypique et fonctionnelle démarquant les stades précoces du développement de la lignée B dans la moelle osseuse de souris.

L'expression des marqueurs de surface indiqués dans le cadre beige a été évaluée par analyses par cytométrie en flux. L'expression des gènes indiqués dans le cadre violet a été évaluée par analyses d'ARNm, d'expression de gène-rapporteur ou de marquage cytoplasmique. La largeur des lignes indique le niveau d'expression des marqueurs correspondants.

**Abréviations :** *HSC*, hematopoietic stem cells; *MLP*, multilineage progenitor; *ELP*, early lymphoid progenitor; *NF B*, newly formed B cell; *Fo B*, follicular B cell; *Tdt*, Terminal deoxynucleotidyl transférase (Figure issue de Richard R. Hardy et al., 2007). De plus, le champ d'expression de CD2 est restreint aux 'petites' cellules pré-B contenant des chaînes *IgH* cytoplasmiques (Milne et al., 2004).

Bien que ces progéniteurs ne soient jusqu'ici pas strictement engagés dans une voie de différenciation spécifique (« *priming* »), ils expriment 'passivement' un ensemble de gènes de différenciation B (tels que *Ebf1*, *Pax5*), prêts à répondre à un signal inductif type cytokinique (Månsson et al., 2007). Au stade suivant, les pro-B précoces, marquées par l'expression de CD19 qui est transactivé spécifiquement par Pax5 (Kozmik et al., 1992), s'engagent stablement, activement et irréversiblement dans la lignée B. Par ailleurs, une fraction de cellules exprimant B220 retient un potentiel de lignage T et est capable de migrer dans le thymus (Rumfelt et al., 2006; Martin et al., 2003).

Dans la moelle osseuse adulte, le développement de la lignée lymphoïde B peut se diviser en stades cellulaires (tableau 2.1) :

- la nomenclature de Bâle (*BaseI*), distinguant les cellules en fonction de leur taille pour identifier les sous-compartiments de progéniteurs lymphoïdes B proliférant.
- la classification de Hardy en fractions phénotypiques se basant sur la cytométrie en flux, et distinguant ainsi les cellules en fonction de l'expression de marqueurs de surfaces.

Chaque phénotype correspond fonctionnellement à une étape du mécanisme de réarrangement des loci *IgH* et *IgL* (Figure 2.4, Richard R. Hardy et al., 2007).

Ces classifications phénotypiques ont été initialement développées en parallèle par les groupes de RR Hardy, (Hardy et al., 1991; Hardy & Hayakawa, 2001; Rumfelt et al., 2006), A. Rolink et F. Melchers (Melchers, 2005; Melchers et al., 1991, 1994) et la caractérisation phénotypique des stades de différenciation lymphoïde B (et sa contrepartie fonctionnelle) continue d'être étoffée et affinée (Figure 2.5, Jensen et al., 2016).

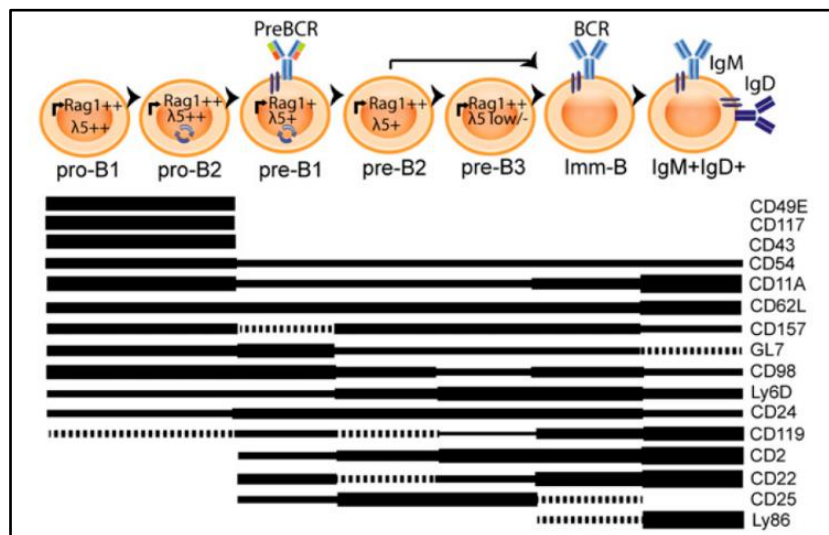


Figure 2.5 : Champs d'expression de marqueurs de surface permettant la distinction des étapes de la différenciation B dans la moelle osseuse de souris.

La largeur des lignes indique les niveaux d'expression relatifs des marqueurs de surface étudiés et les lignes pointillées indiquent que moins de 50% des cellules expriment le marqueur concerné. Figure issue de Jensen et al., 2016.

TABLEAU 2.1: Termes des stades cellulaires B dans la moelle osseuse et statuts de réarrangements des Ig

STADES	Précoce $\longrightarrow$ Tardif						
Localisation	Moelle osseuse						
Fractions de Hardy	Fraction A	Fraction B	Fraction C	Fraction C'	Fraction D	Fraction E	Fraction F
Nomenclature de Philadelphia (stade de différenciation)	Pre-pro-B	Early Pro-B	Late pro-B	Early (large) Pre-B	Late (small) Pre-B	Newly Formed B	Follicular B
Nomenclature de Basel (Rolink et Melchers)	Pro-B	Pre-B-I	Pre-B-I	Large Pre-B-II (cycling)	Small Pre-B-II (resting)	Immature B	Mature B
Développement des pré-BCR et BCR				Pré-BCR fonctionnel		BCR fonctionnel	BCR
Ig de surface				Ig $\mu$ et SLC	Ig $\mu$ intracellulaire	Ig $\mu$ (IgH), Igk et Ig $\lambda$ (IgL)	Ig $\mu$ (IgH), Igk et Ig $\lambda$ (IgL)
Statut de réarrangement du locus IgH	Germinal	Recombinaison D <sub>H</sub> →J <sub>H</sub> en cours	Recombinaison V <sub>H</sub> →DJ <sub>H</sub> en cours	VDJ (Recombinaison V <sub>H</sub> -DJ <sub>H</sub> se termine)	VDJ	VDJ	VDJ
Statut de réarrangement du locus IgL	Germinal	Germinal	Germinal	Germinal	Recombinaison V <sub>L</sub> →J <sub>L</sub> en cours	VJ	VJ

## 2.2.1.2. Mécanismes moléculaires de la lymphopoïèse B précoce

### 2.2.1.2.1. La recombinaison V(D)J des loci des gènes d'immunoglobulines

Le développement des cellules de la lignée lymphoïde B dans la moelle osseuse dépend de réarrangements somatiques aléatoires des loci du gène des chaînes lourdes (H pour *heavy*) puis du gène des chaînes légères (L pour *light*) des immunoglobulines (Ig). Ce mécanisme appelé recombinaison V(D)J se caractérise par des étapes d'assemblage unique des segments géniques V (variable), D (diversité) et J (jonction) des deux loci *d'IgH* et des segments de gène V et J des multiples loci *d'IgL*, ce qui va coder et constituer la région variable du récepteur antigénique. Survenant avant l'interaction avec l'antigène spécifique, la recombinaison V(D)J est ainsi à l'origine de la diversité des sites de liaison à l'antigène, appelée diversité combinatoire du système immunitaire adaptatif. Ce vaste répertoire de récepteurs antigéniques (*B cell receptor (BCR)* ou Ig solubles ultérieurement) créé permet de reconnaître un très large spectre d'antigènes.

La recombinaison V(D)J des chaînes d'Ig est un processus séquentiel qui définit les étapes de formation du BCR et de son précurseur, le pré-BCR (Tableau 2.1). En contrepartie, le pré-BCR et le BCR établissent une signalisation permettant la sélection positive puis négative des cellules B et garantissant l'homéostasie, la différenciation et la fonction des pré-lymphocytes B et lymphocytes B.

La recombinaison V(D)J fait précisément intervenir de courtes séquences cibles flanquant chacun des segments V, D, J appelées séquences « signal de recombinaison » (RSS pour *recombination signal sequences*). Ce sont les enzymes codées par les gènes RAG (*recombination activating genes*) qui assurent le brassage combinatoire (Oettinger et al., 1990). Le complexe enzymatique formé par RAG1 et RAG2 initie les recombinaisons V(D)J en se fixant aux RSS. Les deux motifs RSS doivent au préalable être rapprochés à proximité et assemblés en un complexe synaptique s'assurant que les partenaires de recombinaisons soient juxtaposés (Gellert, 2002) avant que le complexe RAG1/RAG2 n'introduise des cassures d'ADN double-brins entre chaque élément RSS et le segment génique adjacent (Nishana & Raghavan, 2012). Après le clivage d'ADN, les enzymes RAG en coopération avec des facteurs de réparation d'ADN du système NHEJ (*non-homologous end joining*) joignent les extrémités de gènes pour achever la réaction de recombinaison. L'ajout de nucléotides aléatoires par la déoxynucleotidyle transférase terminale (TdT codée par *dntt*) et les erreurs de réparation subséquentes (ajout ou retrait de base·s nucléotidique·s) sont à l'origine de la diversité jonctionnelle (Lieber et al., 2006).

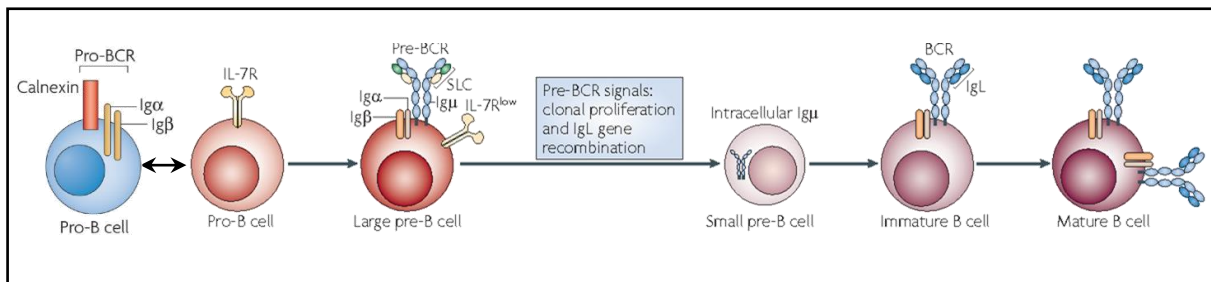


Figure 2.6 : Illustration des stades de différenciation lymphoïde B dans la moelle osseuse.

SLC : *surrogate light chain*, pseudo-chaîne légère composé de VpréB et lamda-5. Figures issues de Cambier et al., 2007 et Herzog et al., 2009.

#### 2.2.1.2.2. Formation et signalisation du pré-BCR

Les cellules pro-B présentent un complexe considéré comme pro-BCR (*B cell receptor*) qui se compose de l'hétérodimère formé par les glycoprotéines CD79a (ou Ig $\alpha$  pour *immunoglobuline-associated alpha* codée par le gène *mb-1*) et CD79b (ou Ig $\beta$  pour *immunoglobulin-associated beta* codée par le gène *B29*) recrutées à la membrane plasmique par la calnexine (Nagata et al., 1997). Cet hétérodimère est indispensable pour transduire les signaux du pré-BCR et du BCR (Melchers et al., 2000).

En premier lieu, la recombinaison VDJ a lieu sur un allèle du locus *IgH* (activation mono-allélique) pour générer une chaîne lourde VDJ- $\mu$  (Ig $\mu$ ) fonctionnelle : la recombinaison D<sub>H</sub>-J<sub>H</sub> débute au stade des pro-B précoces (fraction B) et la recombinaison V<sub>H</sub>-D<sub>H</sub> s'achève dans les cellules pro-B tardives (fraction C), marquant la transition avec le stade des cellules pré-B (fraction D). Jusqu'à un quart des allèles d'*IgH* procèdent déjà à la recombinaison V<sub>H</sub>-D<sub>H</sub> dans les cellules pro-B précoces (fraction B). Aux stades pro-B/pré-B, les pseudo-chaînes légères (SLC pour *surrogate light chain*) sont ensuite exprimées en tant que partenaires de liaison pour la chaîne lourde Ig $\mu$  réarrangée. Chaque pseudo-chaîne légère est formée d'un hétérodimère composé de deux protéines invariantes (codées par des gènes en configuration germinale) : VpréB1 (Igi pour *Immunoglobulin iota chain* ou CD179a) et lambda-5 ( $\lambda$ 5 ou lambda-like-1 ou CD179b ou VpréB2), (Vettermann et al., 2006). L'assemblage de la pseudo-chaîne légère avec la chaîne lourde  $\mu$  des immunoglobulines forment ainsi le pré-BCR (Melchers et al., 2000). Il doit s'associer avec les sous-unités de signalisation Ig $\alpha$  (CD79A) et Ig $\beta$  (CD79B) pour former un récepteur fonctionnel.

Dans un premier temps, l'assemblage du pré-BCR et la signalisation qui en découle constituent un point de contrôle (*checkpoint*) majeur dans la progression de la différenciation de la lignée lymphoïde B car ils fournissent un retour sur la fonctionnalité de la chaîne lourde Ig $\mu$  tout juste réarrangée. Par le biais d'un signal médié par les protéines adaptatrices CD79A/B, le pré-BCR indique ainsi à la cellule qu'une chaîne lourde fonctionnelle issue d'une recombinaison VDJ productive a été générée ; dans le cas contraire, la recombinaison a lieu sur le second allèle d'*IgH*. L'expression d'un pré-BCR fonctionnel induit transitoirement la répression des enzymes de recombinaison RAG1/2 dans les cellules pré-B pour inhiber des réarrangements supplémentaires au locus *IgH* et notamment empêcher la recombinaison V<sub>H</sub>-D<sub>H</sub> du second allèle d'*IgH* (Grawunder et al., 1995). Ce mécanisme d'exclusion allélique garantit l'expression d'un BCR capable de reconnaître un seul antigène (Geier & Schlissel, 2006; Melchers et al., 1995). En conséquence, les clones pré-B (fraction C') avec des réarrangements VDJ-IgH productifs bénéficient d'une prolifération importante (expansion clonale) induite par la signalisation du pré-BCR avant de continuer leur maturation. Ainsi les cellules pré-B précoces (fractions C') qui expriment un pré-BCR fonctionnel sont conventionnellement affublées de l'adjectif 'large' puisqu'elles sont activement en cycle (cyclantes).



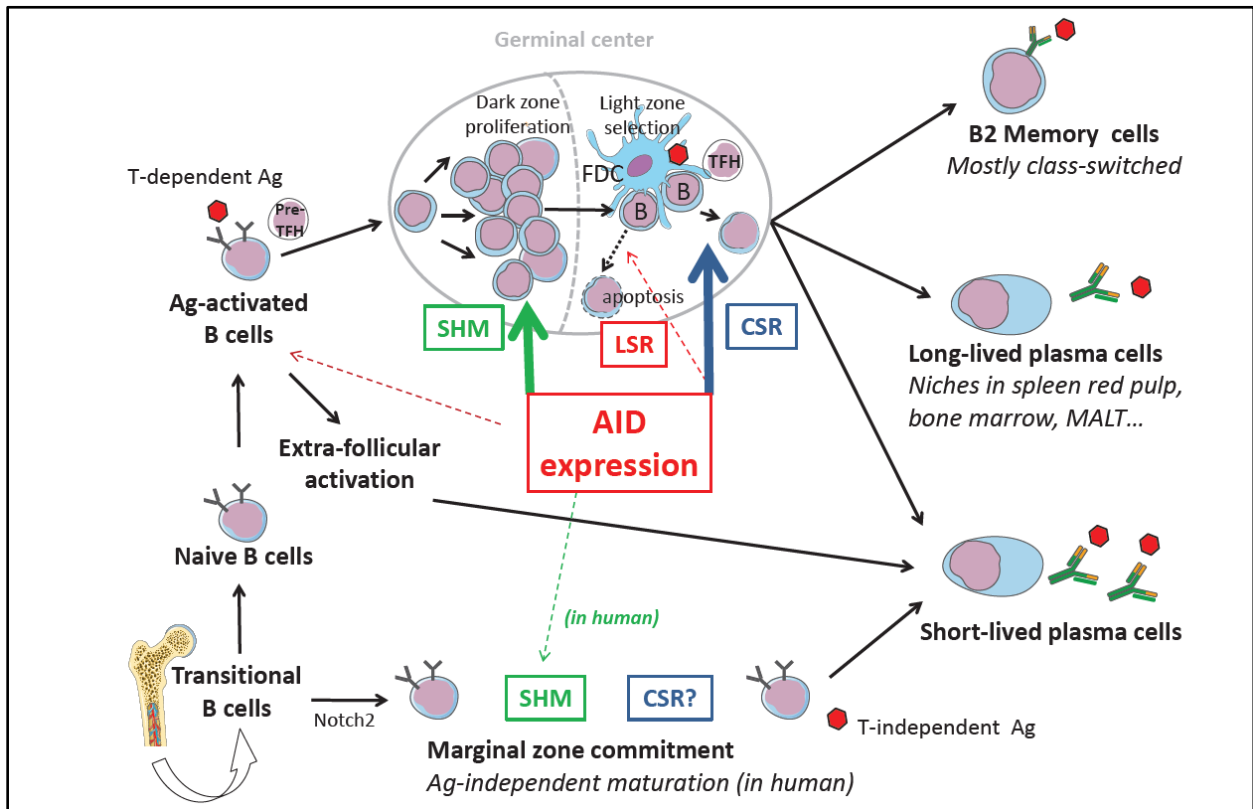


Figure 2.7 : Maturation de la lignée lymphoïde B médiée par l'enzyme AID lors des stades tardifs dans les tissus lymphoïdes.

Après la recombinaison V(D)J et l'expression des IgM de surface, les cellules sortant de la moelle osseuse pour rejoindre la circulation sont nommées cellules transitionnelles. Elles peuvent s'engager vers divers compartiments lymphoïdes en fonction de la signalisation du BCR (l'expression de Notch2 et une signalisation tonique faible du BCR engagent les cellules vers la zone marginale). Après la maturation en cellules IgM+/IgD+, les cellules B naïves qui rencontrent des antigènes (dépendamment des LT) procèdent aux interactions associées avec les pré-lymphocytes pré-TFH (T folliculaires *helper*). Ensuite, les LB s'activent initialement dans les foci extra-folliculaires et se différencient en plasmocytes à courte durée de vie (*short-lived plasma cells*) ou, en coopération avec les LT folliculaires *helper*, participent à la formation des centres germinatifs (GCs) où leurs BCR seront grandement modifiés par les mécanismes de recombinaison de commutation de classe (CSR majoritairement dans la zone claire des GCs) ou d'hypermutation somatique (SHM, majoritairement dans la zone sombre proliférante des GCs), ou éventuellement subiront la recombinaison locus-suicide (LSR). La réaction du centre germinatif génère des LB mémoires et des plasmocytes, soit de courte durée de vie, soit de longue durée de vie survivant alors après leur migration vers des niches protectrices (pulpe rouge de la rate, moelle osseuse, tissus lymphoïdes associés aux muqueuses (MALT)...). Figure issue de (Laffleur et al., 2014).

Dans un second temps, la signalisation pré-BCR induit l'arrêt du cycle cellulaire car les pré-B doivent cesser de proliférer et être 'au repos' (*resting*) comme prérequis avant de poursuivre leur maturation. Elles sont alors conventionnellement nommées 'petites' cellules pré-B (fraction D). Ensuite, la signalisation du pré-BCR enclenche la recombinaison du locus *IgL*. Elle va réguler positivement l'expression des facteurs de transcription promouvant les réarrangements des *IgL* et activer la transcription de gènes des chaînes légères kappa et lambda (Melchers et al., 2000). Par ailleurs la signalisation du pré-BCR active une boucle de rétrocontrôle négatif de ses propres composants (VpréB1 et lambda-5) qui aboutit à la répression de l'expression pré-BCR et de sa signalisation (Parker et al., 2005). La recombinaison productive des gènes de la chaîne légère des *Ig* (*IgL*) médiée par les enzymes RAG mène à la formation du BCR qui conduit à la progression de la différenciation des cellules pré-B tardives en cellules immatures B (fraction F). Le BCR est ainsi constitué des quatre mêmes composants que le pré-BCR à l'exception des pseudo-chaînes légères qui sont remplacées par deux chaînes légères *IgL* conventionnelles qui sont chacune formée d'un homodimère. L'expression constitutive du BCR est essentielle à la survie cellulaire. Elle permet de transduire des signaux « toniques » (signalisation basale en continu) également médiés par l'hétérodimère CD79A/CD79B. La maturation du BCR se poursuit hors de la moelle osseuse.

## 2.2.2. Mécanismes de la lymphopoïèse B tardive hors de la MO

### 2.2.2.1. Stades cellulaires tardifs

La recombinaison V(D)J a généré des cellules lymphoïdes B dites transitionnelles. Pour continuer leur maturation, elles sont relocalisées hors de la moelle osseuse en périphérie vers divers compartiments lymphoïdes secondaires composés de cellules lymphoïdes quiescentes dites naïves. À ce stade, les cellules B naïves expriment des niveaux variables d'IgM ou d'IgD de surface. En périphérie, ces cellules lymphoïdes se différencient en trois types de cellules B matures en fonction du microenvironnement (zones spécifiques spléniques, ganglions lymphatiques, cavité pleurale et cavité péritonéale) et peuvent rejoindre la circulation sanguine (Fairfax et al., 2008; Nutt et al., 2011).

- Les cellules B folliculaires représentent la majorité des lymphocytes B naïfs et sont exclusivement localisées dans les follicules lymphoïdes de la rate et des ganglions lymphatiques ;
- Les cellules B de la zone marginale résident dans le sinus marginal de la rate ;
- Les cellules B1 sont majoritairement localisées aux niveaux des cavités pleurale et péritonéale et à la surface de surfaces muqueuses.

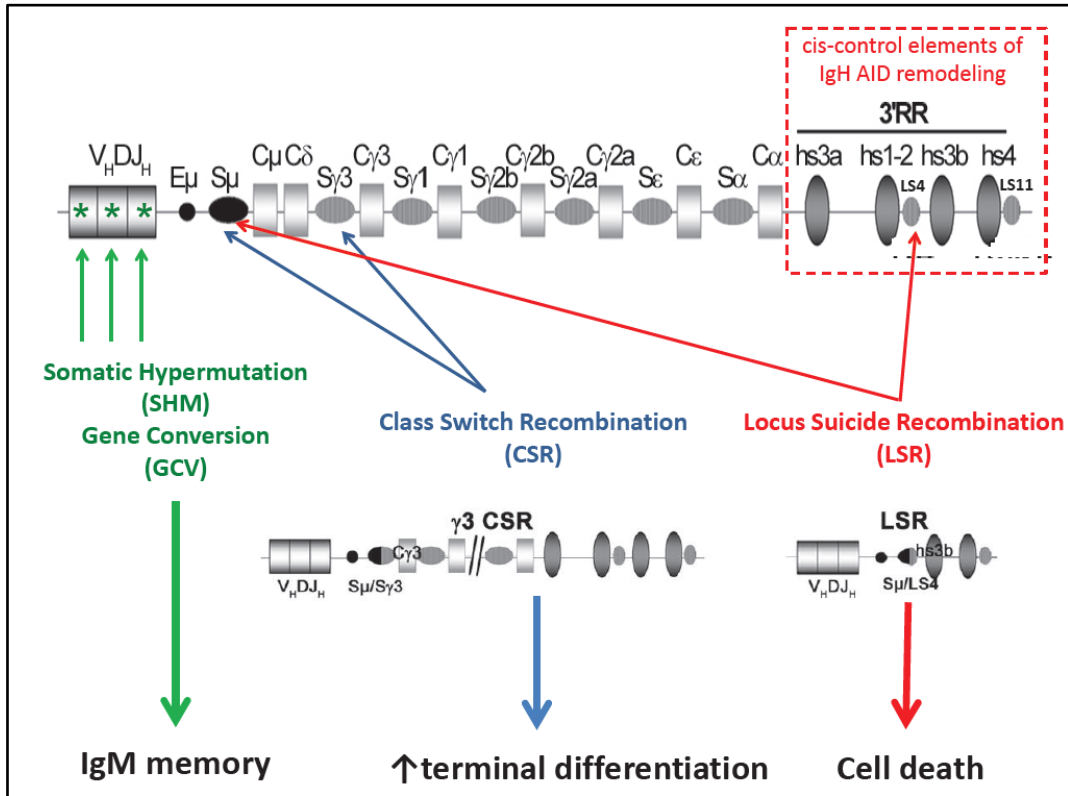


Figure 2.8 : Mécanismes moléculaires médiés par l'enzyme AID au locus *IgH* lors de la lymphopoïèse tardive.

Lors de l'activation des lymphocytes B, l'expression d'AID remodèle les régions variables (V) du gène des Ig par le biais du mécanisme d'hypermutation somatique (SHM) ou éventuellement de conversion génique (GCV), générant des BCR avec une meilleure affinité pour l'antigène considéré. En parallèle ou par la suite, les LB actives diversifient la fonction de leur BCR par le biais du mécanisme de recombinaison de commutation de classe des Ig (CSR ou switch). La recombinaison "suicide" du locus des Ig a éventuellement lieu entre la région *switch*  $\mu$  ( $S_\mu$ ) et des régions répétitives *like-switch* (LSR) localisées dans la région régulatrice 3' (**3'RR**), puis supprime tous les segments géniques constants du locus *IgH* et éteint l'expression du BCR, menant ainsi à la mort cellulaire. Figure issue de Laffleur et al., 2014.

Les cellules B de la zone marginale et B1 répondent rapidement en cas d'infection bactérienne. Quant aux cellules B folliculaires, l'activation de la signalisation de leurs BCR – induite soit par la fixation d'un antigène soit par des stimuli mitogéniques variés – induit leur expansion clonale. Les lymphocytes B (LB) activés forment des centres germinatifs (GCs) en proliférant (zones sombres des GCs). Ensuite, ils se différencient soit en plasmablastes (cyclants, localisés dans des foci extra-folliculaires dans les organes lymphoïdes périphériques) et plasmocytes (non-cyclants, retournant préférentiellement dans la moelle osseuse) soit en LB mémoires (Fairfax et al., 2008; Nutt et al., 2011). Les plasmocytes constituent le stade terminal de la différenciation B et peuvent sécréter leurs BCR sous la forme d'immunoglobulines solubles : les anticorps. La majorité des plasmocytes pérennes (de longue durée de vie) proviennent des cellules B sélectionnées dans les centres germinatifs.

#### 2.2.2.2. Maturation et sélection des gènes des immunoglobulines

Les lymphocytes B (LB) ont développé trois autres mécanismes pour adapter au mieux la réponse immunitaire, notamment après l'interaction avec l'antigène spécifique : soit en optimisant l'affinité et la spécificité des LB, soit en se débarrassant des LB moins efficaces.

##### 2.2.2.2.1. Hypermutation somatique et commutation de classe des Ig

Les LB activées opèrent un processus de maturation des gènes d'Ig : l'hypermutation somatique (*SHM* pour *somatic hypermutation*) et la recombinaison de commutation de classe d'immunoglobulines (*CSR* pour *class switch recombination*), également appelée recombinaison de commutation isotypique ; Laffleur et al., 2014). Ces deux mécanismes requièrent l'enzyme désaminase AID (*activation-induced deaminase*), dont l'expression est activée par la fixation de l'antigène au BCR, et ont pour but d'améliorer l'affinité de l'Ig pour l'antigène, aboutissant ainsi à l'élimination plus efficace du pathogène. La désaminase AID se charge de l'hypermutation somatique qui consiste à introduire des mutations ponctuelles dans les régions variables (V) des gènes d'Ig fonctionnelle (segment VDJ productif). En retour, les LB exprimant l'Ig avec la plus forte affinité pour l'antigène sont sélectionnés. L'enzyme AID initie la CSR en ciblant des séquences d'ADN spécifiques appelées région « switch » du locus *IgH* ( $S\mu$ ) qui vont subir des cassures double-brin afin d'échanger les régions constantes  $\mu$  et  $\delta$  contre des régions constantes  $\gamma$ ,  $\alpha$  et  $\epsilon$  de la chaîne lourde des Ig, générant respectivement des IgG, IgA ou IgE, avec une fonction effectrice du BCR spécifique à chaque classe (Laffleur et al., 2014).

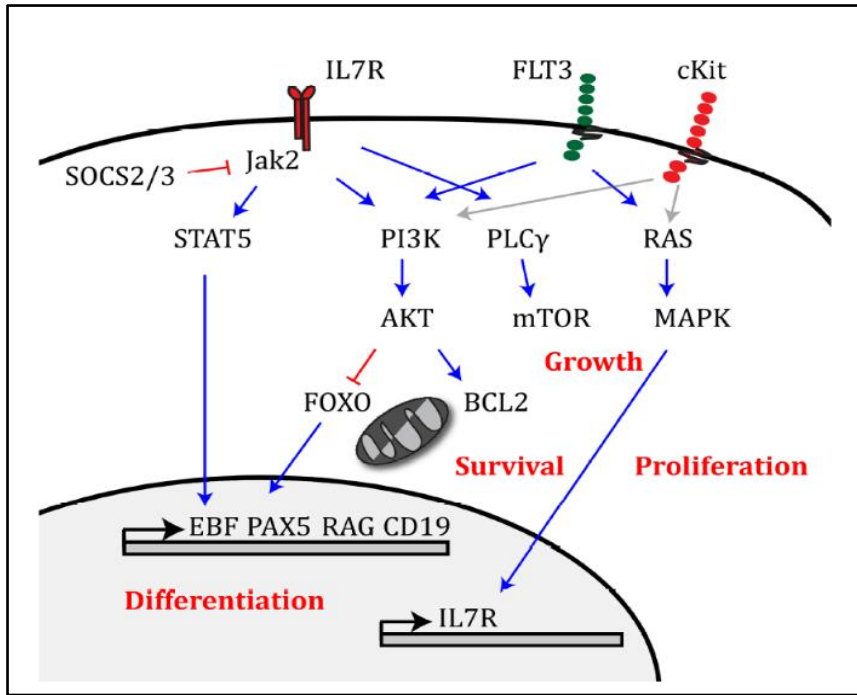


Figure 2.9 : Aperçu des voies de signalisation actives lors de la transition des CLP (*common lymphoid progenitors*) au stade pro-B (pré-B1) ; Figure issue de Petkau & Turner, 2019.

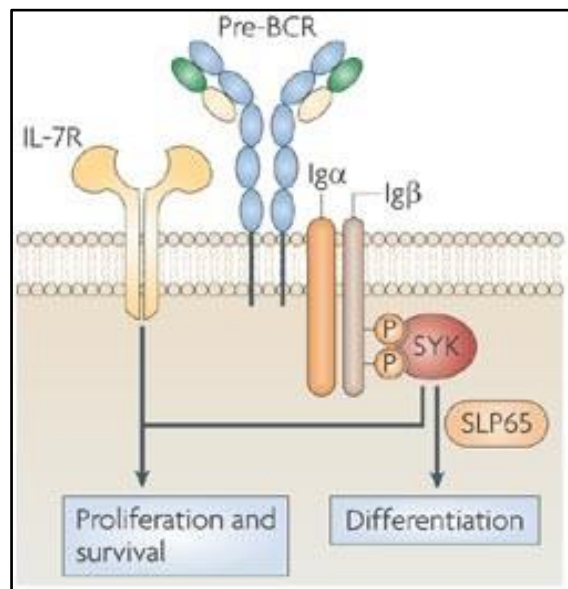


Figure 2.10 : Collaboration entre les signalisations de l'IL-7R et du pré-BCR du stade pro-B au stade immature B. Figure issue de Herzog et al., 2009.

#### 2.2.2.2.2. Recombinaison-suicide du locus des immunoglobulines

L'enzyme AID intervient occasionnellement dans un troisième mécanisme de mutation et de recombinaison affectant la région cis-régulatrice en 3' du locus *IgH* qui est transcrite.

Cette recombinaison-suicide du locus (*LSR* pour *locus suicide recombination*) correspond à une variante stérile de commutation isotypique, causant une délétion de la région constante (C) du locus *IgH*. Cela génère des LB n'exprimant plus d'Ig à leur surface, pourtant indispensable à leur survie cellulaire. Cette perte du BCR entraîne donc l'apoptose du LB (Péron et al., 2012).

### 2.2.3. Voies de signalisation principales de la lymphopoïèse B précoce

Durant le développement lymphoïde B, les cellules sont confrontées à différents microenvironnements qui les exposent à une gamme de cytokines, de chimiokines, de molécules d'adhésion et d'antigènes.

Chez la souris, les circuits de signalisation intracellulaire au cours de la lymphopoïèse B précoces sont principalement initiés par l'activation de l'IL-7R (récepteur à l'interleukine-7) et/ou du pré-BCR. Ils déclenchent des voies de signalisation qui s'intriquent, coopérant et s'équilibrant, pour adéquatement moduler la prolifération et la différenciation des cellules lymphoïdes B (Figure 2.10).

Ne possédant pas encore de pré-BCR, les cellules progénitrices lymphoïdes B (pré-pro-B et pro-B) dépendent essentiellement de la signalisation induite par l'activation de l'IL-7R, suppléée par les récepteurs Kit et Flt3 (Figure 2.9). Ces deux derniers font notamment appel à la voie RAS-MAPK pour promouvoir la prolifération (L.-X. Li et al., 2010; Linnekin, 1999).

#### 2.2.3.1. La signalisation de l'IL-7R

L'activation des récepteurs à l'IL-7 stimule la prolifération et la survie cellulaire en activant les axes de signalisation JAK3-STAT5 et PI3K (Hennighausen & Robinson, 2008; Hofmeister et al., 1999; Figure 2.11, Rickert, 2013).

La signalisation de l'IL-7R induit l'activation du facteur de transcription STAT5 (Goetz et al., 2004) tandis que la signalisation du récepteur Flt3 (Fms-like tyrosine kinase 3) active ERK (appartenant à la voie des MAPK) et Akt (appartenant à la voie PI3K) pour promouvoir synergiquement la prolifération d'ores et déjà au stade pré-pro-B (Åhsberg et al., 2010; Figure 2.9). Par ailleurs, la voie RAS-MAPK est également nécessaire à la prolifération des pro-B, avant la formation du pré-BCR et implique l'activation de la sérine/thréonine kinase Raf par p21Ras (Iritani et al., 1997; Figure 2.9).

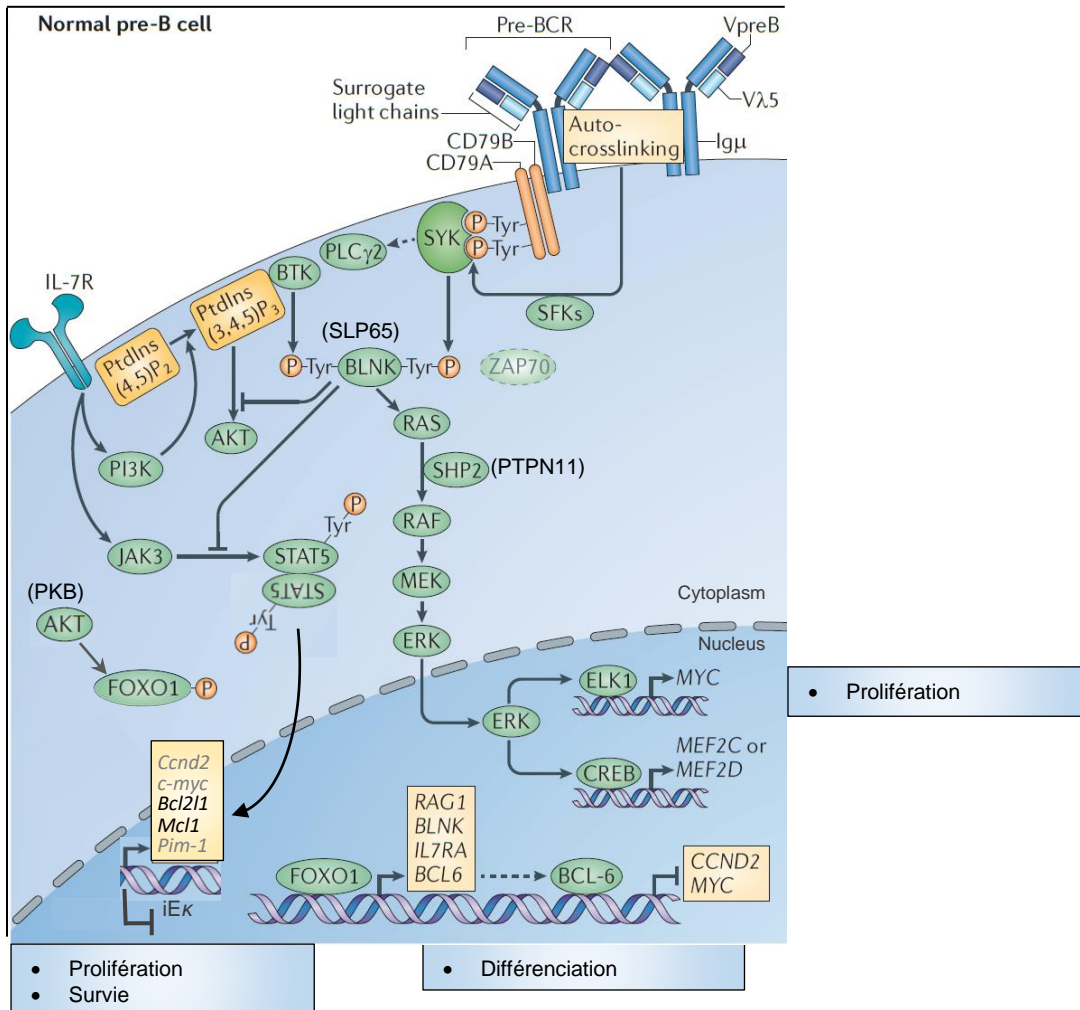


Figure 2.11 : Synergie des signaux de l'IL-7R et du pré-BCR chez la souris

La prolifération, la survie et la différenciation des cellules progéniteurs et précurseurs lymphoïdes B dépendent de l'activation des voies de signalisation JAK3/STAT5, PI3K/AKT(PKB)/FOXO1 et RAS/MAPK. Les cascades de signalisation sont initiées par la fixation d'IL-7 à l'IL-7R et par des interactions entre les pseudo-chaînes légères, représentant une activation ligand-indépendante dite autonome (*auto-crosslinking*). Figure modifiée à partir de Rickert, 2013).

Abréviations dans l'ordre alphabétique: **Bcl2l1**, B-cell lymphoma 2-like 1 (code B-cell lymphoma-extra large (Bcl-xL)); **BCAP**, B-cell PI3K adaptor; **BCL6**, B cell lymphoma 6; **BLNK**, B-cell linker protein; **BTK**, Bruton tyrosine kinase; **CCND2**, cyclin D2 ; **ERK**, extracellular signal-regulated kinase; **FOXO1**, forkhead box protein O1; **GAB2**, GRB2-associated-binding protein 2; **IL-7R**, interleukin-7 receptor; **ITAM**, immunoreceptor tyrosine-based activation motif ; **JAK3**, Janus kinase 3; **LYN**, Lck/Yes novel tyrosine kinase; **Mcl1**, Induced myeloid leukemia cell differentiation protein; **MEF2**, myocyte-specific enhancer factor 2; **MEK**, MAPK/ERK kinase; **PI3K**, phosphoinositide 3-kinase ; **SFKs**, SRC-family protein tyrosine kinases; **PKB**, protein kinase B; **PLCγ2**, phospholipase Cy2; **PtdIns(3,4,5)P<sub>3</sub>**, phosphatidylinositol-3,4,5-trisphosphate; **PtdIns(4,5)P<sub>2</sub>**, phosphatidylinositol- 4,5-diphosphate; **PTPN11**, tyrosine-protein phosphatase non-receptor type 11; **RAG1/2**, recombination-activating gene 1/2; **SHP2**, Src homology region 2 (SH2)-containing protein tyrosine phosphatase 2 ; **SLP65**, SH2-domain-containing leukocyte protein of 65 kDa; **STAT5**, signal transducer and activator of transcription 5; **SYK**, spleen tyrosine kinase; **ZAP70**, ζ-chain-associated protein kinase of 70kDa.

L'IL-7 est présente en abondance dans le stroma environnant des cellules pro-B dans lesquelles s'initie la recombinaison VDJ des *IgH*. En revanche, le niveau d'expression en IL-7 diminue ensuite progressivement. Dès lors que les cellules pro-B tardives commencent à exprimer le pré-BCR, celui-ci induit des boucles de régulation positive et négative de l'IL-7R. Cela conduira à la transition du stade pro-B au stade pré-B.

- L'expression du pré-BCR à la surface des pro-B tardives favorise l'expression de l'IL-7R à la surface cellulaire pour permettre l'expansion clonale des cellules exprimant un pré-BCR fonctionnel malgré une faible concentration en IL-7 environnante (Fleming & Paige, 2001).
- Puis l'expression de l'IL-7R est réprimée en réponse à la signalisation du pré-BCR pour empêcher les 'petites' pré-B tardives de répondre à l'IL-7 pendant qu'elles recombinent les gènes de la chaîne légère de l'immunoglobuline (Marshall et al., 1998).

Plusieurs voies de signalisation collaborent pour promouvoir successivement la prolifération (expansion clonale des clones pré-B) et la progression de la différenciation par l'induction de la recombinaison V(D)J des *IgL*.

#### 2.2.3.1.1. L'axe de signalisation IL-7R-JAK-STAT

Les enzymes Janus kinase 3 (JAK3) sont essentiellement associées aux récepteurs cytokiniques (IL-7R en l'occurrence), sous la membrane plasmique. Dès lors que l'IL-7 se lie à son récepteur, les IL-7R s'agrègent et recrutent les JAK3 qui s'auto-phosphorylent et phosphorylent les queues intracytoplasmiques des oligomères d'IL-7R. La juxtaposition des protéines JAK3 favorisent leur transphosphorylation menant à une activité kinase accrue. Les protéines JAK3 activées recrutent à leur tour les facteurs STAT5 (*signal transducer and activator of transcription 5*) au niveau des IL-7R et les phosphorylent pour transduire le signal : une fois phosphorylés, ils se dissocient du complexe IL-7R-JAK3 et se dimérisent (en homo- ou hétérodimères). Seuls les dimères de STAT phosphorylés peuvent être importés dans le noyau et remplir leur fonction de facteur de transcription.

La voie JAK3-STAT5 joue un rôle prédominant dans le développement des cellules progéniteurs B.

L'activation de STAT5 permet singulièrement de garantir la survie cellulaire des pro-B en activant la transcription du gène anti-apoptotique *Mcl-1*, un autre membre de la famille *Bcl2* (Malin et al., 2010), alors que la survie des pré-B dépend de l'activation de *Bcl-xL* par STAT5 (Dumon et al., 1999).



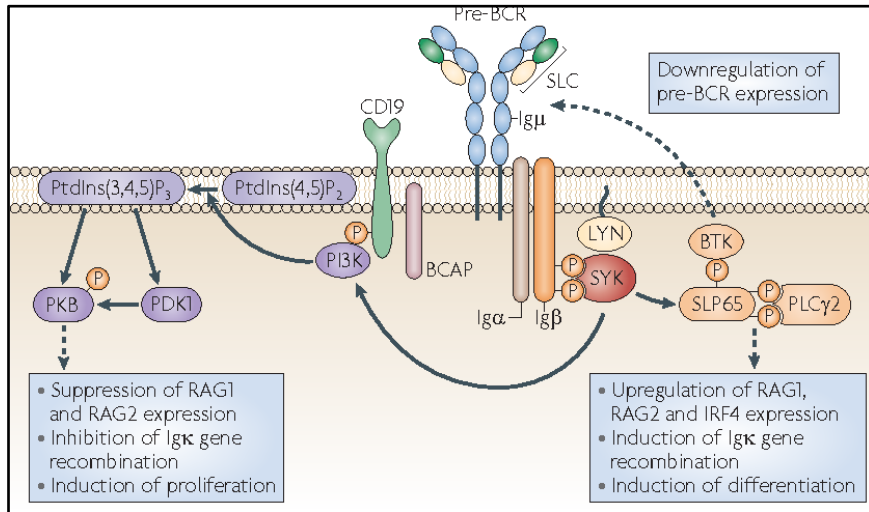


Figure 2.12 : Le rôle central de SYK en aval de la signalisation du pré-BCR.

Figure issue de Herzog et al., 2009.

Au stade pro-B, un rôle majeur de STAT5 activé par la signalisation de l'IL-7R consiste à se lier à l'*enhancer* intronique *iE kappa* pour empêcher la transcription germinale d'*Igk* et les réarrangements prématurés au locus *Igk* (Malin et al., 2010). Par ailleurs, STAT5 est un médiateur de la prolifération des cellules pré-B (Sexl et al., 2000). Il est notamment capable d'activer la transcription de gènes codant des régulateurs du cycle cellulaire tels que la cycline D1 et la cycline D2 (spécifique à la transition G1 / S), ainsi que le proto-oncogène codant le facteur de transcription Myc (Alvarez & Frank, 2004; Matsumura et al., 1999).

L'inactivation de STAT5 survient lorsque l'expression de l'IL-7R est réprimée par la signalisation du pré-BCR dans les petites pré-B tardives (Alvarez & Frank, 2004; Matsumura et al., 1999; Nakayama et al., 2009), ce qui limite l'expansion des pré-B en cours de leur recombinaison critique.

#### 2.2.3.1.2. La voie de signalisation PI3K

En présence abondante d'IL-7 dans la niche des pro-B, la forte activation de l'IL-7R promeut une robuste signalisation PI3K (Figure 2.11) avec le concours de la signalisation précoce du pré-BCR nouvellement formé (*voir la voie PI3K sous le pré-BCR*).

#### 2.2.3.2. La signalisation du pré-BCR

Globalement, la signalisation du pré-BCR induit l'exclusion allélique du gène *IgH*, stimule la prolifération des cellules pré-B et active la recombinaison au locus *IgL* (Geier & Schlissel, 2006). Concernant l'expansion clonale, les signaux mitogéniques induisant l'entrée en G1 du cycle cellulaire et la progression vers la phase S (phase de réplication de l'ADN) sont notamment transmis par la voie PI3K-AKT (AKT étant aussi appelé PKB) et la voie RAS-RAF-MEK1-ERK (MEK1 signifiant *MAPK/ERK kinase 1* et ERK signifiant *extracellular regulated kinase*).

##### 2.2.3.2.1. Initiation de la signalisation du pré-BCR

L'agrégation de pré-BCR fonctionnels à la surface membranaire suite à leur assemblage peut suffire à induire sa signalisation (Ohnishi & Melchers, 2003) ; et il peut être davantage activé par des ligands provenant du stroma tels que la Galectine-1 qui favorise le regroupement des pré-BCR en cluster-s (Gauthier et al., 2002).

##### 2.2.3.2.2. Transduction du signal et plateformes d'interaction

Les glycoprotéines membranaires CD79a et CD79b forment un hétérodimère associé au pré-BCR. Cet hétérodimère assure la transduction du signal du pré-BCR par le biais de leurs motifs d'activation ITAM (*immunoreceptor tyrosine-based activation motif*).



La phosphorylation des motifs ITAMs par les tyrosines kinases de la famille SRC (SFKs, *Src family of protein tyrosine kinases*) ancrées dans la membrane plasmique – telles que LYN (*Lck/Yes Novel tyrosine kinase*) – conduit à l'amplification du signal, en facilitant la liaison et l'activation de la tyrosine kinase SYK (*spleen tyrosine kinase*) et de la tyrosine kinase ZAP70 (*ζ-chain-associated protein kinase of 70kDa*) dans une moindre mesure (Geier & Schlissel, 2006; Rickert, 2013). La kinase ZAP70, faisant partie de la famille de SYK, joue un rôle redondant par rapport à SYK. SYK s'associe à LYN pour, à leur tour, recruter et induire la phosphorylation de divers substrats clés tels que PI3K et BLNK (alias SLP65). BLNK phosphorylée est capable de réguler le cycle cellulaire à la fois positivement et négativement. Ainsi SYK et BLNK jouent des rôles primordiaux à la fois dans la survie cellulaire, la prolifération et la différenciation des précurseurs lymphoïdes (Herzog et al., 2009). Le pré-BCR et l'IL-7R peuvent coopérer pour promouvoir la prolifération et la survie des cellules 'larges' pré-B précoces (Corfe & Paige, 2012; Milne & Paige, 2006). Les signaux prolifératifs de l'IL-7R et du pré-BCR convergent notamment vers l'activation de la voie de signalisation PI3K-AKT (Herzog et al., 2009; Rickert, 2013; Figures 11 et 12).

#### 2.2.3.2.3. La voie PI3K-AKT(PKB)-FOXO1

D'une part, SYK se charge d'activer la voie PI3K (*phosphoinositide 3 kinase*) en synergie avec la signalisation de l'IL-7R (Figures 2.11 et 2.12). SYK participe à la phosphorylation du co-récepteur CD19 et/ou de la protéine adaptatrice BCAP (*B-cell PI3K adaptor*) responsable-s du recrutement et de l'activation de PI3K. PI3K activée permet de phosphoryler PtdIns(4,5)P<sub>2</sub> en PtdIns(3,4,5)P<sub>3</sub> (PIP<sub>3</sub>). Ce second messenger recrute et active alors la sérine/thréonine kinase PKB (*protein kinase B*, alias Akt), qui est le médiateur majeur pour contrôler la prolifération en aval de PI3K. PKB/AKT va notamment inactiver les facteurs de transcription FOXO (*forkhead box protein*, FOXO1, FOXO3a et FOXO4) en les phosphorylant. La phosphorylation du facteur de transcription FOXO1 par AKT le séquestre dans le cytoplasme et le conduit à sa dégradation. D'autre part, SYK multi-phosphoryle la protéine adaptatrice BLNK (*B cell linker protein*) qui présente alors de multiples sites d'amarrage (*docking sites*) tout comme PIP<sub>3</sub>. Ainsi, BLNK s'associe à d'autres médiateurs de signalisation tels que BTK (*Bruton tyrosine kinase*) pré-activée par la voie PI3K et PLCγ2 (*phospholipase C gamma 2*), pour former un complexe moléculaire qui sert de plateforme d'interaction aux diverses voies de signalisation en aval du pré-BCR et du BCR (Herzog et al., 2009). La protéine BLNK phosphorylée participe à l'avancement dans la maturation du pré-BCR en induisant la répression de l'expression la pseudo-chaîne légère (Geier & Schlissel, 2006). Une fois les cellules pré-B (fraction C') multipliées, BLNK sous forme phosphorylée met un terme au signal prolifératif en induisant l'inactivation de STAT5 (en empêchant la phosphorylation de STAT5 par JAK3).



BLNK inhibe également l'activité d'AKT (Herzog et al., 2009), levant ainsi la phosphorylation répressive de FOXO1 exercée par AKT. FOXO1 peut alors activer la transcription des gènes codant les recombinaisons Rag1 et Rag2 (Amin & Schlissel, 2008) ainsi que celle du gène *Bcl-6* (*B cell lymphoma-6*). BCL-6 peut réprimer la transcription de *Myc* et de *Ccnd2* (codant la cycline D2), ce qui atténue davantage la prolifération et induit la mise en quiescence des cellules pré-B tardives qui subissent la recombinaison au locus *IgL* (Duy et al., 2010; Nahar et al., 2011). De plus, la signalisation du pré-BCR induit également l'expression du facteur de transcription IRF4 (*interferon-regulatory factor 4*) qui se fixe à l'*enhancer* en 3' du locus de la chaîne légère  $\kappa$  (*3'Ek*). Ceci mène à la recombinaison au locus *Igk* et permet la progression de la différenciation des cellules pré-B tardives en cellules immatures avec un BCR (Herzog et al., 2009).

#### 2.2.3.2.4. La voie RAS-RAF-MEK-ERK

Le pré-BCR active la voie de signalisation RAS-MAPK (*Mitogen-Activated Protein Kinases*, induisant la mitose) pour transduire des signaux mitogéniques, menant à l'expansion sélective des cellules pré-B à une faible concentration en IL-7 environnante (Fleming & Paige, 2001). Recrutée par la protéine adaptatrice BLNK (*B cell linker protein*), la phospholipase PLC $\gamma$ 2 se lie à la forme activée de RAS pour promouvoir l'activation de la kinase ERK (*extracellular regulated kinase*; Imamura et al., 2009). ERK1 et ERK2 tiennent un rôle déterminant dans la signalisation du pré-BCR. Parmi leurs nombreux substrats, les kinases ERK1/2 phosphorylent notamment le facteur de transcription ELK1 qui active la transcription de *Myc* (Yasuda et al., 2008), et le facteur de transcription CREB1 (*cyclic AMP-responsive element-binding protein 1*) ciblant des gènes de la famille des facteurs de transcription Mef2c/d (*myocyte enhancer factor c/d*) régulant l'activité transcriptionnelle requise pour la progression de la différenciation des cellules pré-B (Herglotz et al., 2016).

#### 2.2.3.3. La signalisation du BCR

##### 2.2.3.3.1. Initiation de la signalisation

À l'inverse du pré-BCR, la signalisation du BCR (ou récepteur antigénique) dépend essentiellement de la fixation d'un antigène. L'expansion clonale des cellules B matures peut survenir en réponse à la fixation d'un l'antigène au BCR, mais également en réponse à des signaux mitogéniques divers. Néanmoins, des signaux toniques émanent du BCR pour garantir la survie des cellules B naïves.

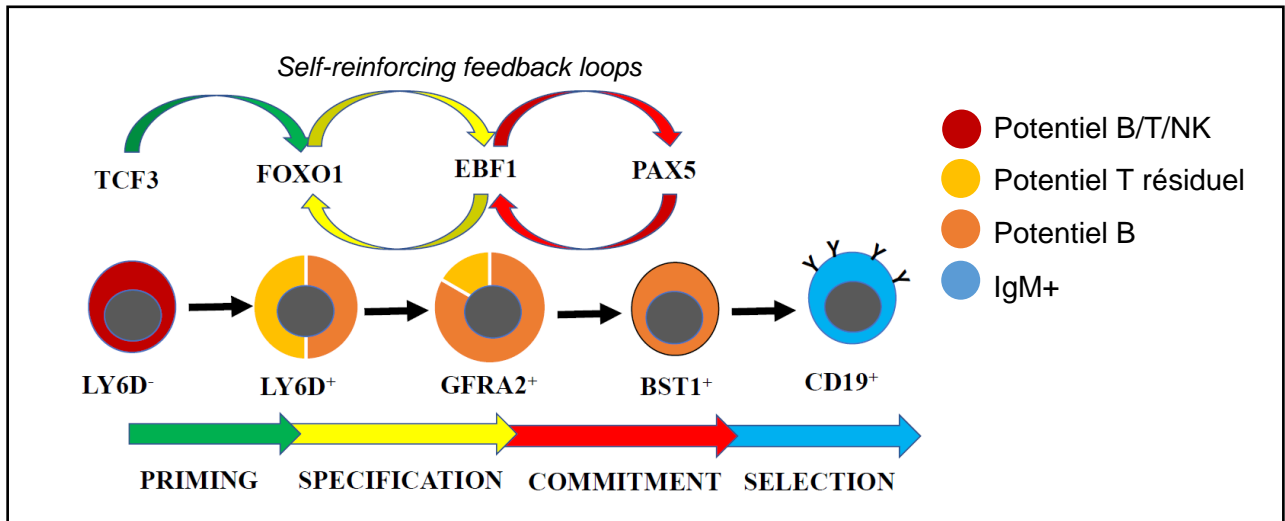


Figure 2.13 : Schéma des réseaux transcriptionnels impliqués dans le « priming », la spécification, l'engagement et la sélection dans le développement précoce des cellules lymphoïdes B.

En aval des ALPs Ly6d-négatives conservant un potentiel lymphoïde B/T et de cellules NK (fond rouge), les BLPs (*B-cell Lymphoid Progenitors*) peuvent être divisés en BLP1 (Ly6d+), BLP2 (*GDNF family receptor 2*<sup>+</sup>) et BLP3 (*bone marrow stroma cell antigen 1*<sup>+</sup>) s'engageant progressivement dans la lignée lymphoïde B (Jensen et al., 2018). Figure adaptée de Sigvardsson, 2018. *Nota Bene* : TCF3 est aussi nommé E2A.

#### 2.2.3.3.2. Voies de signalisation du BCR

La signalisation tonique requiert les mêmes médiateurs que la signalisation du pré-BCR, incluant notamment PLC $\gamma$  qui active la voie PI3K pour assurer la survie des cellules B matures (Srinivasan et al., 2009). Ce médiateur met notamment en jeu la voie NF- $\kappa$ B pour réguler la survie et la prolifération cellulaire. L'axe de signalisation RAS–RAF–MAPK/ERK kinase (MEK)–ERK est aussi induit pour aboutir à l'inhibition de l'apoptose (Lazarus et al., 1993).

#### 2.2.4. Hiérarchie transcriptionnelle orchestrant la différenciation B

Les voies de signalisation ainsi enclenchées par des récepteurs activés (par des vagues de cytokines et facteurs de croissance dépendamment des niches stromales) permettent de contrôler et/ou d'induire l'expression de facteurs de transcription (FTs). En réponse, les FTs régulent ces voies de signalisation, activent des programmes spécifiques d'une lignée au détriment de ceux d'autres lignées alternatives et répriment également les programmes transcriptionnels de maintenance de multipotence. Les programmes de priming, de spécification, d'engagement et de différenciation lymphoïde B sont régulés par des réseaux de FTs (Figure 2.13) collaborant étroitement avec des facteurs épigénétiques (FEs). À tous les stades de développement, ils sont directement connectés les uns aux autres parmi les circuits de régulation. La régulation orchestrée par les FTs et les FEs est essentielle à l'initiation, la promotion et la maintenance des transcriptomes spécifiques de la lignée B. Certains de ces deux types de facteurs ont comme atout de pouvoir alterner de rôles entre activateur et répresseur transcriptionnel (Miyai et al., 2018). Bien que l'hématopoïèse se base sur l'expression séquentielle de tels facteurs, les transitions s'exécutent de façon dynamique et continue. Ainsi sur le plan cellulaire, la lymphopoïèse correspond à un continuum d'états cellulaires transitionnels entre des stades distincts (discrets). Sur le plan moléculaire, la spécification des progéniteurs et leur engagement subséquent spécifiquement dans la lignée lymphoïde B s'organisent en trois grandes étapes rapides (Miyai et al., 2018). Elles se modélisent par trois vagues d'expression de facteurs transcription, complémentées par un réseau d'autres facteurs de transcription à l'expression plus constante (comprenant par exemple Ikaros).

- La phase initiale (*early wave*) réalise un priming aspécifique qui a pour but d'inhiber l'expression de gènes inappropriés tels que ceux impliqués dans le maintien d'un état multipotent ou dans la spécification en lignées alternatifs.
- La phase intermédiaire (*mid-wave*) procède à un changement du statut du cycle cellulaire et du métabolisme.
- La dernière phase de la spécification (*late wave*) fait intervenir les facteurs de transcription qui tiennent un rôle spécifique dans l'engagement de la lignée B (Figure 2.13).



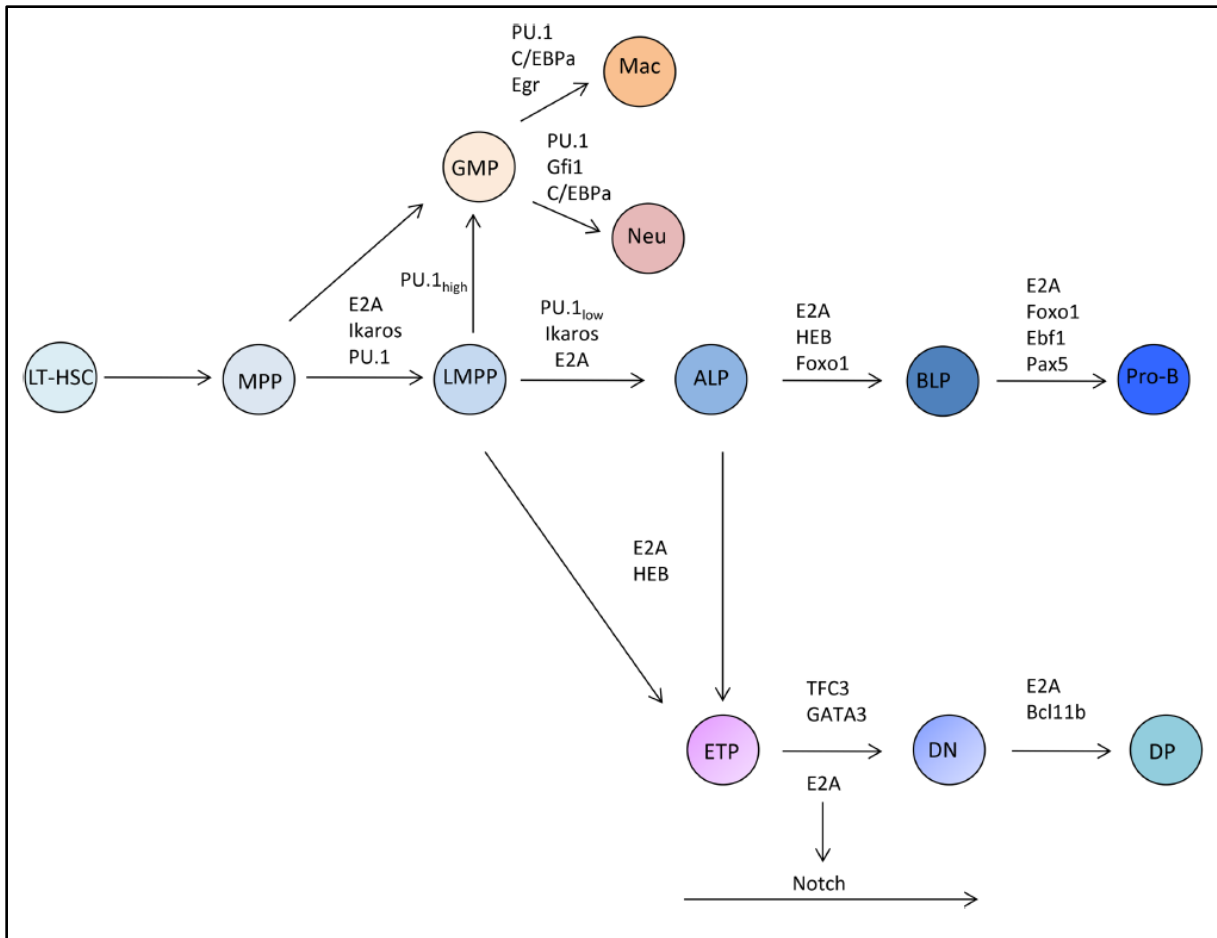


Figure 2.14A : Régulateurs transcriptionnels gouvernant l'hématopoïèse.

Les facteurs de transcription associés aux points de contrôle durant l'hématopoïèse sont indiqués. Figure issue de Murre, 2018.

Abréviations dans l'ordre alphabétique : **ALP**, all lymphoid progenitors ; **BLP**, B-cell-biased lymphoid progenitors ; **DN**, double-negative CD4<sup>+</sup>CD8<sup>-</sup> thymocytes ; **DP**, double-positive CD4<sup>+</sup>CD8<sup>+</sup> cells ; **ETP**, early T-cell progenitors ; **GMP**, granulocyte/macrophage progenitors ; **LMPP**, lymphoid-primed multipotent progenitors ; **LT-HSC**, long-term hematopoietic stem cells ; **Mac**, Macrophages ; **MPP**, multipotent progenitors ; **Neu**, Neutrophils.

	MLP	CLP	Fr. A	Fr. B/C
	GL	50% DJ	90% DJ	200% DJ
PU.1	+++	++	++	+
E2a	+	+	+	++
EBF	-	+	++	+++
Pax5	-	+	++	+++
Rag2-GFP	low (100%)	high (100%)	high (100%)	high (100%)
reconstituting capacity	myeloid T-lineage B-lineage			
in vitro development	myeloid T-lineage B-lineage	+	+	-
	+	+	+	+

Figure 2.14B : Diagramme récapitulatif des stades précoces de développement du lignage B, indiquant les potentiels de différenciation (capacité de reconstitution B/T/myéloïde, et plasticité de lignée *in vitro*). Figure issue de Rumfelt et al., 2006.

#### 2.2.4.1. Réseaux transcriptionnels établissant l'identité cellulaire B

Le rôle primordial des différents facteurs de transcription au cours de la lymphopoïèse B a été essentiellement déterminé à l'aide de modèles murins dont les gènes codant ces facteurs de transcriptions ont été invalidés, ce qui a déterminé une hiérarchie transcriptionnelle de la différenciation B (Hagman & Lukin, 2006) impliquant *Spi1*, *Ikzf1*, *Foxo1*, *Tcf3*, *Ebf1* et *Pax5* (Boller & Grosschedl, 2014; Y. C. Lin et al., 2010; Miyai et al., 2018; Sigvardsson, 2018).

- Priming lymphoïde par Ikzf1, PU.1 et E2A (Figure 2.14)

Certains FTs associés au lignage B sont d'ores et déjà faiblement exprimés dans les progéniteurs non engagés (Månsson et al., 2007). Le priming lymphoïde dépend principalement des FTs *Ikzf1* (également nommé Ikaros, Yoshida et al., 2006), PU.1 (codé par *Spi1*, Scott et al., 1997) et *Tcf3* (également nommé E2A, Dias et al., 2008). Ces FTs permettent de générer un environnement épigénétique permissif en faveur du développement des LB (Månsson et al., 2007) et activent notamment la transcription de gènes codant les enzymes Tdt (*Dnnt*) et Rag (H. Igarashi et al., 2002; Reynaud et al., 2008). E2A active faiblement des gènes tels *Rag1*, *Rag2*, et *Dnnt* dans les BLP (Y. C. Lin et al., 2010). *Ikzf1* induit la transcription de *Gfi1* (*Growth Factor Independent 1 Transcriptional Repressor*) qui va réprimer *Spi1* (Spooner et al., 2009). Cette boucle de régulation mène à la perte du potentiel myéloïde des progéniteurs hématopoïétiques (Figure 2.14B). À forte concentration, PU.1 active un programme transcriptionnel érythro-myéloïde dans les progéniteurs alors qu'à plus faible dose, PU.1 initie l'engagement des MPP dans le lignage lymphoïde (DeKoter & Singh, 2000) en stimulant notamment l'expression du récepteur de l'interleukine-7 (DeKoter et al., 2002). En effet, l'interleukine-7 possède un rôle instructif pro-lymphoïde (Dias et al., 2005), en activant le FT STAT5 dans les CLP. D'autre part, E2A initie un programme transcriptionnel spécifique au lignage B en activant *Foxo1* dans les CLP (Welinder et al., 2011). L'activation combinée de STAT5 et de FOXO1 induit la transcription d'*Ebf1* (Dias et al., 2005; Zandi et al., 2008). La force de ces réseaux transcriptionnels réside dans leur capacité à s'auto-exacerber par des boucles de régulation positive. Ainsi *Ebf1* maintient en retour un niveau d'expression élevé en FOXO1 pour garantir la stabilité de sa propre expression (Mansson et al., 2012).

Une fois stablement exprimé, *Ebf1* coopère avec E2A pour directement activer la transcription de nombreux gènes B-spécifiques tels que les gènes codant la pseudo-chaîne légère du pré-BCR (Y. C. Lin et al., 2010; O'Riordan & Grosschedl, 1999; Sigvardsson et al., 1997).

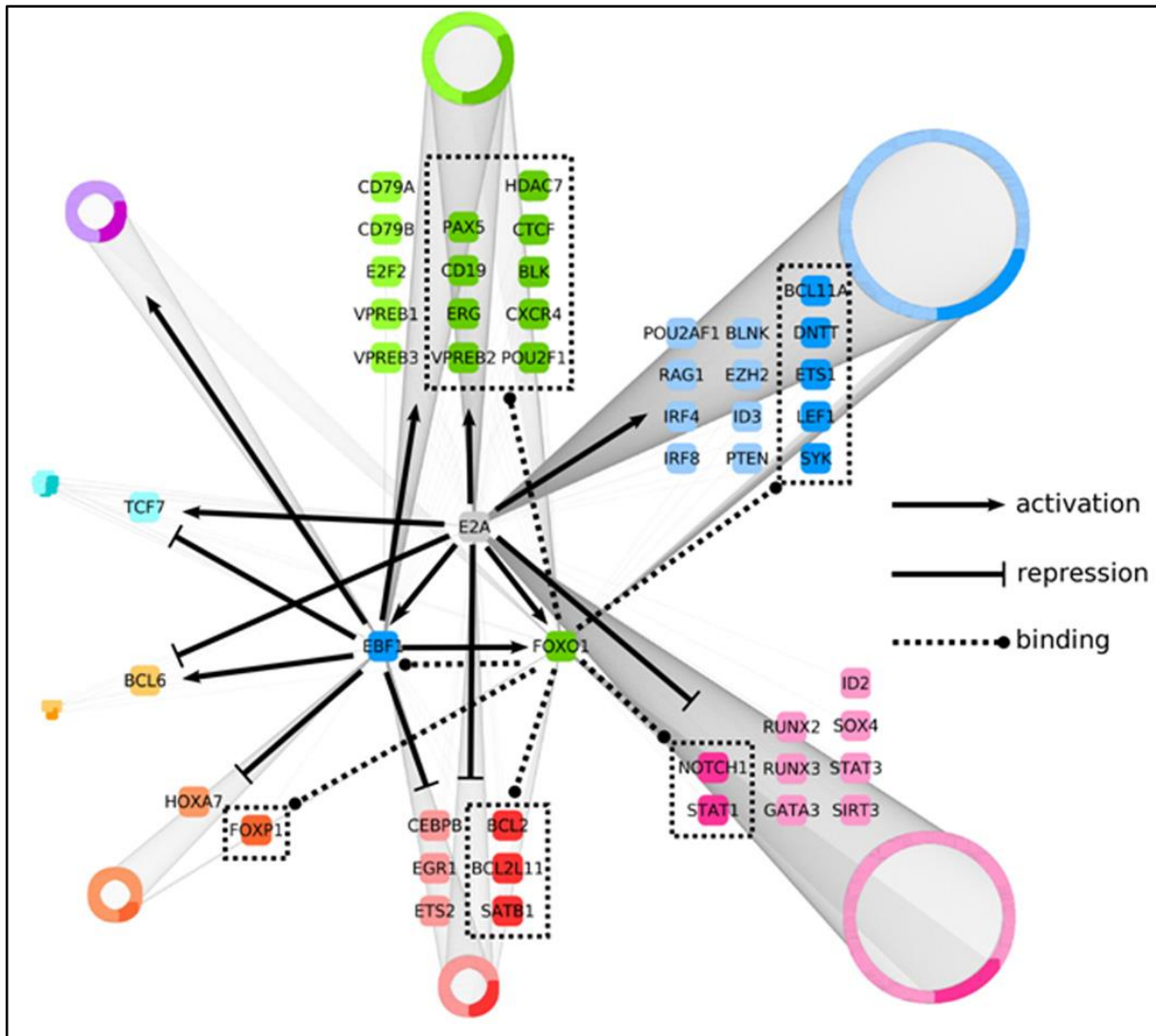


Figure 2.15 : Réseau de régulation reliant dans une voie commune les activités d'un ensemble de régulateurs transcriptionnels, de composants de signalisation et de facteurs de survie dans les cellules lymphoïdes B au cours de la lymphopoïèse.

Une analyse pangénomique intégrative, avec le logiciel Cytoscape, des liaisons entre protéines et ADN et les changements de niveaux de transcrits via l'utilisation de pré-pro-B déficientes en E2A, de pré-pro-B déficientes en EBF1 et de cellules pro-B contrôles a été effectuée pour identifier les cibles transcriptionnelles d'E2A et EBF1, ainsi que les gènes liés par FOXO1. Les 2045 gènes cibles de régulation en résultant ont été divisés en huit groupes basés sur leurs diminution ou augmentation de niveaux de transcrits dans les cellules pro-B en comparaison avec les cellules déficientes en E2A ou EBF1. Les lignes à flèches représentent l'activation de gènes alors que les lignes à barres représentent les répressions transcriptionnelles. Les lignes pointillées indiquent les gènes dont les loci sont occupés par FOXO1. Les gènes mentionnés pour chaque groupe jouent des rôles cruciaux dans l'engagement de cellules B ou sont de potentiels régulateurs. Tous les autres gènes sont regroupés et symbolisés par des anneaux dont la taille est proportionnelle au nombre de gènes membres. Les couleurs plus vives représentent les loci occupés par FOXO1. Figure issue de (Y. C. Lin et al., 2010).

- Spécification et engagement par l'induction d'Ebf1 et de Pax5 (Figures 2.14 et 2.15)

E2a et Ebf1 agissent ainsi en synergie pour finalement induire l'expression de *Pax5* (O'Riordan & Grosschedl, 1999). Ebf1 active directement la transcription de *Pax5* (Zandi et al., 2008) en s'associant avec Stat5 au niveau d'un élément régulateur en amont de l'exon 1A de *Pax5* (Hirokawa et al., 2003) et en collaboration avec PU.1 se fixant à l'*enhancer* B-spécifique de *Pax5* (Decker et al., 2009). Ebf1 et Pax5 sont impliqués dans une ultime boucle d'auto-régulation positive et d'activation réciproque pour obligatoirement maintenir stablement leur expression respective et ainsi préserver l'identité cellulaire B : Pax5 est une cible directe d'Ebf1, tandis que EBF1 et PAX5 ciblent des régions promotrices respectivement proximale et distale d'*Ebf1* (Roessler et al., 2007).

E2A, Ebf1 et Pax5 apportent une contribution majeure à la spécification et à l'engagement dans la lignée lymphoïde B (Nutt & Kee, 2007). Pax5 doit notamment réprimer *Flt3* pour initier l'engagement des cellules dans la lignée lymphoïde B (Holmes et al., 2006). E2a, Ebf1 ainsi que Pu.1 se fixent à des séquences cis-régulatrices du locus *IgH* et/ou activent la recombinaison des clusters D<sub>H</sub>-J<sub>H</sub> (Goebel et al., 2001; Hesslein & Schatz, 2001).

EBF1 et PAX5 collaborent pour initier les programmes transcriptionnels spécifiques de la lignée B, impliqués dans la transduction du signal, l'adhésion, la migration cellulaire, le métabolisme cellulaire et les fonctions immunitaires (Boller & Grosschedl, 2014; Schebesta et al., 2007). D'ailleurs, Pax5 et/ou Ebf1 pourraient être directement impliqués dans la régulation de la totalité des gènes B-spécifiques (Rothenberg, 2014). E2a coopère avec Ebf1 et Pax5 pour transcrire de nombreux gènes spécifiques de la lignée B tels que *Rag1* et *Rag2*, *Tdt* (*Dntt*), *Blk*, *Blnk* et *Pax5* qui active spécifiquement la transcription de *Cd19* codant un récepteur immunitaire. Les gènes codant les composants du complexe pré-BCR (*Cd79a*, *Cd79b*, *VpréB1*, *Igll1* codant *lambda-5*) font notamment partie des cibles transcriptionnelles directes de Pax5 et Ebf1 (Y. C. Lin et al., 2010; Revilla-I-Domingo et al., 2012; Treiber et al., 2010)

Au stade pro-B, *Ikzf1* et Ebf1 modulent l'expression de *lambda-5*. Puis au stade pré-B, Pax5 cible et active *Ikzf3* (codant Aiolos, un membre de la famille Ikaros) qui coopère avec *Ikzf1* pour réprimer la pseudo-chaîne légère (Thompson et al., 2007).

E2A régule également *Irf4* (*interferon regulatory factor 4*) et *Irf8* (Lin et al., 2010) conduisant *Irf4* à notamment intervenir dans le contrôle de la recombinaison du locus *Igk* (Shaffer et al., 2009). Les complexes de protéique impliquant *Irf4* régulent les loci *IgH* et *IgL*, et se lient à l'*enhancer* B-spécifique de la lignée lymphoïde de *Pax5* (Decker et al., 2009; Figure 2.16).

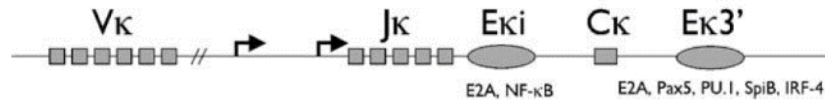


Figure 2.16 : Schéma du locus germinale Igκ incluant les segments géniques V et J, la région exonique constante, l'enhancer intronique (Ei) et l'enhancer en 3'(E3).

Les sites de liaison des facteurs de transcription sont indiqués sous leur-s enhancer-s respectif-s. Les flèches représentent les promoteurs des transcrits germinaux. Figure issue de (Geier & Schlissel, 2006).

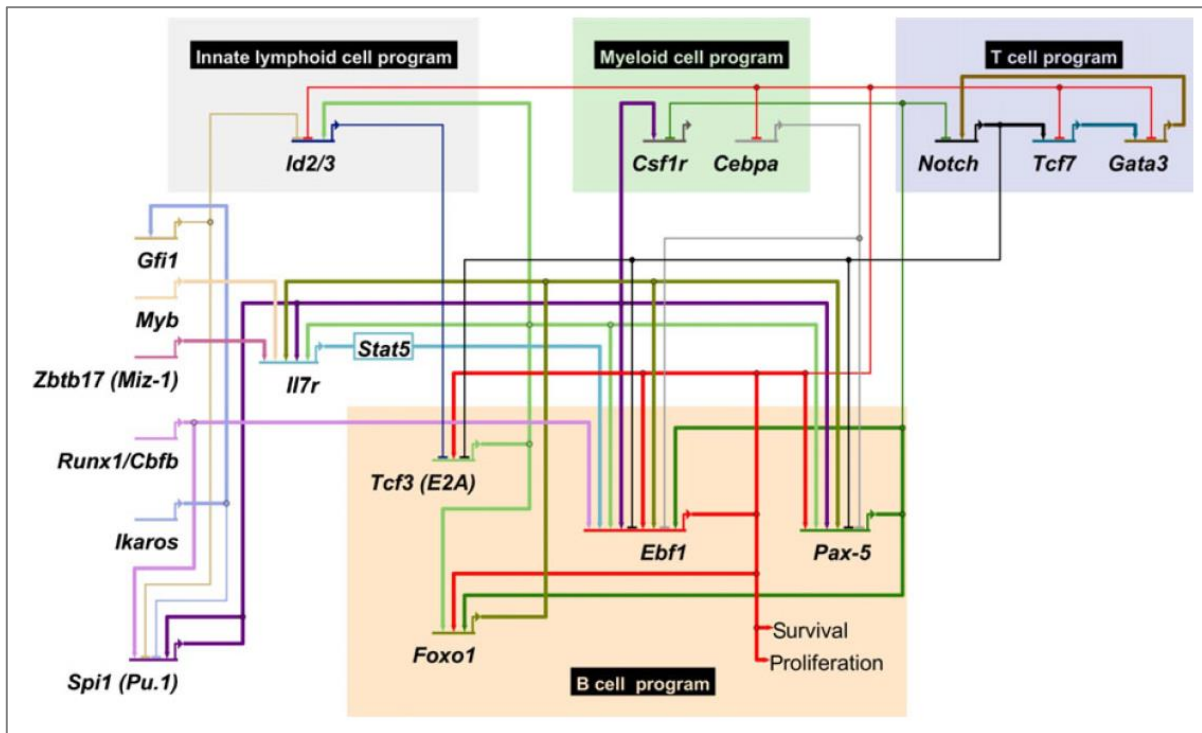


Figure 2.17 : Réseaux de régulation supervisant la lymphopoïèse B précoce.

Les facteurs clés impliqués dans le « priming » sont indiqués sur la gauche du schéma alors que les facteurs de transcription déterminants de la spécification et l'engagement du lignage B sont indiqués dans le cadre de fond orange. Les gènes favorisant les lignées alternatives sont indiqués sous fond gris (lignage lymphoïde inné), vert (lignage myéloïde) et violet (lignage lymphoïde T). Chaque régulation positive est représentée par une épaisse ligne à flèche alors que chaque répression est représentée par une fine ligne à barre. Figure issue de (Boller & Grosschedl, 2014).

TCF3 (E2A) fonctionne seulement comme un activateur transcriptionnel alors que Ebf1 et Pax5 peuvent en outre fonctionner comme des répresseurs transcriptionnels pour empêcher la différenciation dans toute voie alternative autre que lymphoïde B et davantage fixer la différenciation B. Pour ce faire, Ebf1 et Pax inhibent la transcription de gènes requis pour la différenciation des cellules dendritiques, NK, lymphoïdes B ou myéloïdes (Figure 2.16 ; Boller & Grosschedl, 2014; Delogu et al., 2006; Nechanitzky et al., 2013). Par exemple, Pax5 réprime le gène codant le récepteur *Csf1r* (CD115 ou M-CSFR), pour éliminer la possibilité de différenciation macrophagique (Tagoh et al., 2004). Ebf1 réprime la transcription de gènes associés à des voies alternatives tels que *Cebpa*, *Id2*, *Id3* et *Spi1* (Pongubala et al., 2008; Thal et al., 2009; Treiber et al., 2010). C/EBP- $\alpha$  et PU.1 sont des régulateurs transcriptionnels pro-myéloïdes majeurs et *Id2* est un important régulateur du développement des cellules Natural Killer. De plus, *Id2* et *Id3* font partie de la famille des protéines *Id* (*inhibitors of DNA binding*, *Id1-Id4*) qui en l'occurrence inhibent l'activité de liaison à l'ADN d'E2A. Elles remplissent des fonctions communes d'inhibition de la différenciation, de promotion de la prolifération et de maintenance de la multipotence (Lasorella et al., 2014). La répression d'*Id2* et *Id3* par Ebf1 est pertinemment requise pour la progression au-delà du stade pré-pro-B. Cela permet d'amplifier l'expression d'E2A situé à la source des boucles de régulation positive (Figure 2.13) et ainsi bloque irréversiblement la spécification et l'engagement dans la lignée lymphoïde B. Par ailleurs, pour supprimer le potentiel T, Pax5 réprime la transcription de *Notch1*, crucial pour la lymphopoïèse T (Souabni et al., 2002), tandis qu'Ebf1 réprime *Gata3* (Banerjee et al., 2013) et *Tcf7*, codant TCF-1 qui joue un rôle essentiel de *gatekeeper* dans la spécification et la différenciation du lignage T (Nechanitzky et al., 2013). En outre, le facteur *Bach2* (*BTB and CNC homologue 2*), dont l'expression est induite par Pax5 (Schebesta et al., 2007), joue d'ores et déjà un rôle important dans la restriction du lignage myéloïde dès le stade des CLP (Itoh-Nakadai et al., 2014).

Certains FTs comme *Myc* et *Ikaros* (*Ikzf1*) sont exprimés plus fréquemment durant le développement de la lignée lymphoïde B mais ils ont la particularité de basculer d'un rôle de répresseur à celui d'activateurs transcriptionnels, ce qui leur permet d'exercer des fonctions spécifiques à chaque étape de l'engagement bien qu'ils soient exprimés continuellement (Miyai et al., 2018).

Dans les cellules immatures B, suite par la stimulation BCR, *Aiolos* (*Ikzf3*) et *Helios* (*Ikzf2*) contrôlent l'apoptose notamment par le biais de la régulation transcriptionnelle des PKCs, les protéines kinases Cs (Kikuchi et al., 2009, 2011). TCF3 (E2A) est également impliqué dans le contrôle de l'apoptose médiée par la signalisation BCR en régulant la transcription des gènes de la survivine, de l'inhibiteur de l'apoptose-2 (IAP2) et de la caspase-8 (Toyonaga et al., 2009).

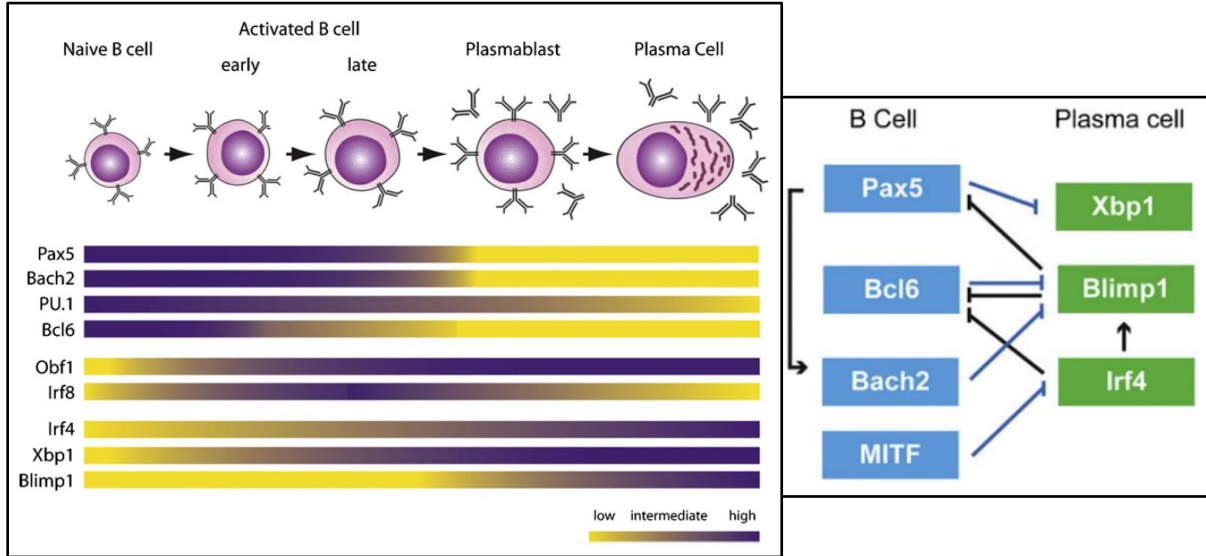


Figure 2.18 (à gauche) : Le réseau de régulation génique contrôlant la transition multi-étape des cellules B matures (LB) en cellules sécrétrices d'anticorps (plasmocytes).

Les stades cellulaires et les profils d'expression des principaux régulateurs transcriptionnels de la différenciation B tardive sont indiqués. Les LB naïfs sont quiescents et expriment un ensemble unique de facteurs de transcription. Lors de l'exposition antigénique *in vivo* ou de l'activation avec soit des facteurs dérivés de LT (CD40L et cytokines) soit des ligands de TLR (*toll-like receptors*) *in vitro*, les LB naïfs entrent en cycle cellulaire et deviennent des cellules B activées. Ci-dessus, les cellules actives sont divisées en deux phases : une première phase, précoce, qui fait directement suite au signal mitogénique initial et une seconde phase, tardive, qui a lieu lors des cycles cellulaires suivants. Certains facteurs de transcription sont induits par activation (tels que Obf1 et Irf8) alors que d'autres sont induits graduellement et ont une activité dose-dépendante. Irf4 est exprimé en quantité intermédiaire dans les LB activés. Lors de l'initiation du programme de différenciation terminale, les LB activés diminuent l'expression des FTs indiqués associés aux LB et augmentent davantage l'expression d'Irf4 pour conduire la différenciation en cellules sécrétrices d'anticorps. Les cellules sécrétrices d'anticorps peuvent se diviser en plasmablastes cyclants de courte durée de vie et en plasmocytes post-mitotiques de longue durée de vie sur la base de l'expression de Blimp1. Figure issue de Stephen L. Nutt et al., 2011.

Figure 2.19 (à droite) : Les principaux régulateurs transcriptionnels supervisant les deux programmes mutuellement exclusifs de la différenciation B tardive.

Modèle du réseau transcriptionnel contrôlant la différenciation terminale des cellules B comme proposé par les équipes de Staudt (Shaffer et al., 2009) et Calame (Shapiro-Shelef & Calame, 2005). Ce modèle est centré sur la répression transcriptionnelle entre les facteurs promouvant les cellules B matures *id est* Pax5, Bcl6, Mitf et Bach2 (en bleu) et les facteurs promouvant la transition plasmocytaire (et la production de cellules sécrétrices d'anticorps) *id est* Xbp1, Blimp1 et Irf4 (en vert) qui assurent le caractère mutuellement exclusif des deux programmes d'expression génique. Les lignes à flèche indiquent une influence positive sur l'expression génique ou la fonction. Les lignes à barre indiquent une activité répressive. Figure issue de Stephen L. Nutt et al., 2011.

En parallèle, les régulateurs épigénétiques participent aussi à fixer la destinée cellulaire. Ezh2 – une unité catalytique du complexe répressif *polycomb 2* (PRC2, *polycomb-repressive complex 2*) qui réprime l'expression génique en marquant les histones par méthylation (H3K27me3) – et Smarca4 – alias BRG1, un régulateur transcriptionnel – collaborent pour réprimer l'expression inappropriée de gènes de lignages non-B (Su et al., 2003; Miyai et al., 2018).

Certains des facteurs épigénétiques jouent le même type de double rôle que les FTs lors de la lymphopoïèse B. Ezh2 et Suz12 – un autre composant de PRC2 – agissent comme activateurs transcriptionnels en phase précoce de la spécification alors qu'ils se comportent comme de puissants répresseurs aux stades tardifs de l'engagement lymphoïde B (Miyai et al., 2018). Spécifiquement dans les pro-LB, Ezh2 se charge de maintenir la répression d'un ensemble de gènes codant des facteurs de transcription associés aux lignées alternatives des lymphocytes T et des cellules lymphoïdes innées de type 2 (ILC2) incluant *Gata3*, *Zbtb16*, *Bcl11b* et *Rora*. C'est ainsi qu'Ebf1 pourrait recruter Ezh2 au niveau du promoteur de *Gata3* pour induire sa répression (Jacobsen et al., 2020). De plus, Ezh2 est spécifiquement requis dans les cellules B engagées pour réprimer *Lin28b* (codant une protéine de liaison à l'ARN associée aux programmes hématopoïétiques fœtaux) pour empêcher la conversion des cellules B2 adultes (dites conventionnelles) en cellules B1 fœtales (Jacobsen et al., 2020).

#### 2.2.4.2. Réseaux transcriptionnels de la différenciation terminale

La transition des lymphocytes B en cellules sécrétrices d'anticorps (plasmocytes) est essentielle pour la réponse immunitaire adaptative. La différenciation terminale en cellules hautement spécialisées est un processus irréversible, qui s'accompagne de la sortie du cycle cellulaire. La lymphopoïèse tardive repose sur un réseau de régulation génique orchestré par deux sous-ensembles de FTs. Ces derniers dirigent deux programmes transcriptionnels mutuellement exclusifs et antagonistes qui assurent au niveau moléculaire la ségrégation des destinées cellulaires – les LB matures versus les plasmocytes sécrétant des anticorps (Figure 2.18, Calame et al., 2003; Stephen L. Nutt et al., 2011, 2015).

Succinctement, Pax5, Bcl-6, Bach2 sont les principaux régulateurs transcriptionnels qui maintiennent l'identité cellulaire B et empêchent la différenciation plasmocytaire alors que Irf4, Blimp1 et Xbp1 sont les FTs pivots qui promeuvent le programme transcriptionnel des cellules sécrétrices d'anticorps (Nutt et al., 2011). Les signaux qui contrôlent la différenciation plasmocytaire doivent finalement éteindre le programme transcriptionnel des LB (Shaffer et al., 2002). Les deux groupes de FTs utilisent donc une boucle de répression transcriptionnelle – impliquant Pax5/Bach2 contre Blimp-1 – renforçant l'exclusivité de chaque transcriptome (Figure 2.19).



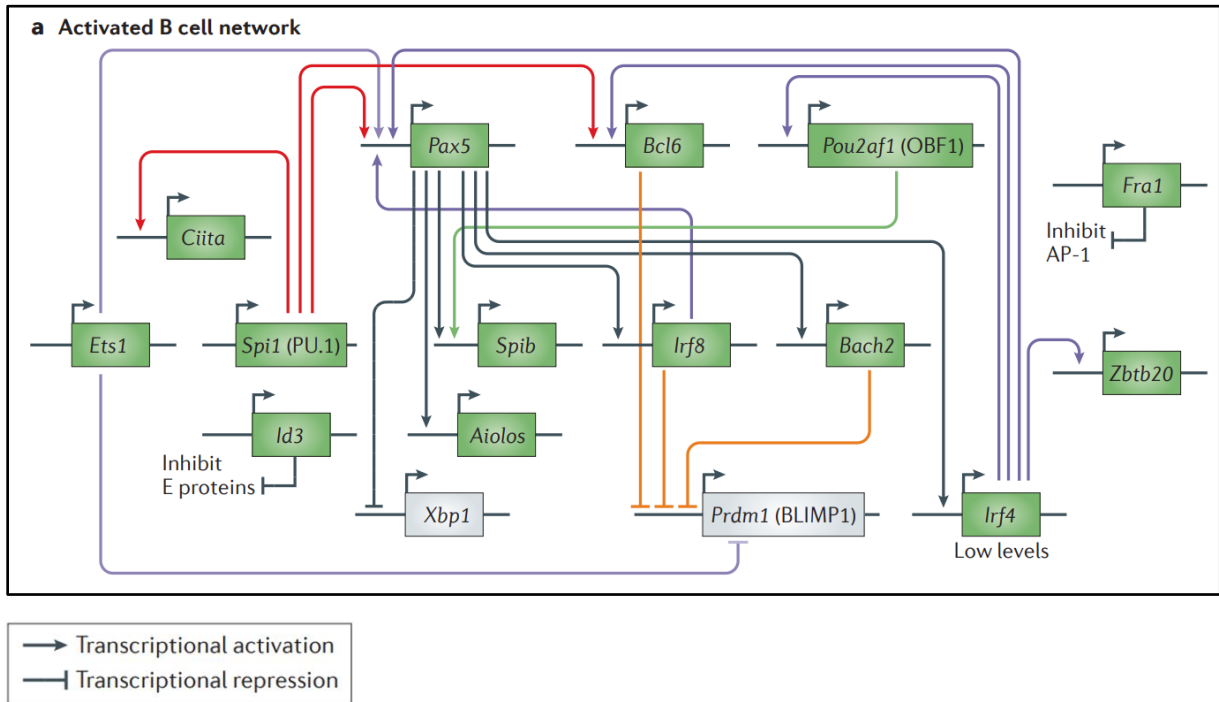


Figure 2.20 : Le réseau de régulation génique contrôlant la maturation LB activés.

Ce schéma modélise le réseau de régulation transcriptionnelle de la lymphopoïèse B tardive, l'activation et la maturation des lymphocytes B y compris la recombinaison de commutation de classe (CSR) et l'hypermutation somatique (SHM). Les lignes indiquent les fixations directes des protéines codées par les gènes indiqués sur les régions régulatrices des gènes cibles. Chaque activation transcriptionnelle est représentée par une ligne à flèche alors que chaque répression est représentée par une ligne à barre. Les flèches horizontales représentent les sites d'initiation de la transcription. Les gènes exprimés sont inscrits sur fond vert alors que les gènes très faiblement exprimés ou pas exprimés sont inscrits sur fond gris.

Abréviations dans l'ordre alphabétique : **AP-1**, activator protein 1 ; **Bach2**, BTB and CNC homologue 2 ; **Bcl6**, B cell lymphoma 6 ; **BLIMP1**, B lymphocyte-induced maturation protein 1 ; **Ciita**, MHC class II transactivator ; **Fra1**, FOS-related antigen 1 ; **Id3**, DNA-binding protein inhibitor ID3 ; **Irf**, interferon-regulatory factor ; **OBF1**, Oct binding factor 1 ; **Pax5**, paired box protein 5 ; **Pou2af1**, POU domain class 2 associating factor 1 ; **Prdm1**, PR domain zinc finger protein 1 ; **UPR**, unfolded protein response ; **Xbp1**, X box-binding protein 1 ; **Zbtb20**, zinc finger and BTB domain-containing protein 20. Figure issue de (Nutt et al., 2015).

Intervenant déjà au cours de la lymphopoïèse B précoce, les FTs Irf4, Pax5, Foxo1 et Ebf1 restent impliqués dans les étapes tardives du développement des cellules B. Dans les cellules B matures (naïves ou activées), Pax5 continue de réguler l'expression des composants du complexe BCR (IgM, CD79A) ainsi que celles d'autres récepteurs immunitaires tels que CD19 et CD21.

Outre *EBF1*, Pax5 active la transcription des gènes codant des FTs plus spécifiques en aval de la hiérarchie transcriptionnelle tels qu'Aiolos (*Ikzf3*), Irf4, Irf8, Bach2 (*BTB and CNC homologue 2*), Bcl-6, OBF1 CIITA et Spi-B qui sont associés au développement des cellules de la lignée lymphoïde B et notamment impliqués dans les réponses antigéniques, les réponses cytokiniques et les réponses à la fixation des ligands aux TLR, ou dans les mécanismes de CSR et SHM (Figures 2.18 et 2.20 ; Holmes et al., 2006; Pridans et al., 2008; Schebesta et al., 2007).

#### 1) Dans les LB naïfs

Dans un premier temps, Mitf (*microphthalmia-associated transcription factor*), répriment l'expression d'Irf4 dans les LB naïfs (L. Lin et al., 2004). Pax5 réprime directement *Cd28* (Delogu et al., 2006) et *Xbp1* (Reimold et al., 1996) codant le facteur de transcription XBP1 (*X-box-binding protein 1*) qui s'avèrera essentiel pour la formation de l'appareil sécrétoire des plasmocytes et nécessaire pour la production d'anticorps (Reimold et al., 2001).

De plus, les répresseurs transcriptionnels Bach2 et Bcl-6 coopèrent pour directement inhiber la transcription prématurée de *Prdm1* (Huang et al., 2014; Ochiai et al., 2006, 2008). Le gène *Prdm1* code le répresseur transcriptionnel Blimp-1 (*B lymphocyte-induced maturation protein-1*) qui remplit une fonction centrale dans la différenciation plasmocytaire (Kallies & Nutt, 2007). EBF1 participe également à la répression de Blimp-1 pour maintenir l'identité cellulaire B (Kikuchi et al., 2012). De surcroît, le double rétrocontrôle positif entre Pax5 et Bach2 entretient le maintien de l'identité cellulaire B et la prévention de la transition plasmocytaire (Figure 2.19).

#### 2) Dans les LB activés

Dans un second temps, suivant l'interaction antigénique, Irf4 et Pax5 sont requis pour activer l'expression d'*Aicda* (codant l'enzyme AID), agissant dans le mécanisme de recombinaison de commutation de classe des immunoglobulines (CSR des Ig) et permettant ainsi la maturation des LB activés (Dege & Hagman, 2013; Sciammas et al., 2006). Irf4 intervient ainsi dans le contrôle de la CSR et également dans la formation des centres germinatifs (Shaffer et al., 2009).

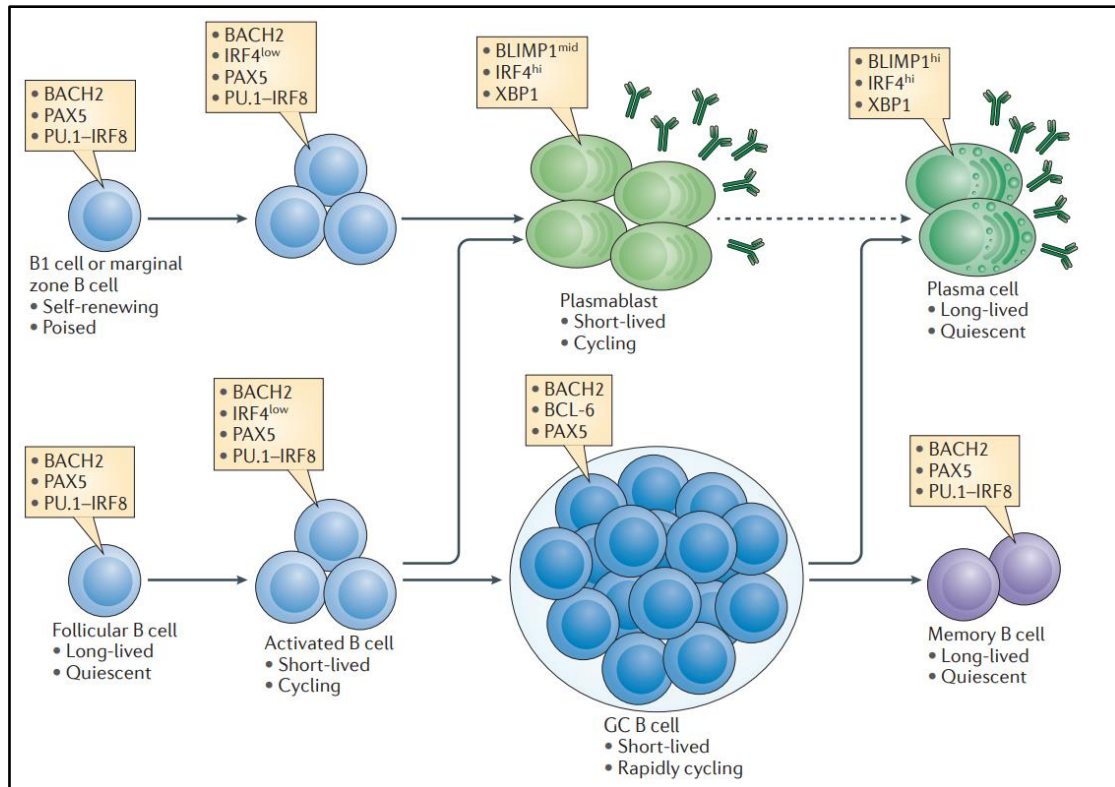


Figure 2.21 : Les stades cellulaires de la différenciation B tardive.

La majorité des cellules B matures sont localisées en follicules dans des organes lymphoïdes et sont alors appelées cellules B folliculaires. Les cellules B de la zone marginales et les cellules B de type B1 sont d'autres types de cellules B spécialisées. Les premières se localisent dans la région entre la pulpe rouge et la pulpe blanche de la rate. Les secondes sont détectées dans la cavité péritonéale et la cavité pleurale. Toutes les cellules B matures expriment PAX5, PU.1, IRF8 et BACH2 et la signalisation BCR induit de faibles (*low*) niveaux en IRF4. Les cellules B activées par une interaction antigénique (également appelé lymphoblastes B) peuvent rapidement proliférer, effectuer la recombinaison de commutation isotypique des immunoglobulines (CSR) et se différencier en plasmablastes de courte durée de vie (*short-lived*) qui expriment des niveaux d'expression élevés (*hi*, pour *high*) en IRF4 et XBP1 et intermédiaires (*mid*) en BLIMP1, et en anticorps sécrétés. Les cellules folliculaires B régulent positivement l'expression de BCL-6 et répriment l'expression d'IRF4 pendant la réaction du centre germinatif (GC) au sein duquel la maturation d'affinité du récepteur antigénique a lieu. Les cellules B comportant des récepteurs à forte affinité pour les antigènes migrent en dehors des centres germinatifs et se différencient soit en cellules B mémoires, qui expriment un programme transcriptionnel similaire à celui des cellules B matures, soit en plasmocytes de longue durée de vie (*long-lived*) qui expriment des niveaux d'expression élevés en IRF4, BLIMP1 et XBP1, et produisent une grande quantité d'anticorps. Les plasmocytes BLIMP1<sup>hi</sup> dérivent des cellules BLIMP1<sup>mid</sup>, mais l'origine des plasmocytes reste indéterminée, pouvant provenir des plasmablastes (ligne en pointillés), ou bien d'un stade plus précoce de plasmocyte engagé. La contribution des cellules B de la zone marginale et des cellules B de type B1 dans le compartiment des plasmocytes est peu caractérisée.

Abréviations dans l'ordre alphabétique : **BACH2**, *BTB and CNC homologue 2* ; **BCL-6**, *B cell lymphoma 6* ; **BLIMP1**, *B lymphocyte-induced maturation protein 1* ; **GC**, *germinal centre* ; **IRF4/8**, *interferon-regulatory factor 4/8* ; **PAX5**, *paired box protein 5* ; **XBP1**, *X-box-binding protein 1*. Figure issue de Stephen L. Nutt et al., 2015.

Lors de la stimulation antigénique, le niveau d'expression de Blimp-1 augmente graduellement tandis que le niveau de Bach2 diminue, mais la quantité restante de Bach2 permet d'une part de maintenir l'expression de Pax5 à un niveau relativement élevé. D'autre part, l'expression de Bach2 reste indispensable pour conserver une répression suffisante de *Prdm1* afin que les LB activés puissent subir la CSR (Muto et al., 2010), étant donné que Blimp-1 inhibe la CSR durant la différenciation plasmocytaire afin de limiter le stress génotoxique engendré par les mécanismes de CSR et d'hypermutation somatique (Shaffer et al., 2002).

Les cellules folliculaires présentent un niveau d'expression protéique nucléaire élevé en Bach2, prolongeant la répression de *Prdm1* (Blimp-1) et permettant de subir la CSR des Ig.

Inversement, les cellules B de la zone marginale et B1 ont un faible niveau d'expression nucléaire de Bach2, ce qui leur permet de rapidement se différencier en plasmocytes IgM<sup>+</sup> sans activer la CSR au préalable (Figure 2.21 ; Muto et al., 2010 ; K. Igarashi et al., 2014).

Les LB mémoires, quiescentes et de longue durée de vie, présentent également un niveau d'expression de Pax5 et Bach2 diminué par rapport aux LB matures naïfs, ce qui facilite les LB mémoires à rapidement se différencier en plasmocytes et donc rend plus efficace une seconde réponse immunitaire adaptative (Figure 2.21 ; Kometani et al., 2013).

### 3) Transition plasmocytaire

L'initiation de la différenciation plasmocytaire (en pré-plasmablastes) ne dépend pas de l'expression de Blimp-1 en soi mais serait due à l'inhibition de l'activité de Pax5, n'impliquant pas forcément la répression de son expression génique (Kallies et al., 2007; Nera et al., 2006). La stimulation antigénique enclenche la signalisation du BCR qui active notamment les kinases ERK1/2 et SYK. Celles-ci vont respectivement phosphoryler des résidus sérines (S189 et S283) et tyrosines (Y264 et Y299) de Pax5 (Inagaki et al., 2016; Yasuda et al., 2012). Chacune des phosphorylations atténue l'activité de Pax5, ce qui induit des variations d'expression graduelles inverses de Blimp-1 et Bach2 (évoquées précédemment). Ainsi ERK1/2 et SYK coopèrent pour inactiver Pax5, ce qui lève progressivement la répression de *Blimp-1*. Parallèlement, Irf4 joue sur les deux tableaux (*dualité de fonction*) en promouvant la maturation des LB activés ou la différenciation terminale en plasmocytes en fonction de son niveau d'expression protéique (Sciammas et al., 2006 ; Ochiai et al., 2013). IRF4 forme des hétérodimères (avec BATF ou PU.1) ou des homodimères en fonction de sa concentration (faible ou élevée respectivement) et ainsi régulent des gènes cibles différents durant la différenciation des cellules naïves en plasmocytes (K. Igarashi et al., 2014).

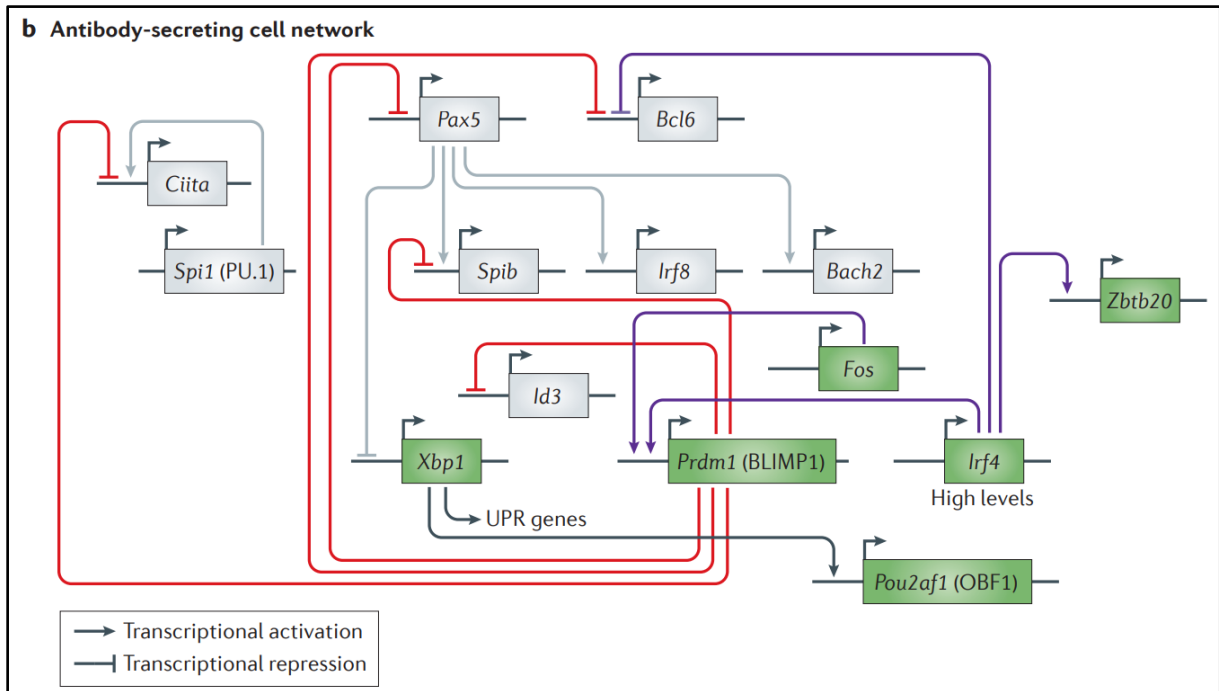


Figure 2.22 : Le réseau de régulation génique contrôlant la différenciation terminale des LB en cellules sécrétant des anticorps (plasmablastes et plasmocytes).

Les lignées indiquent les fixations directes des protéines codées par les gènes indiqués sur les régions régulatrices des gènes cibles. Chaque activation transcriptionnelle est représentée par une ligne à flèche alors que chaque répression est représentée par une ligne à barre. Les flèches horizontales représentent les sites d'initiation de la transcription. Les gènes exprimés sont inscrits sur fond vert alors que les gènes très faiblement exprimés ou pas exprimés sont inscrits sur fond gris.

Abréviations dans l'ordre alphabétique : **AP-1**, activator protein 1 ; **Bach2**, BTB and CNC homologue 2 ; **Bcl6**, B cell lymphoma 6 ; **BLIMP1**, B lymphocyte-induced maturation protein 1 ; **Ciita**, MHC class II transactivator ; **Fra1**, FOS-related antigen 1 ; **Id3**, DNA-binding protein inhibitor ID3 ; **Irf**, interferon-regulatory factor ; **OBF1**, Oct binding factor 1 ; **Pax5**, paired box protein 5 ; **Pou2af1**, POU domain class 2 associating factor 1 ; **Prdm1**, PR domain zinc finger protein 1 ; **UPR**, unfolded protein response ; **Xbp1**, X box-binding protein 1 ; **Zbtb20**, zinc finger and BTB domain-containing protein 20. Figure issue de (Nutt et al., 2015).

À faible concentration, *Irf4* induit l'expression d'AID impliquée dans la maturation des LB activés (Figure 2.20) tandis qu'à forte concentration, *Irf4* peut à la fois réprimer *Bcl-6* (Saito et al., 2007) et activer l'expression de *Prdm1* codant Blimp-1, ce qui induit la transition plasmocytaire (Figure 2.22, Sciammas et al., 2006; Klein et al., 2006; Ochiai et al., 2013).

*Irf4* entre en compétition avec d'autres FTs pour déterminer la destinée cellulaire. *Irf8* et PU.1 jouent un rôle dans l'inhibition de la différenciation plasmocytaire à travers une compétition avec *Irf4* (Carotta et al., 2014). Le locus de *Prdm1* comporte des sites de liaison pour *Irf4* ainsi que pour *Bach2* et *Bcl-6*, signifiant une compétition entre la répression exercée par le duo *Bach2/Bcl-6* et l'activation croissante d'*Irf4* (K. Igarashi et al., 2014).

Pour mener stablement à la différenciation terminale, *Irf4* et Blimp-1 forment une boucle de régulation positive ; à laquelle s'ajoute une boucle de régulation négative : l'expression croissante de Blimp-1, ayant pris l'ascendant, impose la diminution progressive du niveau d'expression d'*Irf8*, *Aicda*, *Pax5*, *Bcl-6* et *Bach2* jusqu'à leur extinction (Delogu et al., 2006; Kometani & Kurosaki, 2015; K.-I. Lin et al., 2002).

Blimp-1 et *Irf4* collaborent pour contrôler la différenciation plasmocytaire, combinant la répression du programme transcriptionnel de l'identité B caractérisée par l'extinction (par Blimp-1) des gènes codant *Spi-B*, *Id3*, *CIITA*, *Myc* (Figure 2.22 ; Stephen L. Nutt et al., 2011, 2015) et l'activation de gènes associés aux cellules sécrétrices d'anticorps tels que *IgJ*, et *Xbp1* activés par Blimp-1 et requis pour la fonction sécrétrice des plasmocytes (Shapiro-Shelef & Calame, 2005). De surcroît, *Cd28*, dont la répression est levée par l'inactivation de *Pax5* (Delogu et al., 2006), régule positivement *Irf4* et Blimp-1 et favorise la survie cellulaire des plasmocytes 'de longue durée de vie' (Rozanski et al., 2015; Utley et al., 2020).

La régulation épigénétique complète le réseau de FTs de la différenciation B tardive. La méthyltransférase d'histone EZH2, intervient dans le cadre de la répression des destinées cellulaires alternatives en déposant des marques épigénétiques répressives au cours des réactions contribuant à la formation et la prolifération des centres germinatifs (GCs). *BCL-6* et Blimp-1 recrutent différenciellement EZH2 au niveau de loci distincts pour respectivement réprimer des gènes associés au type plasmocytaire ou au type lymphocytaire B (Scharer et al., 2018).



## Chapitre 3 de l'introduction

Selon l'Institut National du Cancer (INCa), le cancer représente la première cause de mortalité prématurée en France métropolitaine, déplorant 150'000 décès par an. L'incidence des cancers est de 382'000 nouveaux cas par an. En France, près de 2500 diagnostics de cancers pédiatriques sont établis chaque année, dont la moitié survenant avant l'âge de 5 ans, parfois dès la naissance. Approximativement 1 700 enfants ont moins de 15 ans et 800 patients sont des adolescents (15-18 ans). Chaque année, 500 enfants en décèdent et plus de 50% des enfants qui y survivent vivent avec de graves séquelles ou souffrent de l'apparition de pathologies chroniques.

Certains types de cancers sont spécifiques à l'enfant et la plupart des cancers de l'adulte n'affectent pas les enfants. Une plus longue période d'exposition aux facteurs de risque socio-environnementaux et l'accumulation nécessaire de mutations somatiques adéquates acquises dans un ordre précis pour mener à un cancer peuvent expliquer cette divergence (Tomasetti & Vogelstein, 2015; Vogelstein & Kinzler, 1993, 2015). La nature des cancers pédiatriques varie également avec l'âge (nourrisson, enfant, adolescent, voire jeune adulte). En Europe et aux États-Unis, les cancers pédiatriques les plus courants sont les leucémies, les tumeurs cérébrales et les lymphomes. Il est intéressant de remarquer que les cancers de l'enfant les plus fréquents (Tableau 3 ci-dessous, source INCa) correspondent aux cancers résultant notamment d'altérations génétiques des gènes de développement *PAX*.

Type de cancer	Organe ou tissu d'origine
leucémie	moelle osseuse et sang
lymphome	ganglions et système lymphatique
néphroblastome	reins
neuroblastome	glandes surrénales et système nerveux sympathique
ostéosarcome, tumeur d'Ewing	os
rhabdomyosarcomes et autres sarcomes des parties molles	muscles et tissus de soutien
rétinoblastome	rétine
tumeur cérébrale	cerveau et moelle épinière
autres tumeurs	ovaires, testicules, système nerveux central, thorax, abdomen

Tableau 3 : Les principaux cancers pédiatriques. Source : site web de l'INCa.



<b>Sous-familles</b>	<b>Gènes PAX</b>	<b>Domaines d'expression (liste non-exhaustive)</b>	<b>Pathologies associées hormis cancers (liste non-exhaustive)</b>	<b>Cancers associés (liste non-exhaustive)</b>
<b>Groupe I</b>	<b>PAX1</b>	Sclérotome, squelette, thymus	Syndrome de Klippel-Feil ; syndrome de Jarcho-Levin ;	Tumeurs des glandes salivaires
	<b>PAX9</b>	Sclérotome, squelette, dents, thymus	Oligodontie	Ostéosarcomes et sarcomes d'Ewing (cancers des os), LAL-T
<b>Groupe II</b>	<b>PAX2</b>	SNC, yeux, oreilles, reins, glandes mammaires	Syndrome rein-colobome	néoplasmes rénaux dont renal cell carcinoma (RCC) ; Wilms tumor (néphroblastome)
	<b>PAX5</b>	SNC, testicules, Lignée lymphoïde B		Médulloblastomes, neuroblastomes de type N, carcinomes neuroendocriniens, cancer de la prostate, hémopathies lymphoïdes (lymphomes à cellules B, LAL-B, LAM)
	<b>PAX8</b>	SNC, thyroïde, reins	Hypothyroïdie congénitale causée par dysplasie thyroïdienne	Cancer de thyroïde ; néoplasmes rénaux dont carcinomes rénaux (RCC)
<b>Groupe III</b>	<b>PAX3</b>	Crête neurale, SNC, somites/muscles	Syndrome de Waardenburg	Rhabdomyosarcomes alvéolaires (ARMS), mélanome, sarcome d'Ewing (cancer des os) sarcomes nasosinusiens biphénotypiques
	<b>PAX7</b>	Crête neurale, SNC, somites/muscles	Myopathies	Rhabdomyosarcomes alvéolaires (ARMS), mélanome, sarcome d'Ewing (cancer des os) carcinome pulmonaire épidermoïde
<b>Groupe IV</b>	<b>PAX4</b>	Pancréas, intestins, glande pinéale, SNC	Diabète de types I et II	Insulinomes (nésidioblastomes)
	<b>PAX6</b>	Pancréas, intestins, SNC, glande pinéale, yeux, nez	Diabète ; aniridie, cataracte	Glioblastomes, rétinoblastomes

Tableau 4. Pathologies associées aux altérations génétiques et dérégulation d'expression des gènes PAX.

Données bibliographiques extraites de la littérature citée dans le texte et/ou indiquées dans les revues de (Chi & Epstein, 2002), et (Mahajan et al., 2015).

### 3.1. Implication des protéines PAX dans la pathogénèse

Étant données les fonctions primordiales des protéines PAX à de multiples étapes de l'embryogenèse, dans les différenciations tissulaires et l'organogenèse (cf. section 1.2), les mutations perte-de-fonction des gènes *PAX* conduisent à **diverses anomalies du développement**, dont le mécanisme pathogénique repose essentiellement sur **l'haploinsuffisance**.

L'importance des gènes *Pax* dans le développement embryonnaire a ainsi été mise en évidence par divers phénotypes résultant de l'inactivation des gènes *Pax* chez la drosophile et la souris (Tremblay & Gruss, 1994). Par exemple chez la souris, les mutations hétérozygotes des gènes *Pax1*, *Pax3*, ou *Pax6* mènent respectivement aux phénotypes "ondulé" (*undulated phenotype* ; Balling et al., 1988), "tacheté" (*spotch phenotype* ; D. J. Epstein et al., 1991) ou "de petit œil" (*small eye (Sey) phenotype* ; Hill et al., 1991). Plus drastiquement, les embryons exprimant une mutation homozygote de *Pax6* (*Sey/Sey*) ne développent ni d'œil ni de cavités nasales, empêchant les nouveau-nés de respirer pendant la succion, précipitant la mort des souriceaux (Hill et al., 1991). Similairement, l'inactivation génique de *Pax2* et *Pax8* se manifeste notamment par une absence de reins (Bouchard et al., 2002; Favor et al., 1996) tandis que l'inactivation homozygote de *Pax5* entraîne une létalité périnatale due à un défaut majeur de développement du SNC (Urbánek et al., 1994).

De plus, les gènes *PAX* sont fréquemment exprimés dans les cancers (Muratovska et al., 2003). La dérégulation d'expression et/ou d'activité des gènes *PAX* est **associée à l'oncogenèse de divers types de tumeurs malignes** (Chi & Epstein, 2002; Mahajan et al., 2015; Robson et al., 2006) – en particulier des cancers pédiatriques – qui découlent de leurs territoires d'expression physiologique respectifs (Tableau 4). Bien que tous les gènes *PAX* semblent pouvoir contribuer à l'oncogenèse, les gènes des groupes II (*PAX2/PAX5/PAX8*) et III (*PAX3/PAX7*) jouent des rôles particulièrement importants dans l'initiation et la progression tumorales alors que les gènes *PAX* des groupes I et IV sont moins souvent impliqués dans l'oncogenèse (Chi & Epstein, 2002; Mahajan et al., 2015; Robson et al., 2006). Le degré de sévérité tumorale peut être corrélé au niveau de complexité de structure protéique : les altérations des protéines *PAX* des groupes II et III (qui contiennent le motif octapeptide ET un homéodomaine partiel ou complet) sont souvent associées à des cancers plus agressifs, alors que les altérations des protéines *PAX* des groupes I et IV (qui contiennent le motif octapeptide OU un homéodomaine complet) sont plutôt associées à un pronostic favorable (Robson et al., 2006). Ainsi, le potentiel oncogénique des gènes *PAX* est le reflet de leurs propriétés cellulaires physiologiques. Pour rappel, les facteurs de transcription *PAX* sont largement impliqués dans la régulation de la différenciation cellulaire, la prolifération, la migration et la survie de différents types cellulaires.

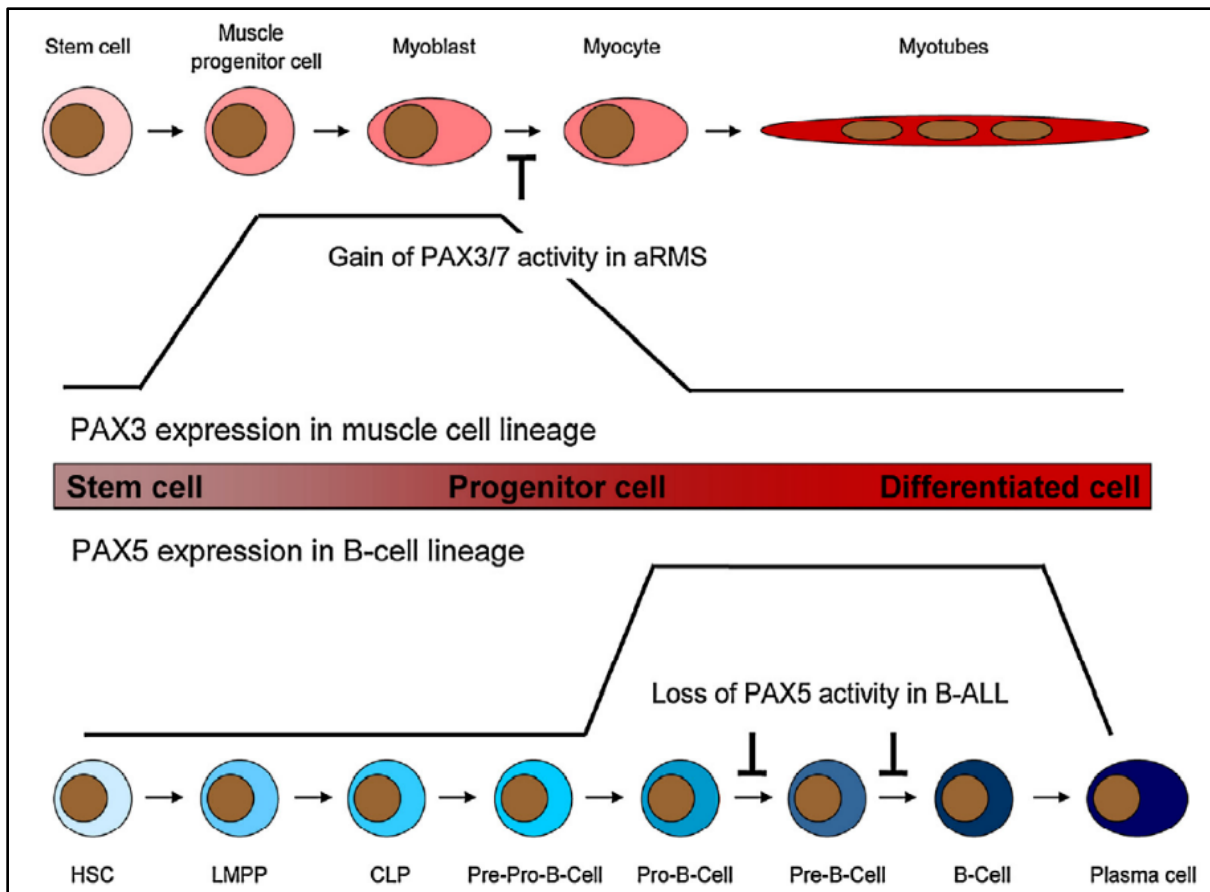


Figure 3.01 : Exemples représentatifs des mécanismes gain-de-fonction et perte-de-fonction conduisant à l'expression et/ou l'activité aberrante de protéines PAX, aboutissant à un blocage de différenciation combiné au maintien des cellules progénitrices mutées dans un état prolifératif. Figure issue de (Wachtel & Schäfer, 2015).

Les cellules cancéreuses peuvent tirer profit de ces puissantes propriétés physiologiques et les détourner en faveur de la progression tumorale. En particulier, les cinq gènes *PAX* des groupes II (*PAX2/PAX5/PAX8*) et III (*PAX3/PAX7*) confèrent des propriétés de motilité cellulaire, de survie cellulaire et d'auto-suffisance en signaux de croissance cellulaire, qui sont des caractéristiques favorisant la progression tumorale (Robson et al., 2006).

La dérégulation de l'expression ou de l'activité des protéines *PAX* peut être causée soit par des **anomalies quantitatives (variation du nombre de copies, CNV)**, soit par des **altérations de leur structure ou séquence génique** (Mahajan et al., 2015; Robson et al., 2006; Wachtel & Schäfer, 2015).

- Anomalies quantitatives

L'expression endogène des gènes *PAX* se trouve souvent requise pour la prolifération et la survie des cellules cancéreuses (Gibson et al., 2007; Muratovska et al., 2003). L'expression inappropriée de protéines *PAX* sous forme sauvage constitue un mécanisme récurrent de l'oncogenèse. Ainsi, la surexpression de protéines *PAX* (sous forme sauvage) peut davantage promouvoir la progression tumorale. Par ailleurs, l'expression ectopique de protéines *PAX* peut également jouer un rôle oncogénique en dérégulant les signaux de prolifération et de survie cellulaire. La réexpression aberrante des protéines *PAX* est également fréquemment observée dans les cancers, permettant de réattribuer des fonctions embryonnaires ou développementales (Morgenstern et al., 2010). La persistance anormale de l'expression de *PAX* dans des tissus partiellement différenciés peut provoquer un blocage de la différenciation (en empêchant la différenciation terminale) et/ou participer à la formation d'hyperplasie.

*A contrario*, la perte de fonction d'une protéine *PAX* peut bénéficier à la cellule tumorale : étant donné que les protéines *PAX* sont indispensables au processus de différenciation tissulaire, la réduction ou la perte d'expression de gènes *PAX* mènent à un blocage de la différenciation à un stade cellulaire précoce, caractéristique de nombreuses tumeurs.

- Altérations génétiques

L'activité aberrante des protéines *PAX*, résultant d'altérations de structure génique et notamment des translocations caractéristiques, peut être attribuée à divers mécanismes gain-de-fonction ou perte-de-fonction (Wachtel & Schäfer, 2015 pour revue). Ces aberrations génétiques perturbent généralement l'activité transcriptionnelle des facteurs de transcription *PAX* de sorte à générer un état cellulaire permissif pro-tumoral favorisant la survie, la prolifération et la migration des cellules tumorales au détriment de la différenciation cellulaire.

En définitive, l'activité aberrante des protéines *PAX* entraînent typiquement un blocage de la différenciation maintenant les cellules progénitrices mutées dans un état très prolifératif, permettant leur expansion exponentielle (Wachtel & Schäfer, 2015 pour revue ; Figure 3.01).



### 3.1.1. Groupe I : PAX1 & PAX9

Le défaut d'expression en *PAX1* engendre des malformations vertébrales congénitales, telles que le syndrome de Klippel-Feil et le syndrome de Jarcho-Levin. La surexpression de *PAX1* est associée à des tumeurs des glandes salivaires (Francioso et al., 2002). Une mutation perte-de-fonction de *PAX9* (causant un décalage du cadre de lecture à partir de la phase codant le domaine PAIRED) peut provoquer une oligodontie, définie par une agénésie dentaire d'au moins six dents (Stockton et al., 2000). L'expression de *PAX9* se trouve significativement réduite dans la majorité des carcinomes invasifs et dysplasies épithéliales précancéreuses de l'œsophage (Gerber et al., 2002) alors que *PAX9* est surexprimé dans des tumeurs malignes pédiatriques d'origine osseuse (ostéosarcomes et sarcomes d'Ewing) et les LAL-T pédiatriques (Mahajan et al., 2015).

### 3.1.2. Groupe II : PAX2, PAX5 & PAX8

Les altérations affectant *PAX2*, *PAX5* et *PAX8* sont davantage des mutations dites perte-de-fonction.

- L'haploinsuffisance de *PAX2* cause le Syndrome Rein-Colobome (SRC) caractérisé par une hypodysplasie rénale et une dysplasie du nerf optique (Eccles, 1998; Favor et al., 1996).
- Les mutations hétérozygotes de *PAX8* provoquent une dysgénésie thyroïdienne, une forme d'hypothyroïdie congénitale (Macchia et al., 1998). Des mutations faux-sens affectant le domaine PAIRED de *PAX8* ont notamment été identifiées et ont pour conséquence de réduire voire supprimer son habilité de liaison à l'ADN.
- Les délétions de *PAX5*, considéré comme un gène suppresseur de tumeur, sont très fréquemment détectées dans les leucémies lymphoblastiques B (Mullighan et al., 2007).

Par ailleurs, l'expression aberrante de *PAX2* ou *PAX8* est associée à divers cancers du rein et notamment à la formation de tumeurs de Wilms (néphroblastome) représentant 5 à 14% des cancers de l'enfant (Ozcan et al., 2012). La surexpression de *PAX2* favoriserait notamment la progression tumorale des carcinomes à cellules rénales (CCR) qui est le type le plus courant de cancer du rein (Gnarra & Dressler, 1995; Winyard et al., 1996).

*PAX8* joue également un rôle dans des cas de cancers de la thyroïde (Mahajan et al., 2015 pour revue). *PAX8* se situe plus fréquemment dans le cytoplasme dans les cancers de la thyroïde réduisant le champ d'activité transcriptionnelle de *PAX8* et conduisant ainsi à la diminution de la transcription des gènes spécifiques de la thyroïde (Scouten et al., 2004).



### 3.1.3. Implication de PAX5 dans l'oncogénèse

Corrélée à son rôle clé dans la différenciation et la prolifération des cellules de la lignée lymphoïde B (Nutt et al., 1999; Wakatsuki et al., 1994), la dérégulation de l'expression de *PAX5* est fortement associée à des hémopathies malignes, principalement du tissu lymphoïde B. Néanmoins, *PAX5* est impliqué dans la pathogénèse de diverses tumeurs malignes non-hématopoïétiques et joue un rôle oncogénique notoire dans les tumeurs solides malignes non-hématopoïétiques, soulignant son potentiel oncogénique (ne se résumant pas à la perte de son rôle de suppresseur de tumeur).

- **Hémopathies malignes**

*PAX5* est ciblé par diverses altérations génétiques somatiques dans plus d'un tiers des leucémies aiguës lymphoblastique de la lignée B (Coyaud, Struski, Prade, et al., 2010; Familiades et al., 2009; Mullighan et al., 2007) et dans plusieurs types de lymphomes à cellules B (Poppe et al., 2005). *Ces sujets sont ultérieurement développés dans des sections spécifiques.*

Par ailleurs, l'expression aberrante de *PAX5* est associée aux leucémies aiguës myéloblastiques (LAM) qui présentent une translocation t(8;21)(q22;q22) produisant l'oncoprotéine de fusion RUNX1-ETO (Tiacci et al., 2004). Les LAM t(8 ;21) font partie des leucémies aiguës biphénotypiques (MPAL, *mixed-phenotype acute leukemia*) et se caractérisent par l'expression simultanée de gènes spécifiques des lignages myéloïde et lymphoïde (Weinberg & Arber, 2010). L'expression inappropriée de *PAX5* est en partie responsable de l'activation aberrante de gènes spécifiques de la lignée B dont CD19 (Walter et al., 2010).

- **Tumeurs malignes non-hématopoïétiques**

*PAX5* est exprimé de façon ectopique dans les médulloblastomes (affectant le cervelet), qui constituent la tumeur cérébrale maligne la plus fréquente de l'enfant. Notamment dans les médulloblastomes desmoplastiques, *PAX5* s'exprime dans une zone tumorale proliférant et contenant des cellules neuronales indifférenciées (Kozmik et al., 1995).

L'expression de *PAX5* est également dérégulée dans les neuroblastomes (NB), qui constituent la tumeur maligne solide extra-crânienne la plus fréquente de l'enfant. Les neuroblastomes sont composés de types cellulaires malins distincts, incluant des cellules neuroblastiques et des cellules stromales (Brodeur, 2003). *PAX5* est exprimé dans des lignées cellulaires dérivées de la sous-population de phénotype agressif (type N) alors que son expression n'est pas détectée dans les cellules de phénotype bénin (sous-population de type S, cellules adhérentes au substrat) (Baumann Kubetzko et al., 2004). Les neuroblastomes sont des tumeurs dérivées des cellules de la crête neurale, à l'instar du cancer bronchique à petites cellules (CBPC ; *small-cell lung carcinoma, SCLC*) qui est une tumeur agressive.





L'expression ectopique de *PAX5* est aussi détectée dans des lignées cellulaires de CBPC (Baumann Kubetzko et al., 2004). Fonctionnellement, l'expression aberrante de *PAX5* semble conférer des propriétés de malignité aux cellules tumorales de NB et CBPC, incluant un taux de prolifération accru et/ou une croissance cellulaire indépendante de l'adhésion au stroma (Baumann Kubetzko et al., 2004).

*PAX5* est fréquemment exprimé dans les carcinomes neuroendocrines peu différenciés – tels que le carcinome à cellules de Merkel qui est un cancer agressif de la peau, et les carcinomes neuroendocrines bronchiques à petites cellules (déjà susmentionnés), de haut grade – mais jamais dans les tumeurs carcinoïdes, des tumeurs neuroendocrines différenciées (Dong et al., 2005; Kanteti et al., 2009; Kolhe et al., 2013; Torlakovic et al., 2006).

Par ailleurs, *PAX5* se trouve impliqué dans le cancer de la prostate de mauvais pronostic, étant également exprimé dans les tissus prostatiques sains des patients. *PAX5* est capable d'activer la transcription de *FOXP4* (*forkhead box P4*) et *FOXP4-AS1* (*FOXP4 antisense RNA 1*), un long ARN non-codant remplissant un rôle « d'éponge à miR ». La surexpression de *PAX5* régule positivement l'axe oncogénique *FOXP4/FOXP4-AS1* et ainsi concoure à réduire l'apoptose et à promouvoir la prolifération cellulaire (Wu et al., 2019).

#### 3.1.4. Groupe III : *PAX3* & *PAX7*

Des mutations de *PAX3* ont été identifiées chez les patients atteints du Syndrome de Waardenburg (de type I ou III), caractérisé par des anomalies de pigmentation (des yeux, des cheveux et de la peau) et par différents degrés de surdité (Tassabehji et al., 1992). Certains de ces patients présentent des troubles musculaires (Hoth et al., 1993).

Des variants bi-alléliques de *PAX7* ont été identifiés dans des cas de myopathies familiales de degrés modéré à sévère, caractérisées par une hypotonie, une atrophie musculaire, des scolioses et des traits du visage légèrement dysmorphiques. *PAX7* étant normalement responsable du maintien du réservoir de cellules souches musculaires (les cellules satellites), le drastique défaut d'expression de *PAX7* chez ces patients conduit spécifiquement à l'épuisement de ce réservoir de cellules satellites (Feichtinger et al., 2019).

Par ailleurs, *PAX3* joue un rôle anti-apoptotique dans les cellules tumorales des mélanomes (Scholl et al., 2001).

#### 3.1.5. Groupe IV : *PAX4* & *PAX6*

Parmi les mutations perte-de-fonction de *PAX6* ont été identifiées des mutations non-sens tronquant *PAX6* au sein du domaine PAIRED (N-terminal) ou du domaine PST (C-terminal). Ce dernier mutant conserve une activité transcriptionnelle partielle. L'hétérozygotie en *PAX6* provoque de sévères défauts crânio-faciaux et/ou du SNC (Glaser et al., 1994).

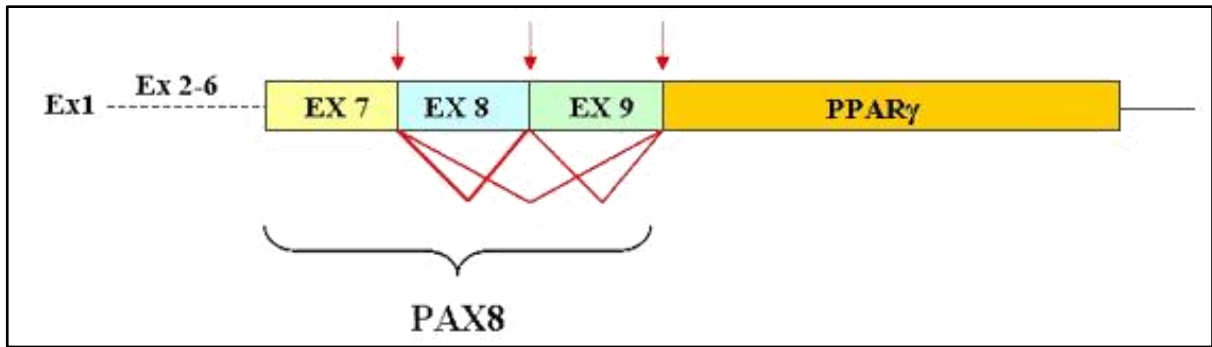


Figure 3.02 : Schéma des structures des chimères de fusion PAX8-PPAR $\gamma$ , caractéristiques des tumeurs de la thyroïde. Figure de Atlas Genetics (de, Biase D ; Morandi, L ; Tallini, G PAX8 (paired box 8) Atlas Genet Cytogenet Oncol Haematol. 2010;14(4):412-415.)

Étant donné le rôle prépondérant de *PAX6* dans le développement des yeux, l'haploinsuffisance en *PAX6* résultant de mutations perte-de-fonction hétérozygotes de *PAX6* peut causer une cécité, des cataractes et des malformations congénitales oculaires telles qu'une aniridie et des colobomes oculaires (Glaser et al., 1994; Jordan et al., 1992). D'ailleurs, *PAX6* pourrait être impliqué dans la prolifération et la survie des cellules tumorales de rétinoblastome (Bai et al., 2011). En revanche, *PAX6* exerce un rôle de suppresseur de tumeur en freinant la croissance et les propriétés invasives des cellules tumorales de glioblastomes (Mayes et al., 2006; Zhou et al., 2005).

Pour rappel, les protéines *PAX4* et *PAX6* sont toutes deux impliquées dans la régulation des cellules endocrines du pancréas. Des variants de *PAX4* ont été associés aux diabètes de type 1 et 2, notamment des variants hypomorphes résultant de mutations faux-sens du domaine PAIRED de *PAX4* ont été identifiés chez des patients diabétiques (Lorenzo et al., 2017; Shimajiri et al., 2001). *PAX6* est exprimé dans les adénocarcinomes pancréatiques (Lang et al., 2008) tandis que *PAX4* est fortement exprimé dans les insulinomes. Il s'agit notamment d'une forme tronquée de la protéine *PAX4* capable de se lier à l'ADN aussi efficacement que la forme sauvage mais va préférentiellement activer les gènes cibles plutôt que réprimer leur transcription (Miyamoto et al., 2001), ce qui favoriserait la survie des cellules tumorales (Brun et al., 2007).

### 3.1.6. Translocations chromosomiques des PAX associées aux cancers

Les gènes *PAX*, et plus particulièrement les gènes des groupes II et III, sont fréquemment ciblés par des réarrangements chromosomiques, une caractéristique de certaines tumeurs malignes. *PAX3*, *PAX5*, *PAX7* et *PAX8* sont notamment impliqués dans des translocations chromosomiques récurrentes considérées comme des événements *drivers* de l'oncogenèse (Barr, 1997; Wachtel & Schäfer, 2015) et se traduisent par divers mécanismes.

- L'oncogène de fusion *PAX8-PPAR $\gamma$* , résultant de la translocation récurrente t(2;3)(q13;p25), est exprimé dans plusieurs cancers de la thyroïde (Castro et al., 2006; Kroll et al., 2000; Marques et al., 2002). L'oncoprotéine juxtapose le domaine de liaison à l'ADN de *PAX8* à la quasi-totalité du récepteur (nucléaire) gamma activé par les « proliférateurs » de peroxyosomes (*PPAR- $\gamma$*  ou *PPAR $\gamma$* , *peroxisome proliferator-activated receptor gamma* ; Figure 3.02). L'oncoprotéine exerce un effet dominant-négatif sur la protéine normale *PPAR $\gamma$*  qui tient normalement un rôle de suppresseur de tumeur, alors que *PAX8-PPAR $\gamma$*  conserve la capacité à transactiver les promoteurs activés par *PAX8* (Gregory Powell et al., 2004; Placzkowski et al., 2008).

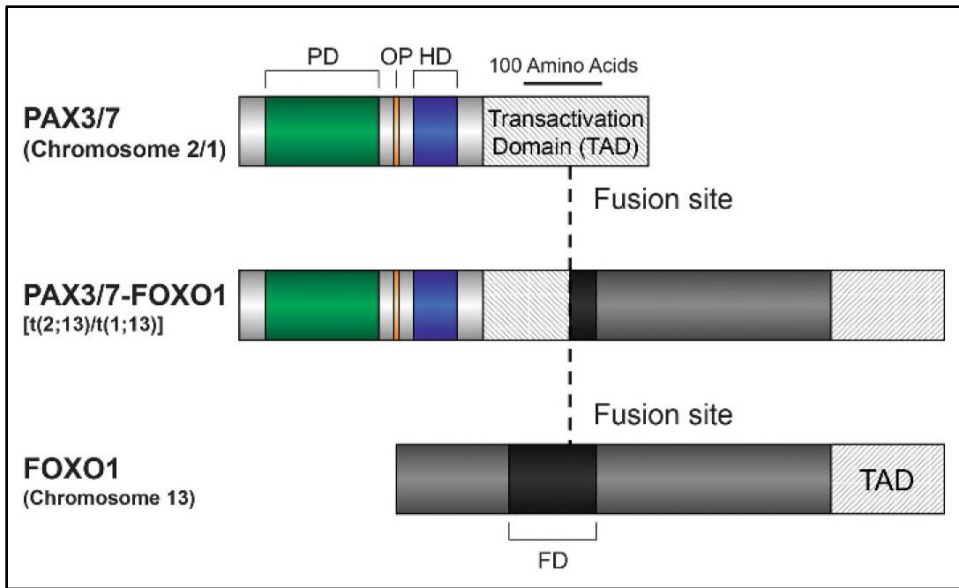


Figure 3.03 : Schéma de la structure des oncogènes de fusion PAX-FOXO1 (PAX-FKHR), caractéristique des RMS-A.

Les translocations *PAX-FOXO1* impliquent FOXO1 (chromosome 13) et PAX3 (chromosome 2) ou PAX7 (chromosome 1). Elles préservent intacts le domaine de liaison à l'ADN (PAIRED, PD), le motif octapeptide (OP) et l'homéodomaine complet (HD) des protéines PAX et le domaine de transactivation (TAD) du facteur de transcription FOXO1 (FKHR). FD, forkhead domain. La région N-terminale de PAX3 et PAX7 contient le domaine de cis-répression chevauchant le domaine PD. Les points de cassures ne tronquant que partiellement les TAD de PAX3 et PAX7. Figure issue de (Mahajan et al., 2015).

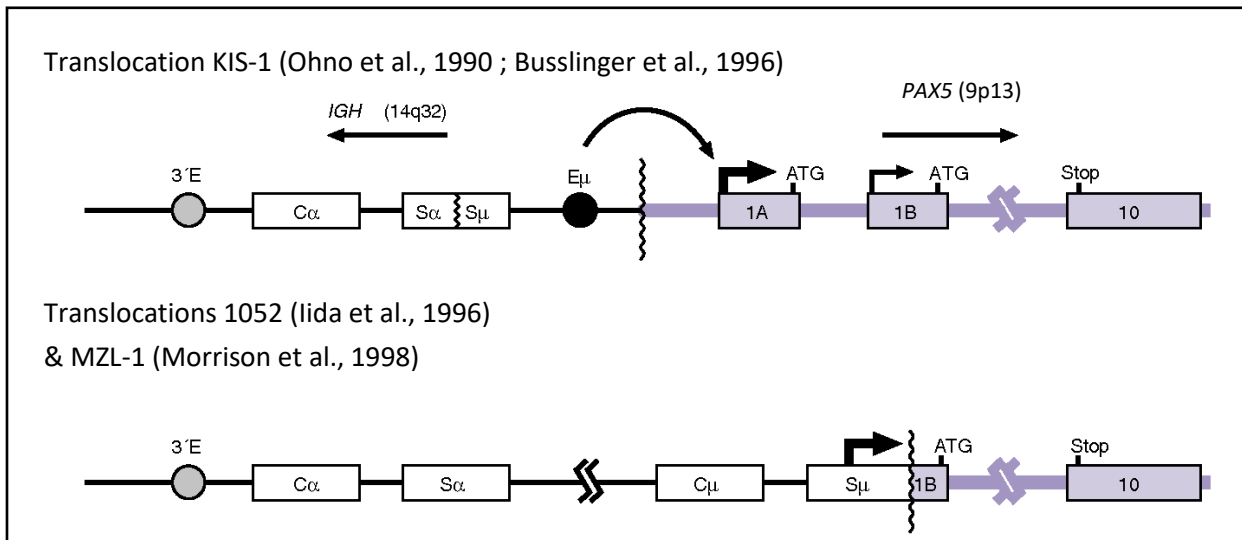


Figure 3.04 : Structures et régulation de la translocation *IGH-PAX5* associée aux lymphomes.

La structure du gène *PAX5* reste intact mais sa transcription est dérégulée de deux manières différentes. Figure issue de (Cobaleda, Schebesta, et al., 2007).

- Les translocations chromosomiques  $t(2;13)$  *PAX3-FKHR* et  $t(1;13)$  *PAX7-FKHR* caractérisent les rhabdomyosarcomes alvéolaires (RMS-A), un cancer pédiatrique du muscle squelettique (tissu mou). Elles génèrent des chimères de fusion préservant les domaines de la région N-terminale des PAX (PAIRED, OP et HD complet) et le domaine de transactivation transcriptionnelle C-terminal du facteur de transcription FKHR (forkhead in rhabdomyosarcoma, plus communément appelé FOXO1, forkhead box protein O1 ; Figure 3.03). Les oncoprotéines *PAX3-FKHR* et *PAX7-FKHR* utilisent un mécanisme gain-de-fonction similaires dans les RMS-A (Bennicelli et al., 1996, 1999). Elles fonctionnent comme de puissants activateurs transcriptionnels, avec des propriétés de transactivation exacerbées. Le TAD de FKHR résiste à l'inhibition des domaines de cis-répression de *PAX3* et *PAX7* chevauchant leur domaine PAIRED respectif (Bennicelli et al., 1996, 1999). De surcroît, *PAX3-FKHR* résiste à la répression normalement exercée par hDaxx sur *PAX3* (Hollenbach et al., 1999).
- *PAX5* est notamment impliqué dans un réarrangement chromosomique récurrent avec le locus de la chaîne lourde des immunoglobulines (*IGH*, *immunoglobulin heavy chain*) spécifiquement associé à la survenue de plusieurs lymphomes non-hodgkiniens à cellules B (LNH-B), pouvant être de type indolent tel que le lymphome lymphoplasmocytaire (Iida et al., 1996), ou de type agressif tel que le lymphome diffus à grandes cellules B (LDGCB ; Poppe et al., 2005). Il s'agit de la translocation  $t(9;14)(p13;q32)$  générée par une mauvaise commutation isotypique dans les cellules B des centres germinatifs. La translocation  $t(9;14)$  amène le locus entier et intact de *PAX5* sous le contrôle du locus *IGH* qui possède des promoteurs et activateurs transcriptionnels puissants (Figure 3.04). Notamment dans le cas des LDGCB, elle place le puissant *enhancer* E $\mu$  en amont des promoteurs de *PAX5* (Busslinger et al., 1996). Elle se traduit par la surexpression de *PAX5* (Busslinger et al., 1996; Iida et al., 1996; Ohno et al., 1990) qui promeut la croissance des néoplasmes lymphoprolifératifs susmentionnés via la stimulation de la signalisation du BCR (Cozma et al., 2007). Par ailleurs, la translocation *PAX5-IGH* a été modélisée chez la souris par expression ectopique constitutive d'un minigène *PAX5* à partir du locus *IgH* murin (Souabni et al., 2007). Ces souris transgéniques développent des lymphomes lymphoblastiques (immatures) de type T, démontrant le potentiel oncogénique de cette translocation et soulignant l'importance d'une stricte régulation d'expression de *PAX5* à tous les stades du développement, notamment pour éviter d'interférer avec le développement lymphoïde T ; tandis qu'à l'inverse, la délétion totale expérimentale prématurée de *Pax5* dans les cellules B matures peut également induire le développement de lymphomes agressifs immatures de type B (n'exprimant plus que le marqueur B220) chez la souris (Cobaleda, Jochum, et al., 2007), corroborant son rôle de suppresseur de tumeur en tant que garant de la différenciation et de l'identité lymphoïde B.

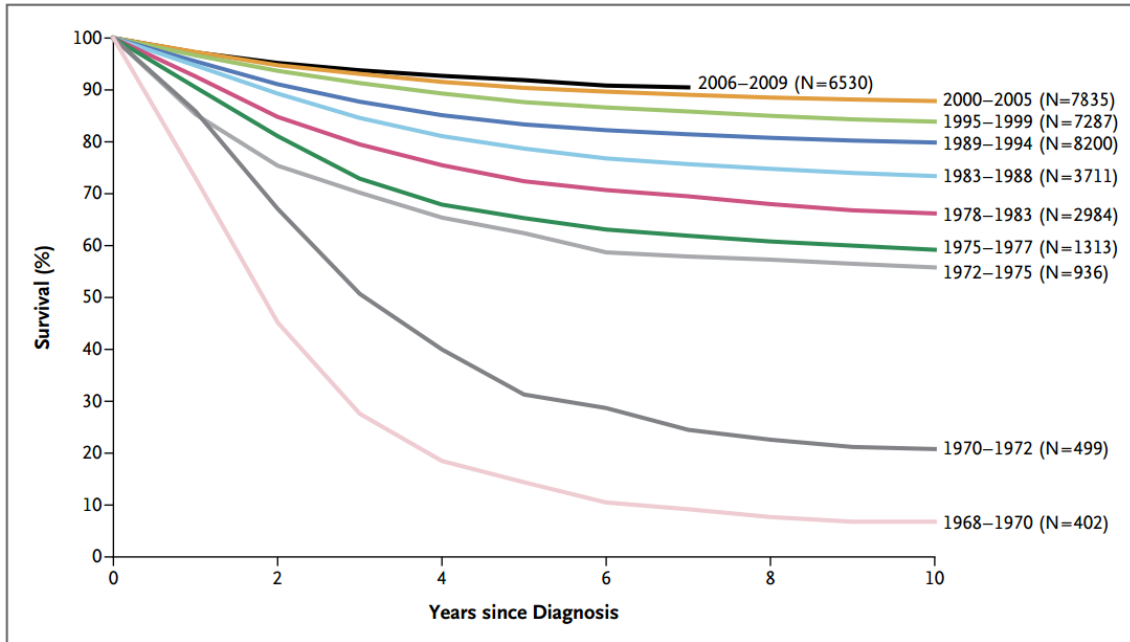


Figure 3.05 : Survie globale des enfants atteints de LAL-B et recrutés dans des essais cliniques de 1968 à 2009. Figure issue de (Hunger & Mullighan, 2015).

Tranches d'âge	< 20 ans			
Incidence annuelle aux États Unis	3/100'000			
Tranches d'âge	0-14 ans		15-19 ans	
Incidence annuelle aux États Unis	4.1/100'000		1.7/100'000	
Tranches d'âge	Nourrissons	2-3 ans	8 ans	> 10 ans
Incidence annuelle aux États Unis	2.25/100'000	> 9/100'000	< 3/100'000	1.8-2.25/100'000

Tableau 5 : Incidences annuelles des LAL en fonction de l'âge aux États-Unis (Hunger & Mullighan, 2015) ; Howlader N, Noone AM, Krapcho M: SEER Cancer Statistics Review (CSR) 1975-2013. Bethesda, Md: National Cancer Institute, 2015. Available online. Voir sections 28 et 29.

## 3.2. Les leucémies aiguës lymphoblastiques de la lignée B

Les altérations somatiques du facteur de transcription lymphoïde *PAX5* représentent un trait caractéristique (*hallmark*) des leucémies aiguës lymphoblastiques de la lignée B (LAL-B) en survenant dans un tiers des cas pédiatriques et adultes (Familiades et al., 2009; Mullighan et al., 2007).

### 3.2.1. Généralités sur les LAL-B

#### 3.2.1.1. Incidences

Les leucémies aiguës lymphoblastiques (LAL) constituent le cancer pédiatrique le plus fréquent, représentant approximativement 25% des diagnostics de cancers de l'enfant de moins de 15 ans (Howlader et al., 2015, référence sous Tableau 5). Les LAL représentent 80% des leucémies aiguës de l'enfant contre 20% des leucémies aiguës de l'adulte. Près de 85 % des LAL de l'enfant dérivent de la lignée B tandis qu'approximativement 75-80% des cas de LAL chez l'adulte sont de la lignée B.

Le taux d'incidence globale des LAL-B se situe entre 1 et 2 nouveaux cas par an pour 100'000 individus (2/100'000 en 2015), affectant ainsi entre 800 et 1000 personnes par an en France avec un âge médian au diagnostic de 17 ans chez les hommes et 22 ans chez les femmes (en 2012 ; Le Guyader-Peyrou et al., 2016). L'incidence des LAL varie en fonction de l'âge. Un pic de fréquence est notamment relevé entre 2 et 5 ans, avec 75% des diagnostics de LAL établis avant 6 ans. Il existe un second pic de fréquence chez l'adulte âgé de plus de 50 ans. De façon similaire aux États-Unis, le taux d'incidence annuelle des LAL est d'environ 3 nouveaux cas pour 100'000 individus de moins de 20 ans, avec un pic de fréquence ayant lieu entre 3 et 5 ans (Ries et al., 1999) ; Tableau 5 ; Howlader et al., 2015). En moyenne, le taux d'incidence des LAL-B pédiatriques est de 5/100'000 (Hunger & Mullighan, 2015).

Les LAL affectent plus fréquemment les garçons/hommes (~60% des cas) que les filles/femmes, avec un sexe-ratio d'incidence de 1.6 en France et mondialement (en 2012 ; Le Guyader-Peyrou et al., 2016) et de 1.3 en moyenne en Europe (Sant et al., 2010).

La fréquence des LAL varie également en fonction des origines ethno-rationnelles puisque le taux d'incidence des LAL est significativement plus faible chez les individus de couleur Noire que chez les individus Caucasiens et se trouve le plus élevée chez les enfants Hispaniques, avec des taux d'incidence respectivement de 1.48, 3.56 et 4.3 cas pour 100'000 individus (Barrington-Trimis et al., 2015; Lim et al., 2014).





### 3.2.1.2. Caractéristiques biologiques et cliniques

Ces hémopathies malignes s'initient dans la moelle osseuse et sont caractérisées par un blocage de la différenciation des lymphocytes B au stade progéniteur pro-B ou précurseur pré-B. Cet arrêt de la différenciation B s'accompagne d'une prolifération clonale excessive de ces lymphoblastes anormaux, envahissant rapidement la moelle osseuse puis les organes lymphoïdes secondaires tels que la rate et les ganglions lymphatiques, et le sang (crise blastique).

La présentation clinique des LAL-B comprend généralement des anomalies de numération de la formule sanguine telles qu'une anémie, une thrombopénie et une neutropénie d'où résultent respectivement une fatigue et une pâleur (syndrome anémique), des ecchymoses et des saignements spontanés (syndrome hémorragique), et des infections (syndrome infectieux). Les infiltrations des lymphoblastes leucémiques dans la rate et les ganglions lymphatiques ainsi que dans le foie sont fréquentes, provoquant respectivement une splénomégalie, une polyadénopathies (hypertrophie ganglionnaire) et une hépatomégalie, observables au diagnostic (syndrome tumoral ; Hunger & Mullighan, 2015).

### 3.2.1.3. Prise en charge thérapeutique

Le traitement standard des LAL-B repose principalement sur de longues et lourdes polychimiothérapies, qui réussissent particulièrement à soigner les patients pédiatriques (Hunger et al., 2012). Le schéma thérapeutique actuel est complexe, comprenant plusieurs phases de traitement pour assurer une rémission complète durable (Hunger & Mullighan, 2015 pour revue):

- une thérapie d'induction pendant 4 à 6 semaines, dont le but est de parvenir à une rémission,
- une thérapie de consolidation (intensification) pendant 6 à 8 mois,
- puis un traitement d'entretien de moindre intensité pendant 2 ans.

Les cycles de traitement incluent notamment des glucocorticoïdes, de la vincristine, de la L-asparaginase, de la cyclophosphamide, des anthracyclines, de la cytarabine, de la 6-mercaptopurine et de la méthotrexate.

Une chimiothérapie intrathécale est usuellement administrée dans le fluide cérébro-spinal pour traiter ou prévenir les atteintes méningées. Les patients avec d'importantes infiltrations leucémiques extra-médullaires dans le CNS et/ou les testicules peuvent éventuellement recevoir une radiothérapie complémentaire.



La polychimiothérapie peut être complétée par une allogreffe de cellules souches hématopoïétiques (CSH) lors de la première rémission, en cas de LAL-B à risque élevé de faible réponse ou de persistance de maladie résiduelle, ou après une récurrence si la rémission partielle ou complète peut être atteinte. Les médecins ont plus fréquemment recours aux transplantations allogéniques de CSH pour des patients adolescentes et adultes car ils répondent moins bien à la chimiothérapie que les enfants (Curran & Stock, 2015).

Par ailleurs, les patients qui expriment le transcrite de fusion *BCR-ABL1*, résultant de la translocation t(9;22) caractérisée par la génération aberrante du chromosome de Philadelphie, *Ph1<sup>+</sup>*, bénéficient du succès des inhibiteurs de tyrosine kinases (TKI) ciblant *BCR-ABL1*, octroyés en combinaison avec la chimiothérapie de rigueur (Aricò et al., 2010; Fielding, 2015; Slayton et al., 2018).

Plusieurs générations de TKI ont été développées (imatinib, dasatinib, nilotinib, bosutinib, ponatinib) et représentent la première et unique thérapie ciblée couramment indiquée dans le cadre des LAL-B. Ils pourraient également être efficaces chez des patients n'exprimant pas *BCR-ABL1* mais d'autres fusions de classe ABL et présentant ainsi un profil d'expression génique similaire aux LAL-B *Ph1<sup>+</sup>* (nouvelle catégorie nommée *BCR-ABL1-like* ou *Ph-like*), (Lengline et al., 2013; Tanasi et al., 2019; Weston et al., 2013).

La chimiothérapie est associée à une toxicité considérable – les effets indésirables peuvent persister – et des effets secondaires peuvent se développer à long terme tels que des tumeurs malignes secondaires, une cardiotoxicité, des anomalies endocriniennes et de fertilité (Fulbright et al., 2011). La chimiothérapie, à l'instar de la radiothérapie, peut engendrer des troubles neurocognitifs. Par conséquent, un suivi clinique et biologique régulier est requis sur le long terme pour contrôler ces complications tardives et pour dépister les éventuelles rechutes.

Le CD19 est un marqueur de surface précoce et spécifique de la lignée B, et est généralement exprimé au diagnostic de leucémies et lymphomes de type B. Par conséquent, il est communément utilisé comme marqueur de diagnostic, de caractérisation et de suivi des hémopathies malignes de la lignée B (Scheuermann & Racila, 1995). Exploitant cette caractéristique, d'innovantes approches thérapeutiques ont récemment émergé et été adaptées au ciblage du CD19 :

- Il s'agit notamment des immunothérapies médiées par les propres lymphocytes T du patient transduites *ex vivo* avec des constructions lentivirales codant des récepteurs antigéniques chimériques dirigés contre les cellules tumorales (*Chimeric Antigen Receptor (CAR)-T cell therapy*), (Grupp et al., 2013; Maude et al., 2015).



- Un autre type de thérapie ciblée a été développé, les conjugués anticorps-médicament (*Antibody-Drug Conjugates*), couplant en l'occurrence un anticorps dirigé contre le CD19 à un maytansinoïde, un dérivé de la maytansine qui est un puissant agent cytotoxique des microtubules (E. E. Hong et al., 2015).

#### 3.2.1.4. Taux de survie

Le taux de survie dépend de la réponse et de la tolérance aux traitements qui sont essentiellement corrélées à l'âge des patients et aux types d'aberrations génétiques à l'origine des leucémies. Les progrès significatifs qui ont été accomplis dans la prise en charge thérapeutique ont globalement fait chuter le taux d'échecs thérapeutiques et de rechutes tout en améliorant la survie et la qualité de vie des patients (Hunger & Mullighan, 2015) pour revue).

- Ainsi, dans les pays développés, le taux de survie globale à 5 ans a atteint 90% chez les enfants de moins de 15 ans, qui tolèrent relativement bien les agents de chimiothérapie (Hunger et al., 2012; Pui et al., 2015; Figure 3.05). Il s'avère qu'approximativement 98% des enfants souffrant de LAL atteignent la rémission et 80% à 90% d'entre eux sont considérés guéris.
- Tandis que le taux de survie (survie globale à 5 ans) des adolescents âgés de 15 à 19 ans a dépassé 75% (Howlader et al., 2015).
- Présentement, environ 85% des patients âgés de 1 à 18 ans nouvellement diagnostiqués avec une LAL et traités selon les régimes de polychimiothérapies actuels sont pressentis pour survivre à long terme sans occurrence d'événement pathologique en lien avec leur LAL, avec une survie à 5 ans dépassant 90% (Peters et al., 2016).
- Cependant le taux de survie à long-terme décline en-dessous de 50% (stagnant entre 45% et 50%) chez les adultes âgés de 40 ans et plus, et leur taux de guérison est estimé entre 20% et 40% (Bassan & Hoelzer, 2011; Frey & Luger, 2015; Jabbour et al., 2015).

La cause la plus fréquente d'échec thérapeutique reste la rechute. La maladie récidive dans 15% à 20% des cas de LAL pédiatriques et chez 40% à 75% des patients adultes. Le pronostic vital des patients qui souffrent de rechute est mauvais (Forman & Rowe, 2013; Hunger & Mullighan, 2015; Liew et al., 2012).

Le pronostic plus péjoratif des LAL-B de l'adulte peut s'expliquer par une faible tolérance à la toxicité de la chimiothérapie intensive (corrélation inverse à l'âge) et à des facteurs de comorbidité préexistants. Les enfants et les adultes présentent des aberrations génétiques similaires mais dans des proportions différentes. Ainsi, le mauvais pronostic des LAL-B de l'adulte peut en partie être attribué la fréquence élevée de sous-types génétiques de LAL-B défavorables (Tableau 6 au dos).

Sous-types génétiques de LAL-B	Pronostic associé	Age médian	Fréquence chez les enfants et les AJAs	Fréquence chez les adultes
<b>Sous-types établis par l'OMS en 2016</b>				
<b>Hypodiploïdie</b> (< 40 chromosomes) <i>Low-hypodiploidy</i> (32-39 chromosomes) <i>Near-haploidy</i> (24-31 chromosomes)	Mauvais Intermédiaire	47 ans 5,4 ans	1% < 3% à tout âge	10-15% < 3% à tout âge
<b>Hyperdiploïdie</b> <i>High hyperdiploidy</i> (51-67 chromosomes)	Excellent pronostic	* 4 ans	25-30%	10%
Réarrangement de <i>KMT2A</i> ( <i>MLL</i> )	Mauvais pronostic	40 ans	7% (60% des enfants de < de 2 ans, 80% des nourrissons)	10-16%
ETV6-RUNX1 ( <i>TEL-AML1</i> )	Excellent pronostic	4 ans	25%	3%
BCR-ABL1 ( <i>Ph+</i> )	Défavorable, pronostic amélioré avec les ITK	40-45 ans	5% (3%)	>30% 40-50%
TCF3-PBX1 ( <i>E2A-PBX1</i> )	Bon pronostic	8 ans	5-8%	5%
TCF-HLF	Très mauvais pronostic	15 ans	0.5%	<1% à tout âge
<b>Entités provisoires (OMS 2016)</b>				
<i>Ph-like</i> (BCR-ABL1- <i>like</i> ) (entité provisoire)	Mauvais pronostic	21 ans	Jusqu'à 15% chez les enfants, 25-30% chez les AJAs	~ 16%
iAMP21 (entité provisoire)	Intermédiaire (bon pronostic avec thérapie intensive)	10 ans	<3% chez les enfants et les AJAs	<1%
<b>Nouveaux sous-types identifiés depuis 2016</b>				
ETV6-RUNX1- <i>like</i>	Intermédiaire à favorable	3 ans	3%	
Réarrangement de <i>MEF2D</i>	Intermédiaire à défavorable (rechute précoce)	14 ans	4% chez les enfants, 7% chez les AJAs	~ 1%
Réarrangement de <i>ZNF384</i> (et <i>ZNF384-like</i> )	Intermédiaire	15 ans	5% chez les enfants et les AJAs	~ 1%
Réarrangement de <i>DUX4</i>	Excellent pronostic	14,3 ans	5% chez les enfants, 8% chez les AJAs	~ 3%
<b>PAX5alt</b> (altérations de <i>PAX5</i> initiatrices autres que <i>PAX5 P80R</i> )	Intermédiaire	10 ans	~ 10%	~ 8%
<b>PAX5 P80R</b>	Intermédiaire à favorable	22 ans	1%	4%
IKZF1 N159Y			<1% à tout âge	
Réarrangement de <i>NUTM1</i>	Excellent pronostic	3 ans	1-2%	
Réarrangements de <i>BCL2</i> et <i>MYC</i>	Mauvais pronostic	48 ans	3% chez les AJAs	3% chez les adultes

Tableau 6 : Prévalence et pronostic associés à chaque sous-type génétique de LAL-B.

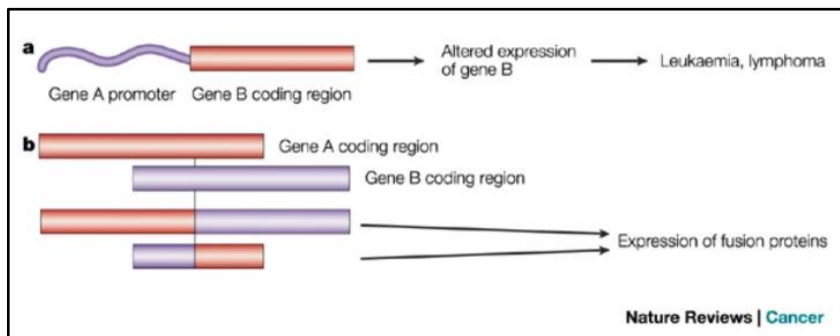
La distribution des altérations génétiques diverge en fonction de l'âge des patients. L'âge médian au diagnostic est indiqué. Les AJAs, Adolescents et Jeunes Adultes, correspondent à la catégorie de patients âgés de 15 à 39 ans ; les adultes correspondent à la catégorie de patients âgés de 40 ans et plus. Tableau établi d'après la classification de l'OMS 2016 (Arber et al., 2016) et nouveaux sous-types d'après les revues (Kimura & Mullighan, 2020; Mullighan, 2019).

*Nota Bene : Un biais de distribution peut subsister car il y a plus de criblages génomiques de cohortes pédiatriques que de cohorte de patients adultes. Par ailleurs, cette distribution et ces pronostics s'appliquent probablement davantage aux patients de pays développés.*

### 3.2.2. Bases moléculaires des LAL-B

Les réarrangements chromosomiques sont des altérations fréquentes et caractéristiques des LAL-B. De ce fait, les LAL-B sont en premier lieu classées selon des anomalies cytogénétiques récurrentes (Tableau 6), comprenant des aneuploïdies (anomalies quantitatives, gain ou perte de chromosomes entiers) et des réarrangements chromosomiques (anomalies quantitatives, qualitatives ou structurales). En effet, des translocations chromosomiques spécifiques (non aléatoires) sont fréquemment associées aux LAL-B (Korsmeyer, 1992) et sont **considérées comme des événements initiateurs de la leucémogénèse** (Rowley, 2001).

- Il existe deux classes fonctionnelles de réarrangements chromosomiques :



- Soit un oncogène est juxtaposé aux régions régulatrices de gènes fortement exprimé causant la dérégulation de l'expression d'une protéine intacte. Par exemple, les réarrangements *CRLF2/IGH*, *EPOR/IGH* et *MYC/IGH* placent les gènes codant les **récepteurs cytokiniques** *CRLF2* (*cytokine receptor-like factor 2*), *EPOR* (*erythropoietin receptor*) et le facteur de transcription *MYC* sous le contrôle du locus *IGH* dans des LAL-B pédiatriques et adultes (Y. Li et al., 2018; Russell, Capasso, et al., 2009; Russell, De Castro, et al., 2009), menant à leur surexpression, à l'instar de *PAX5/IGH* dans le lymphome B précédemment mentionné (Busslinger et al., 1996).
- Soit la translocation chromosomique conduit à la production d'une chimère de fusion oncogénique (oncoprotéine de fusion). Ces translocations impliquent fréquemment des gènes qui régulent normalement l'hématopoïèse et le développement lymphoïde, tels que les gènes codant les **facteurs de transcription** *ETV6*, *RUNX1*, *TCF3* et *PAX5* (O'Neil & Look, 2007), ou les **facteurs épigénétiques** *KMT2A* (lysine N-méthyltransférase 2A, alias *MLL*, *mixed-lineage leukemia*), (McCarthy, 2010) et *NUMT1* (*NUT Midline Carcinoma Family Member 1*), (Hormann et al., 2019). Il peut également s'agir de translocations de gènes impliqués dans la **signalisation cellulaire** comme celle fusionnant les **tyrosine kinases** *BCR* et *ABL1*, menant à leur activation constitutive (Hermans et al., 1987), (Figure 3.06B).



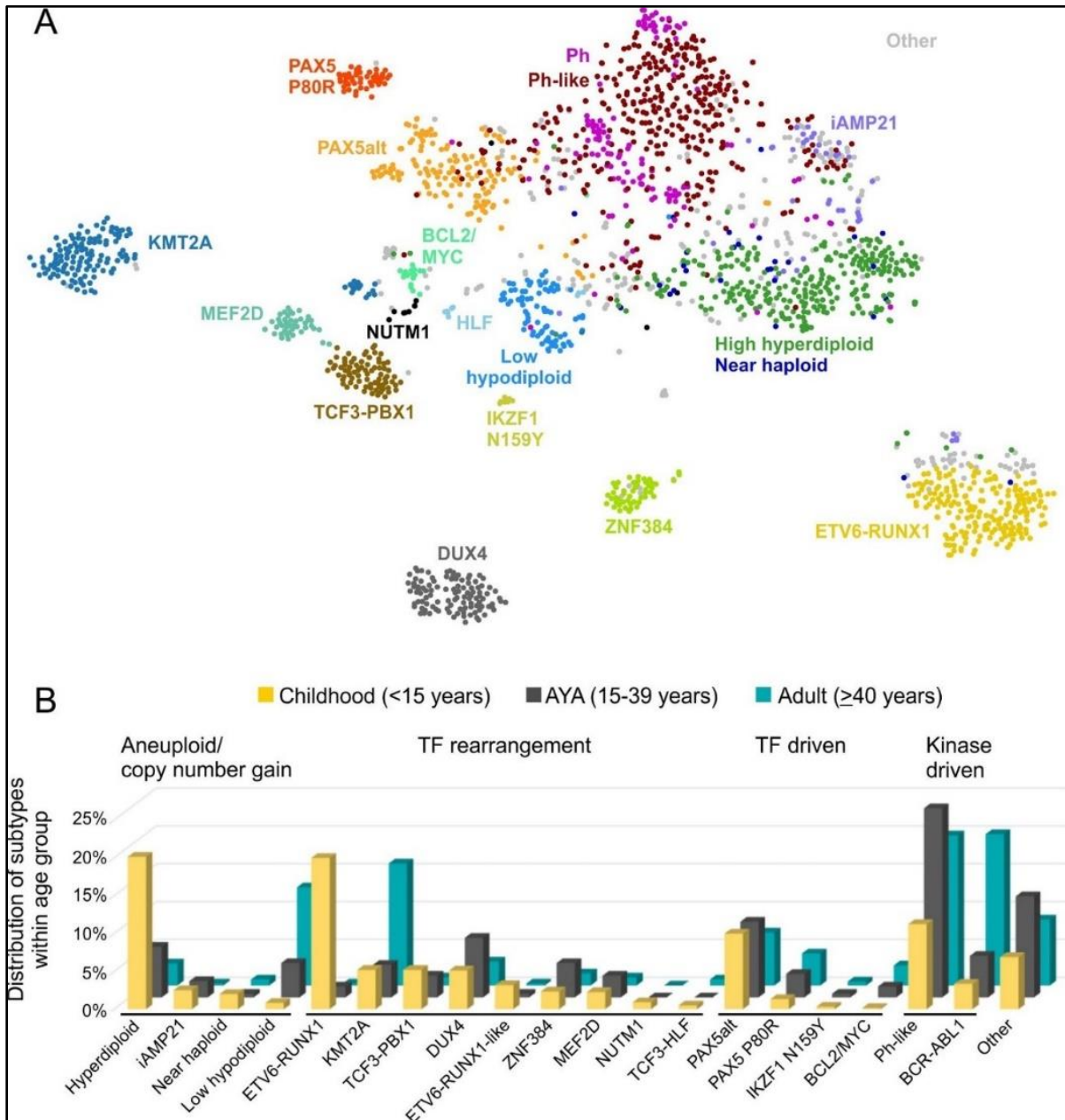


Figure 3.06 : Classification révisée des sous-types moléculaires de LAL-B (2019-2020).

**A.** Principaux sous-types génétiques de LAL-B basés sur le profil d'expression génique de 1988 cas classés en 3 catégories d'âge : enfants entre 0.2 et 15 ans, adolescents et jeunes adultes (AYAs) de 15 à 39 ans, et adultes de 40 ans et plus. **B.** Distribution des sous-types génétiques de LAL-B au sein de chaque catégorie d'âge. Les sous-types sont regroupés entre anomalies chromosomiques de nombre (*Aneuploid/copy number gain*) et anomalies de structure comprenant les réarrangements chromosomiques impliquant des facteurs de transcription (*TF rearrangement*), les autres altérations de FTs (*TF-driven*) ou de kinases (*kinase-driven*), ou tous autres aberrations qui n'entrent pas dans ces sous-groupes ('*other*'). Figures A (t-SNE plot) et B issues de (Gu et al., 2019).

La caractérisation moléculaire des LAL-B se précise au fur et à mesure des progrès technologiques couplés aux collaborations entre les centres de soin et les laboratoires de biologie moléculaire, enrichissant la classification cytogénétique d'ores et déjà établie (par l'OMS en 2008, Campo et al., 2011) et permettant de lever le voile sur les LAL-B sans spécification (par l'OMS en 2016 et avancées ultérieures ; Tableau 6).

En 2008, approximativement 25% des cas de LAL-B ne correspondaient à aucun des sous-types génétiques et présentaient principalement un pronostic intermédiaire à défavorable. Ces LAL-B constituaient alors un sous-groupe **hétérogène** (*B-other ALL*) défini comme n'exprimant aucune des translocations chromosomiques connues (*ETV6-RUNX1*, *BCR-ABL1*, *TCF3-PBX1*, réarrangements de *KMT2A*...) et ne présentant ni d'hyperdiploïdie, ni d'hypodiploïdie.

Le sous-type *BCR-ABL1-like* s'est d'abord démarqué au sein de cette catégorie, étant de mauvais pronostic et mimant un profil d'expression génique similaire à celui LAL-B *Ph*<sup>+</sup> (exprimant *BCR-ABL1*) sans avoir subi la translocation *t(9;22)(q34;q11.2)* (Den Boer et al., 2009; Roberts et al., 2017; Roberts, Li, et al., 2014; Roberts, Pei, et al., 2014). Puis, à l'instar des LAL-B *Ph-like*, d'autres sous-types « *like* » (*ETV6-RUNX1-like*, *ZNF384-like*) ont été décrits, présentant des profils d'expression génique et immunophénotypes similaires à ceux des sous-types princeps (*ETV6-RUNX1*<sup>+</sup>, *ZNF384* réarrangé) sans pour autant exprimer le réarrangement chromosomique en question (Lilljebjörn et al., 2016; Zaliouva et al., 2017).

Ces similarités permettent aux cliniciens de définir avec précision les risques pour le patient, d'adapter l'intensité des chimiothérapies (réduction ou intensification) et d'envisager d'élargir le champ d'application de thérapies ciblées (repositionnement d'ITK), contre les LAL-B *Ph-like* en l'occurrence, associées à un mauvais pronostic quel que soit l'âge et à un risque élevé de rechute.

Il s'avère que tous les sous-types génétiques de LAL-B peuvent être associés des **signatures spécifiques d'expression génique** (Figure 3.06A), (Gu et al., 2019; Y. Li et al., 2018), validant les sous-types cytogénétiques antérieurs. Les algorithmes de profilage transcriptionnel permettraient actuellement de classifier plus de 90% des LAL-B en sous-types génétiques distincts. Ces récentes analyses transcriptomiques effectuées sur de grandes cohortes ont notamment permis d'identifier de nouveaux sous-types (cyto)génétiques caractérisés par les réarrangements de *DUAX*, *MEF2D*, *PAX5*, *ZNF384* et *NUTM1* (Gu et al., 2016, 2019; J.-F. Li et al., 2018; Lilljebjörn et al., 2016).

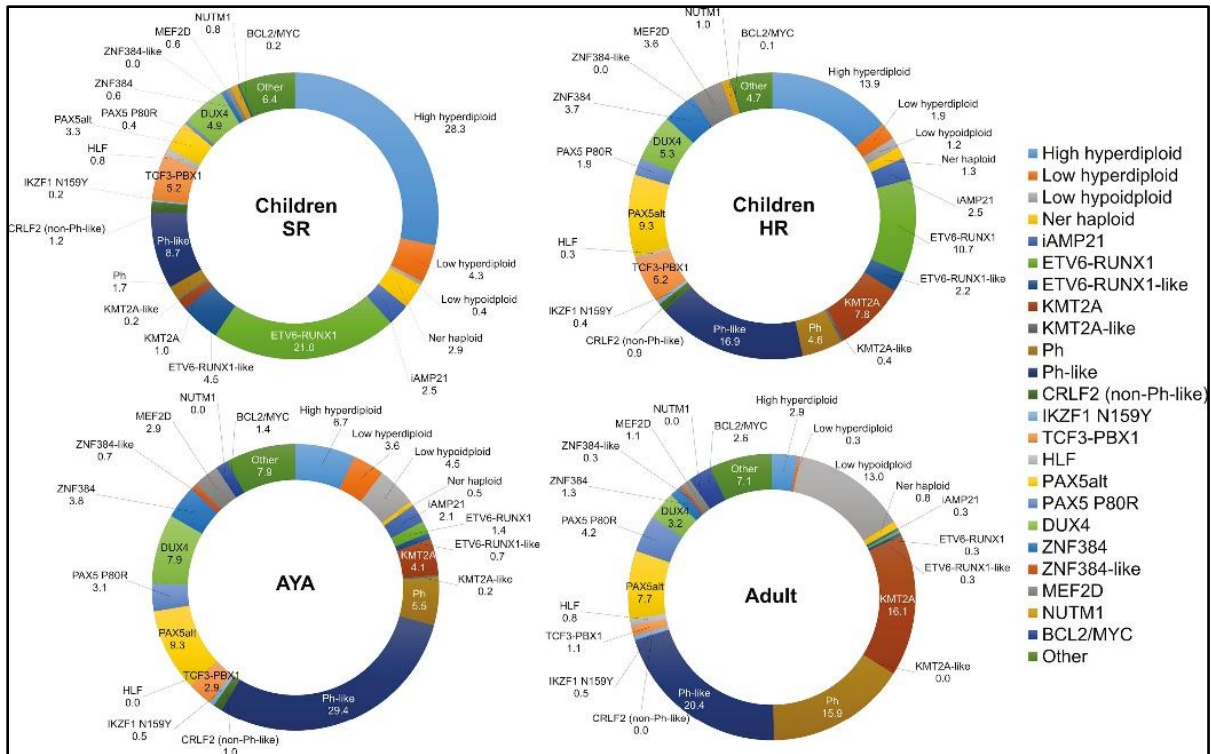


Figure 3.07 : Prévalence des principaux sous-types génétiques de LAL-B selon la catégorie d'âge et le niveau de risque

Cette figure met en évidence la différence d'étiologie moléculaire en fonction de l'âge, en partie à l'origine de la différence de pronostic vital lié à l'âge.

Abréviations : SR, risque standard ; HR, risque élevé ; AYA, adolescents et jeunes adultes.

Figure issue de Kimura & Mullighan, 2020, d'après (Gu et al., 2019).

À toutes ces altérations chromosomiques récurrentes s'ajoutent d'autres altérations génomiques récurrentes, définissant parfois un sous-type à part entière (Tableau 6) :

- des mutations ponctuelles de facteurs de transcription lymphoïdes comme *PAX5* P80R et *IKZF1* N159Y qui sont associées à des signatures moléculaires spécifiques (Gu et al., 2019; Passet et al., 2019) ;
- des amplifications de gènes et gains d'ADN telles que les amplifications intrachromosomiques du chromosome 21 (iAMP21 ; Yilong Li et al., 2014), qui englobent *RUNX1* ; et les amplifications intragéniques de *PAX5* (*PAX5*<sup>AMP</sup> ; C. Schwab et al., 2017) ;
- des délétions partielles ou totales de gènes : les délétions d'*IKZF1* sont présentes dans 15% des LAL-B (Mullighan, Su, et al., 2009) et l'inactivation (mono- ou bi-allélique) de *PAX5* survient dans un tiers des cas (Familiades et al., 2009; Mullighan et al., 2007).

Cette classification génétique vise principalement à stratifier les risques pour les patients

La classification actuelle des LAL-B inclut ainsi de multiples sous-types caractérisés par des altérations génétiques initiatrices récurrentes, sous-types dont la fréquence varie significativement avec l'âge au diagnostic et qui sont corrélés à des valeurs pronostiques différentes (Figure 3.07 ; Tableau 6). Le taux de survie élevé des enfants est partiellement expliqué par une fréquence élevée en anomalies chromosomiques associées à un pronostic favorable comme la translocation la plus fréquente chez les enfants, ETV6-RUNX1 (25% des cas ; Shurtleff et al., 1995) et les hyperdiploïdies (25% à 30% des cas pédiatriques) ; contrairement aux adultes (au taux de survie médiocre) qui sont davantage affectés par des altérations génétiques associées à un pronostic défavorable telles que la translocation BCR-ABL1 (10-16%) et les hypodiploïdies (10-15% des cas adultes). Les LAL du nourrisson sont particulièrement de mauvais augure car 80% d'entre elles présentent un réarrangement de *KMT2A*, qui est associé à un mauvais pronostic (Kimura & Mullighan, 2020 pour revue).

Par ailleurs, ces récentes études pangénomiques et transcriptomiques permettent aussi de spécifier les voies de signalisation dérégulées, potentiellement ciblables pharmacologiquement (Kimura & Mullighan, 2020 pour revue).



### 3.2.3. Altérations génétiques de *PAX5* dans les LAL-B

Les patients présentant des altérations fréquentes de *PAX5* peuvent se classer en deux sous-types génétiques se distinguant par leur signature d'expression génique d'après des algorithmes de profil transcriptionnel (Tableau 7 ci-dessous), (Bastian et al., 2019; Gu et al., 2019; Passet et al., 2019; Zaliova et al., 2019) :

- **Le sous-type *PAX5alt*** (globalement 7.4% des cas), rassemblant une variété d'altérations de *PAX5* (comprenant des réarrangements chromosomiques, des amplifications intragéniques ou des mutations de séquence de *PAX5*) ;
- **Le sous-type *PAX5 P80R***, se définissant par la mutation ponctuelle *PAX5* p.Pro80Arg et des altérations bi-alléliques de *PAX5*.

Sous-types génétiques	Nombre de cas analysés	Enfants à Risque Standard	Enfants à Risque élevé	Adolescents et Jeunes Adultes	Adultes ≥ 40 ans	Critères
<b><i>PAX5alt</i></b>	148	3.3%	9.3%	9.3%	7.7%	<i>Hierarchical gene expression profile cluster enriched with PAX5 alterations</i>
<b><i>PAX5 P80R</i></b>	44	0.4%	1.9%	3.1%	4.2%	<i>PAX5 p.Pro80Arg (P80R) mutation or clustered with PAX5 P80R subtype</i>

Tableau 7 : Prévalence des sous-types *PAX5alt* et *PAX5 P80R* en fonction de l'âge.

Catégories d'âge : enfants entre 0.2 et 15 ans, adolescents et jeunes adultes (AYAs) de 16 à 39 ans, et adultes de 40 ans à 79 ans. *PAX5alt* et *PAX5 P80R* sont des sous-types définis par intégration du profil d'expression génique, du caryotype et/ou des mutations génétiques. *Critères* : Une analyse prédictive de puces à ADN (PAM, Prediction Analysis of Microarrays) a été utilisée pour calculer la similarité du profil d'expression génique avec un/des sous-types de LAL-B spécifiques. D'après les données de (Gu et al., 2019).

#### 3.2.3.1. Fréquence élevée et diversité des altérations de *PAX5*

En somme, *PAX5* est la cible d'altérations somatiques la plus fréquente dans les LAL-B, étant altéré dans un tiers des cas à la fois chez les patients enfants et les patients adultes. En effet, les altérations génétiques somatiques de *PAX5* (délétions, amplifications, mutations, réarrangements chromosomiques) se produisent dans 38,9% des LAL-B pédiatriques (enfants de moins de 15 ans) et dans 34% des LAL-B de l'adulte (Mullighan et al., 2007; Familiades et al., 2009).



De surcroît, parmi tous les gènes altérés dans les LAL-B, *PAX5* se démarque par sa diversité d'altérations génétiques (Coyaud, Struski, Prade, et al., 2010), d'autant plus que chaque type d'altérations est attaché à sa propre diversité de mécanismes d'altération : multiples mécanismes d'inactivation et multiples altérations de structure et de séquence de *PAX5*.

#### 3.2.3.1.1. Diversité des mécanismes d'inactivation de *PAX5*

*PAX5* est inactivé dans approximativement 32% des cas de LAL-B pédiatriques (Mullighan et al., 2007) et adultes (Familiades et al., 2009), ce qui permet de qualifier *PAX5* de gène suppresseur de tumeur haplo-insuffisant dans les LAL-B (Dang et al., 2015).

Les inactivations de *PAX5* se présentent fréquemment sous forme de délétions mono-alléliques allant d'une perte complète du chromosome 9 à la perte d'un sous-ensemble d'exons. Des délétions partielles ou totales sont ainsi détectées dans près de 30% des LAL-B (29.7% chez les enfants, Mullighan et al., 2007; 27% chez les adultes, Familiades et al., 2009), entraînant une modification du niveau d'expression de *PAX5*. Les délétions de *PAX5* peuvent découler d'une délétion d'une portion chromosomique plus large, en l'occurrence du 9p. La majorité des délétions complètes de *PAX5* se produisent simultanément aux délétions de *CDKN2A/B* (gènes suppresseurs de tumeur localisé au locus voisin en 9p21.3) dans les LAL-B (chez 83.3% des enfants et 100.0% des adultes avec une délétion de *PAX5*; Kim et al., 2011).

Les amplifications intragéniques focales de *PAX5* (*PAX5*<sup>AMP</sup>) sont également des événements récurrents [Familiades et al. décrivent un cas isolé chez l'enfant (Familiades et al., 2009); Schwab et al. décrivent 2 cas isolés (C. J. Schwab et al., 2010); Ofverholm et al. établissent la récurrence chez les enfants (Ofverholm et al., 2013)], souvent associés à une trisomie 5 (C. Schwab et al., 2017; Zaliova et al., 2019). Il s'agit de transcrits aberrants de *PAX5* contenant des exons amplifiés, ce qui entraîne une perte de fonction de *PAX5*. Les amplifications de *PAX5* impliquent typiquement les exons 2 à 5 ou seulement l'exon 5 (C. Schwab et al., 2017) et il existe des amplifications atypiques impliquant les exons 4 à 8 (Zaliova et al., 2019). La présence de *PAX5*<sup>AMP</sup> revêt une potentielle valeur pronostique et est préférentiellement associée aux rechutes (Ofverholm et al., 2013; C. Schwab et al., 2017).

L'inactivation de *PAX5* peut résulter d'une inversion entre *PAX5* et le gène adjacent *ZCCHC7* (Figure 1.16A), conduisant à une fusion réciproque qui peut altérer le cadre de lecture de *PAX5* (Roberts et al., 2012).

Les insertions nucléotidiques et micro-délétions (variations de séquence '*indel*') ainsi que les délétions partielles peuvent également générer un décalage du cadre de lecture.



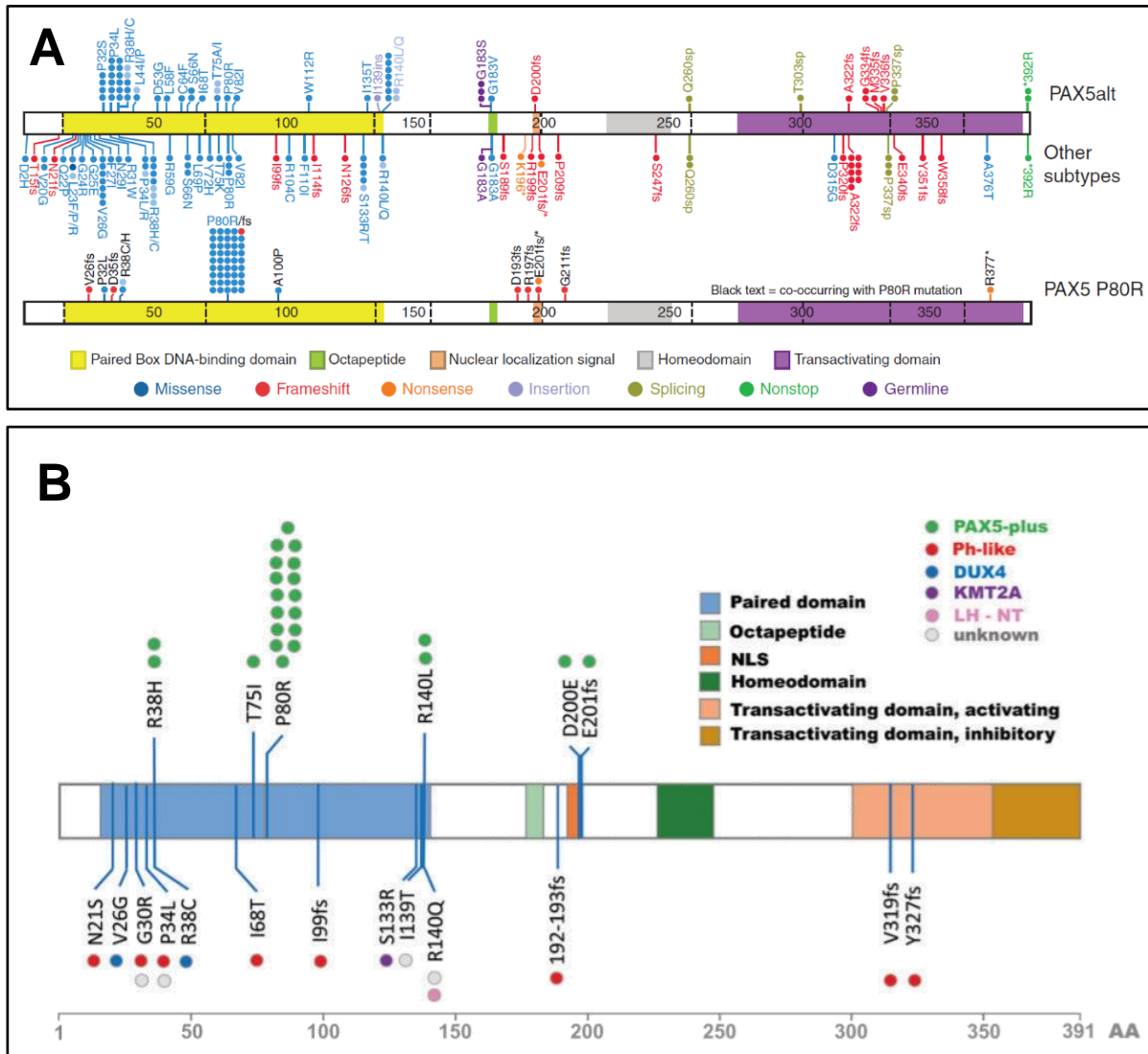


Figure 3.08 : Diversité des mutations de séquence de *PAX5* parmi différents sous-types génétiques de LAL-B

**A.** Schéma des domaines protéiques de *PAX5* recensant les mutations détectées dans le sous-type *PAX5alt* (panel supérieur), dans le sous-type *PAX5 P80R* (panel inférieur) ainsi que dans d'autres sous-types génétiques de LAL-B (panel du milieu). Le nombre de points est proportionnel à la fréquence des altérations et les couleurs indiquent les types d'altération. Figure issue de Gu et al., 2019. **B.** Graphisme représentant l'expression de mutation de séquence de *PAX5* en association avec des sous-types génétiques de LAL-B distincts (*PAX5-plus* = altération bi-allélique de *PAX5* ; profil *Ph-like* ; réarrangement de *DUX4* ; réarrangement de *KMT2A* ; LH-NH = Hypodiploïdies, Low-Hypodiploid, NH, Near-haploid). Le nombre de points est proportionnel à la fréquence des altérations et les couleurs indiquent les sous-groupes moléculaires. Figure issue de Bastian et al., 2019.

#### 3.2.3.1.2. Diversité des mutations de séquence de *PAX5*

Il s'agit de mutations ponctuelles de la phase codante de *PAX5*, identifiées dans 6-7% des LAL-B totales (7.2% chez les enfants dans Mullighan et al., 2007; 6% chez les adultes dans Familiades et al., 2009). De nombreuses mutations de séquence de *PAX5* se répartissent tout le long de *PAX5* mais se concentrent principalement sur les domaines fonctionnels (Figure 3.08). Par exemple, cette diversité de mutations ponctuelles contraste avec *DNMT3A* (codant la DNA (cytosine-5)-méthyltransférase 3A) dont la position hotspot R882 (Ley et al., 2010) est la cible dans 40% des cas de mutations, tandis que les autres 60% des altérations de *DNMT3A* sont des mutations perte-de-fonction (D. J. Park et al., 2020).

Les mutations affectant la moitié N-terminale de *PAX5* sont supposées conférer un gain de fonction (aberrant), avec le hotspot mutationnel P80R (Bastian et al., 2019; Gu et al., 2019; Passet et al., 2019; Zaliouva et al., 2019) comme exemple représentatif de ce consensus (*si un événement génétique devient récurrent, il est plus probable que ce soit une mutation gain-de-fonction*). À l'inverse, les mutations affectant la moitié C-terminale de *PAX5* sont supposées entraîner une perte de fonction, notamment à cause de la perte de fonctionnalité du module de régulation transcriptionnel (domaine de transactivation accolé au domaine d'inhibition à l'extrémité C-ter).

#### 3.2.3.1.3. Impact fonctionnel des inactivations et mutations de *PAX5*

Tous ces mécanismes d'inactivation génétique de *PAX5* ont pour conséquence soit la diminution ou la perte de l'expression de *PAX5*, soit la perte de son activité de liaison à l'ADN ou de régulation transcriptionnelle, pouvant mener à une haploinsuffisance de *PAX5* (Mullighan et al., 2007).

Par exemple, un décalage du cadre de lecture (résultant d'une micro-*indel* ou d'une mutation ponctuelle non-sens) peut se traduire par l'expression d'une protéine tronquée ayant perdu un ou plusieurs domaines fonctionnels (Figure 3.08) ; tout comme une mutation ponctuelle faux-sens peut endommager un domaine fonctionnel. Les mutations de séquence de *PAX5* sont majoritairement des mutations perte-de-fonction (à l'instar de *TP53*, un illustre autre gène suppresseur de tumeur). Les mutations ponctuelles (sans distinction) aboutissent généralement à une réduction significative de la transactivation transcriptionnelle médiée par *PAX5* (Mullighan et al., 2007). Cependant, les altérations de *PAX5* sont hétérozygotes chez la majorité des patients atteints de LAL-B (délétion hétérozygote partielle ou totale, ou mutation perte-de-fonction) et l'expression des gènes cibles de *PAX5* (comme *CD19*, *CD79A* et *BLNK*) peut être garantie par l'expression de l'autre copie de *PAX5* (Mullighan et al., 2007).

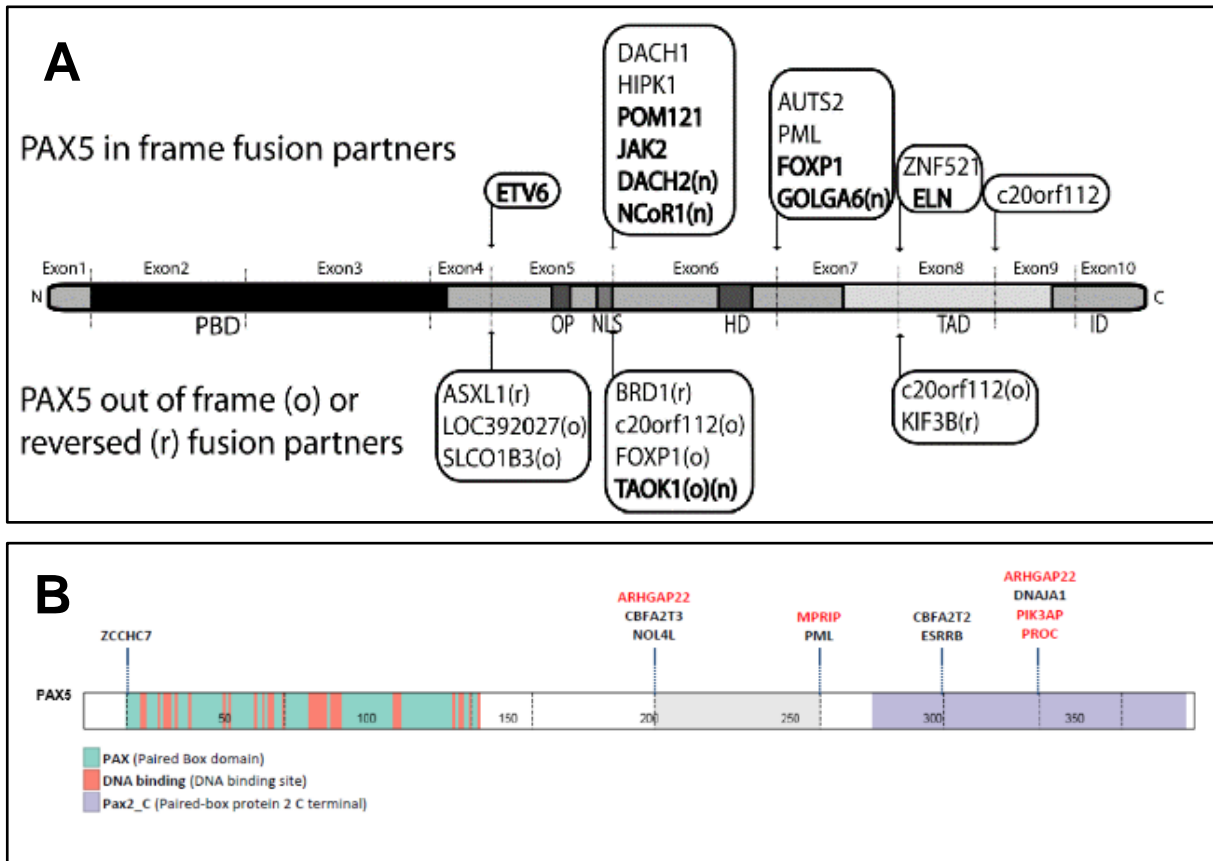


Figure 3.09 : Représentations schématiques des points de fusion de certains partenaires de fusion de *PAX5* impliqués dans les différents types de mécanisme de translocation de *PAX5*. Figure supérieure (A) issue de Coyaud et al., 2010. Figure inférieure (B) issue de Zaliouva et al., 2019.

#### 3.2.3.1.4. Le hotspot mutationnel *PAX5* P80R

La principale mutation ponctuelle est *PAX5* P80R, retrouvée dans 5% des LAL-B et touchant le domaine de liaison à l'ADN en substituant une proline par une arginine, ce qui altère le domaine de liaison à l'ADN (domaine PAIRED) de *PAX5* (5.3% dans LAL-B de l'adulte, substitution c.239C.G, p.P80R identifiée dans tous les cas de LAL-B rapportés par Passet et al., 2019). Sa prévalence augmente avec l'âge et est associée à un pronostic relativement favorable (Passet et al., 2019 ; Tableau 7 en 3.2.3 reprenant les données de la Figure 3.07, d'après Gu et al., 2019). Elle est typiquement combinée à une inactivation du second allèle normal de *PAX5* (par délétion ou mutation), une délétion bi-allélique de *CDK2A* et des altérations de la voie RAS. La mutation *PAX5* P80R est pressentie pour être une mutation gain-de-fonction (c'est-à-dire attribuant un potentiel oncogénique à *PAX5*), étant donné sa récurrence, son exclusivité envers les translocations chromosomique et sa capacité à causer une signature d'expression génique spécifique, résultant de la combinaison unique d'altérations ciblant *PAX5*, *CDKN2A* et *RAS*, définissant ainsi un sous-type génétique de LAL-B (*PAX5* P80R) à part entière (Zaliova et al., 2019, Passet et al., 2019, Gu et al., 2019, Bastian et al., 2019). De cette manière, *PAX5* P80R se distingue à la fois des autres mutations de séquences mais également des altérations de structure (réarrangements chromosomiques de *PAX*) qui sont associées à une autre signature d'expression génique (*PAX5* altéré ; Gu et al., 2019 ; Tableau 7).

#### 3.2.3.1.5. Diversité des réarrangements chromosomiques de *PAX5*

Les réarrangements structuraux du gène *PAX5* (*PAX5-X* où *X* désigne le partenaire de fusion) correspondent (Figure 3.09A) :

- soit à des translocations chromosomiques respectant le cadre de lecture (*in-frame fusion*),
- soit à des fusions ne respectant pas le cadre de lecture (*out-of-frame*),
- soit à des fusions de *PAX5* à des gènes dans l'orientation opposée (*reversed fusion*).

Il existe deux mécanismes de formation de transcrite de fusion *in-frame* impliquant *PAX5* :

- Le type de fusion le plus fréquent comprend la partie 5' de *PAX5* (exons 1 à 5/6/7/8) fusionnée à divers partenaires de fusion, tandis que le domaine de transactivation de *PAX5* est partiellement ou totalement perdu.
- Le second type est généré par inversion intrachromosomique et représenté par *ZCCHC7-PAX5*, impliquant la partie 3' de *PAX5* en plus du domaine de liaison à l'ADN (exons 2-10), (Figure 3.09B).

**Toutes les chimères de fusion impliquant *PAX5* conservent systématiquement le domaine de liaison à l'ADN (domaine PAIRED) situé à l'extrémité N-terminale de *PAX5* et perdent (en partie ou dans sa totalité) le module de régulation transcriptionnelle incluant le domaine de transactivation (TAD/ID) de *PAX5*, situé à l'extrémité C-terminale.**



Les réarrangements de *PAX5* conduisent à l'expression de transcrits de fusion, comme *PAX5-ETV6* (le plus fréquent des *PAX5-X*) et *PAX5-ELN* (notre favorite), dans approximativement **2.5%** des LAL-B (Coyaud, Struski, Prade, et al., 2010; Nebral et al., 2009).

De nombreux partenaires de fusion de *PAX5* ont été identifiés, constituant **le recombino**  
**de *PAX5*** (Huret et al., 2013 ; et se référer à la section références supplémentaires). En 2008, au moins 12 fusions *in-frame* de *PAX5* avaient été décrites, élevées au nombre de 16 en 2014 (An et al., 2008; Bousquet et al., 2007; G. Cazzaniga et al., 2001; Coyaud, Struski, Prade, et al., 2010; Kawamata et al., 2008; Lee et al., 2012; Mullighan et al., 2007; Nebral et al., 2007, 2009). Le recombino de *PAX5* ne cesse de s'étoffer de divers partenaires de fusion grâce au séquençage de larges cohortes internationales. **Plus d'une cinquantaine de réarrangements chromosomiques impliquant *PAX5* sont actuellement recensés dont au moins une vingtaine de fusions *in-frame*** (listés dans la base de données Mitelman), se rapprochant du recombino de *KMT2A* doté de plus de 80 partenaires de fusion (Winters & Bernt, 2017). Les partenaires de fusion de *PAX5* constituent un groupe de gènes très **hétérogènes**, codant des facteurs de transcription (ETV6, FOXP1, ZNF521), des récepteurs nucléaires (ESRRB), des protéines transmembranaires (TMEM14B), des protéines structurales nucléaires ou extracellulaires (PML, ELN), des kinases (JAK2), des modulateurs de l'épissage alternatif (MBLN1), voire des partenaires aux fonctions encore inconnues (AUTS2, C20orf112), (Coyaud, Struski, Dastugue, et al., 2010; Gu et al., 2016; Lilljebjörn et al., 2016; Mullighan et al., 2007; Nebral et al., 2007, 2009). Les réarrangements chromosomiques de *PAX5* les plus fréquents sont *PAX5-ETV6* (de loin en tête), et *PAX5-NOL4L* (*PAX5-C20orf112*), notamment suivis de *PAX5-ELN*, *PAX5-FOXP1* et *PAX5-JAK2*, chacun d'eux ayant été identifié dans au moins trois cas non-apparentés de LAL-B. La plupart des fusions impliquant *PAX5* sont rares et de nombreux réarrangements ont seulement été identifiés chez un ou deux patients non-apparentés (suffisants à les qualifier de translocations récurrentes).

Les modes d'actions et mécanismes moléculaires sous-jacents à ces diverses translocations de *PAX5* sont encore peu décrits (c'est pourquoi il est intéressant de comparer avec les effets des translocations impliquant d'autres gènes *PAX* que *PAX5*). Ses divers partenaires influencent différemment le rôle des chimères de fusion *PAX5-X*. Certes, les protéines de fusion chimériques se localisent dans le noyau et sont capables de se fixer aux promoteurs des gènes cibles de *PAX5*, conformément à la conservation et l'intégrité du domaine de localisation nucléaire et du domaine de liaison à l'ADN de *PAX5*, et certaines ont un effet dominant-négatif *in vitro* (N. Kawamata et al., 2012; Norihiko Kawamata et al., 2010; Bousquet et al., 2007; Mullighan et al., 2007).



De plus, certaines protéines de fusion de *PAX5* se caractérisent par une aptitude à auto-interagir alors que d'autres activent des programmes transcriptionnels et participent à l'hyperactivation de voies de signalisation cellulaire oncogéniques et ainsi n'agissent pas seulement comme des inhibiteurs compétitifs (V. Cazzaniga et al., 2015; Fazio et al., 2013; Fortschegger et al., 2014; Smeenk et al., 2017).

### 3.2.3.2. Différents degrés d'implication dans la leucémogénèse et exclusivité d'expression des altérations de *PAX5*

La nature des altérations de *PAX5* est corrélée à sa contribution à la leucémogénèse des LAL-B, soit en tant qu'événement initiateur précoce (aussi appelé lésion fondatrice ou *princeps*), soit en tant qu'événement secondaire coopératif. Un événement peut être considéré comme initiateur par la littérature le définissant ainsi (étude-s fonctionnelle-s spécifique-s), par la récurrence de l'altération et par la présence de l'altération à la fois au diagnostic et à la rechute.

Les altérations de *PAX5* de type délétion de *PAX5* sont considérées comme des événements secondaires car elles sont toujours associées à la présence d'autres anomalies cytogénétiques (telles que *BCR-ABL1* ou *TCF3-PBX1*) ou à un caryotype complexe (au moins 3 altérations non-relées ; Coyaud, Struski, Prade, et al., 2010; Familiades et al., 2009; Iacobucci et al., 2010) alors que les altérations de *PAX5* telles que les translocations chromosomiques (comme *PAX5-ELN*) et certaines mutations somatiques ponctuelles (se résumant à *PAX5 P80R*) semblent être des événements initiateurs (*princeps*) car ils sont détectés isolés au diagnostic (Coyaud, Struski, Prade, et al., 2010). En règle générale, les translocations chromosomiques associées aux LAL-B sont détectées dans tous les clones et sont présentes au diagnostic et souvent à la rechute (Hunger & Mullighan, 2015). Les translocations chromosomiques de *PAX5* ne sont associées ni à l'expression d'un autre transcrite de fusion oncogénique (tel que *KMT2A-AF4* ou *BCR-ABL1*) ni à un caryotype complexe (Coyaud, Struski, Prade, et al., 2010).

En théorie, au-delà du cas par cas, les mutations ponctuelles affectant la moitié N-terminale de *PAX5* sont considérées comme des lésions *driver* initiateuses de leucémogénèse (telle que *PAX5 P80R*) tandis que les mutations ponctuelles affectant la moitié C-terminale de *PAX5*, ainsi que toute mutation qui correspondrait à une mutation strictement perte-de-fonction, interviennent essentiellement dans la leucémogénèse en tant qu'événements secondaires et ce dans plusieurs sous-types génétiques de LAL-B (Figure 3.08). Les altérations de type perte-de-fonction (délétions, mutations de séquence, *et cetera*) de *PAX5* peuvent être concomitantes à une autre altération sur le second allèle de *PAX5* (Bastian et al., 2019; Gu et al., 2019; Passet et al., 2019; Figure 3.08).





Parmi les *exceptions qui confirment la règle*, quelques cas isolés de translocations chromosomiques concomitantes ont récemment été décrits :

- Le transcrite de fusion *EMILIN2-PAX5* a été identifié chez un patient (homme de 27 ans) exprimant le transcrite de fusion *NUP214-ABL1* avec une délétion d'*IKZF1*, caractérisant un profil *Ph-like*, de mauvais pronostic (Liu et al., 2017).
- Les transcrits de fusion *PAX5-SPECC1* (accolés en direction opposée) et *TCF3-PBX1* ont été tous les deux détectés au diagnostic et à la rechute chez un patient de 16 ans à l'issue fatale (Barbosa et al., 2018).
- Les transcrits de fusion *PAX5-NOL4L* et *PAX5-C20orf112* sont coexprimés chez un patient au profil *Ph-like* avec le gène de l'IL-7R muté (Gu et al., 2016).
- Les transcrits de fusion *PAX5-NCOA5* et *P2RY8-CRLF2* sont coexprimés chez un patient avec une LAL-B assimilée au sous-type génétique avec réarrangement de *CRLF2* (Gu et al., 2016).

Les réarrangements ne produisent pas systématiquement une oncoprotéine de fusion. Les réarrangements *PAX5-X* concomitants à d'autres translocations *drivers* visent à perturber la fonction de *PAX5* de façon similaire à une délétion de *PAX5* (Barbosa et al., 2018). Ainsi, *PAX5-AUTS2* est classifié dans le sous-type *PAX5-altéré* s'il survient en tant qu'événement *princeps* (isolé) ou au sous-type *CRLF2*-réarrangé s'il est coexprimé avec *P2RY8-CRLF2* (Gu et al., 2016; Liu et al., 2017).

Faisant écho à l'exclusivité des événements initiateurs précoces (isolés au diagnostic), les fusions géniques de *PAX5*, les amplifications de *PAX5* et la mutation faux-sens *PAX5 P80R* constituent des altérations mutuellement exclusives (C. Schwab et al., 2017; Zaliouva et al., 2019).

Sous-types de LAL-B selon l'EGIL		Expression des marqueurs déterminants				Fréquences des sous-types immunophénotypiques	
		(s)CD19, (c/s)CD22, (c)CD79A	(s)CD10	(c) $\mu$	(s)Ig	Fréquence chez les patients pédiatriques	Fréquences chez les patients adultes
<b>B-I</b>	→ LAL pro-B	+	-	-	-	5%	10%
<b>B-II</b>	→ LAL-B commune	+	+	-	-	60%	60%
<b>B-III</b>	→ LAL pré-B	+	+/-	+	-	15%	10%
<b>B-IV</b>	→ LAL mature B	+	+/-	+/-	+	3%	4%

Tableau 8 : Immunophénotype des blastes leucémiques de LAL-B selon la classification établie par l'EGIL (European Group for the Immunological characterization of Leukaemias) et distribution des sous-types immunophénotypiques dans les LAL-B pédiatriques et adultes. (c) cytoplasmique ; (s), surface. D'après (Bene et al., 1995; Lemmers et al., 2000).

### 3.2.4. Classification immunophénotypique des LAL-B

#### 3.2.4.1. Sous-types immunophénotypiques selon l'EGIL

L'immunophénotypage des blastes leucémiques par cytométrie en flux est la procédure diagnostique de référence pour l'identification de la lignée d'où dérive la LAL et la définition du sous-type de LAL.

La classification immunophénotypique des LAL-B se base sur l'utilisation d'un panel d'anticorps monoclonaux dont l'expression détermine le stade de différenciation/maturation auquel les blastes leucémiques sont bloqués. L'EGIL (*European Group for the Immunological characterization of leukaemias*) classe ainsi les LAL-B en quatre sous-types, de B-I à B-IV, essentiellement en fonction de l'expression des marqueurs CD19, CD22 (cytoplasmique ou membranaire), CD79A (cytoplasmique), CD10 (Bene et al., 1995), de l'expression de la chaîne lourde  $\mu$  (cytoplasmique, c $\mu$ ) et des chaînes légères  $\kappa$  et  $\lambda$  des immunoglobulines (Ig). Tous les lymphoblastes de LAL-B expriment le CD19 et/ou le CD22 et/ou le CD79A (au moins deux de ces marqueurs). D'autres marqueurs sont importants pour le diagnostic et la caractérisation d'une LAL-B, tels que le CD20 et le CD24.

- **Sous-type B-I (*Pro-B ALL*)** : L'expression d'au moins deux des trois marqueurs B les plus précoces (CD19, CD22, CD79A), sans expression d'autres marqueurs de la différenciation B, définit les LAL pro-B.
- **Sous-type B-II (*Common B-ALL*)** : La co-expression du CD19 et du CD10 (ou antigène *CALLA* pour *Common Acute Lymphoblastic Leukemia Antigen*, une métalloendopeptidase de surface) définit les LAL-B communes. Elles représentent le sous-type de LAL-B le plus fréquent (60%) chez l'enfant et l'adulte.  
De plus, les lymphoblastes des LAL-B de sous-types B-I et B-II expriment souvent le CD24 et réagissent à l'anticorps 4G7 spécifiquement dirigé contre les pseudochaînes légères à la surface des pro-B et pré-B.
- **Sous-type B-III (*Pre-B ALL*)** : L'expression cytoplasmique de la chaîne lourde  $\mu$  des Ig définit les LAL pré-B.
- **Sous-type B-IV (*Mature B-ALL*)** : L'expression membranaire des immunoglobulines (Ig de surface, sIgM) définit les LAL B-matures.  
De plus, l'expression membranaire de CD20 et CD22 est variable au-delà du stade pro-B (sous-type B-I) alors que l'expression de TdT est généralement perdue dans le sous-type B-IV.

Par ailleurs, le CD34 (marquant les cellules souches et progéniteurs multipotents hématopoïétiques) peut être exprimé, tout comme les antigènes 'myéloïdes' CD13 et/ou CD33 qui sont exprimés dans approximativement 30% des cas.



### 3.2.4.2. Corrélation immunophénotype-cytogénétique

Certains sous-types cytogénétiques sont associés à un sous-type immunophénotypique. Par exemple :

- Les réarrangements de *KMT2A (MLL)* mènent à des LAL pro-B (sous-type B-I) souvent associées à l'expression de marqueurs myéloïdes (dont CD15).
- Les LAL-B exprimant BCR-ABL1 sont toujours de sous-type B-II et expriment souvent CD34, CD38, CD25 et CD13/33. Les LAL-B avec une hyperdiploïdie ou la translocation *ETV6-RUNX1* sont très souvent de sous-type B-II (*common B-ALL*).
- Les translocations *TCF3-PBX1* et les altérations de *PAX5* (sous-type *PAX5alt*) causent des LAL pré-B (*pre-B ALL* ; sous-type B-III). Les réarrangements de *MEF2D* mènent également à des leucémies très similaires aux LAL pré-B.
- Par ailleurs, certains sous-types génétiques peuvent être associés à des immunophénotypes distincts, comme les réarrangements de *DUX4* qui caractérisent spécifiquement les LAL-B exprimant CD371 (Schinnerl et al., 2019).

Inversement, les réarrangements de *ZNF384* sont retrouvés dans plusieurs sous-types immunophénotypiques, tels que des LAL pré-B (sous-type B-III), des LAL pro-B très précoces (sous-type B-I) exprimant les marqueurs myéloïdes CD13/33, ou encore des leucémies aiguës bi-phénotypiques (leucémies aiguës de phénotype mixte, LAPM) (Alexander et al., 2018).

→ Les LAL-B avec altérations bi-alléliques de *PAX5* (dont les LAL-B avec *PAX5 P80R*) ont un immunophénotype caractéristique de LAL-B commune (Bastian et al., 2019) alors que les patients recensés avec une LAL-B exprimant *PAX5-ELN* présentent différents immunophénotypes (B-I, B-II, B-III, (Denk et al., 2014) pour revue).

Bien que la classification cytogénétique et moléculaire ait supplanté la classification immunophénotypique en termes de facteur pronostique, l'immunophénotypage reste indispensable pour confirmer le diagnostic de LAL-B et demeure utile pour surveiller la maladie résiduelle minimale (*MRD, minimal residual disease*).

### 3.2.5. Pathogenèse des LAL-B

#### 3.2.5.1. Les cancers résultent de l'accumulation de mutations somatiques

Il est admis que les mutations s'accumulent naturellement dans les cellules somatiques au cours de la vie, d'autant plus lorsque le tissu possède un fort potentiel réplicatif, comme le tissu hématopoïétique qui se renouvelle rapidement (Blokzijl et al., 2016; Tomasetti & Vogelstein, 2015). À l'instar de certains cancers solides (du système digestif) qui seraient induits seulement par quelques mutations somatiques *drivers* acquises dans un ordre spécifique, la pathogenèse des hémopathies malignes résultent de l'acquisition séquentielle d'altérations génétiques qui de concert établissent le clone leucémique (Mullighan, 2012; Ortmann et al., 2015; Vogelstein & Kinzler, 2015). En moyenne, les génomes des LAL pédiatriques comportent seulement 10 à 20 mutations codantes non-silencieuses lors du diagnostic, et environ deux fois plus lors de la rechute (Hunger & Mullighan, 2015; Ma et al., 2015).

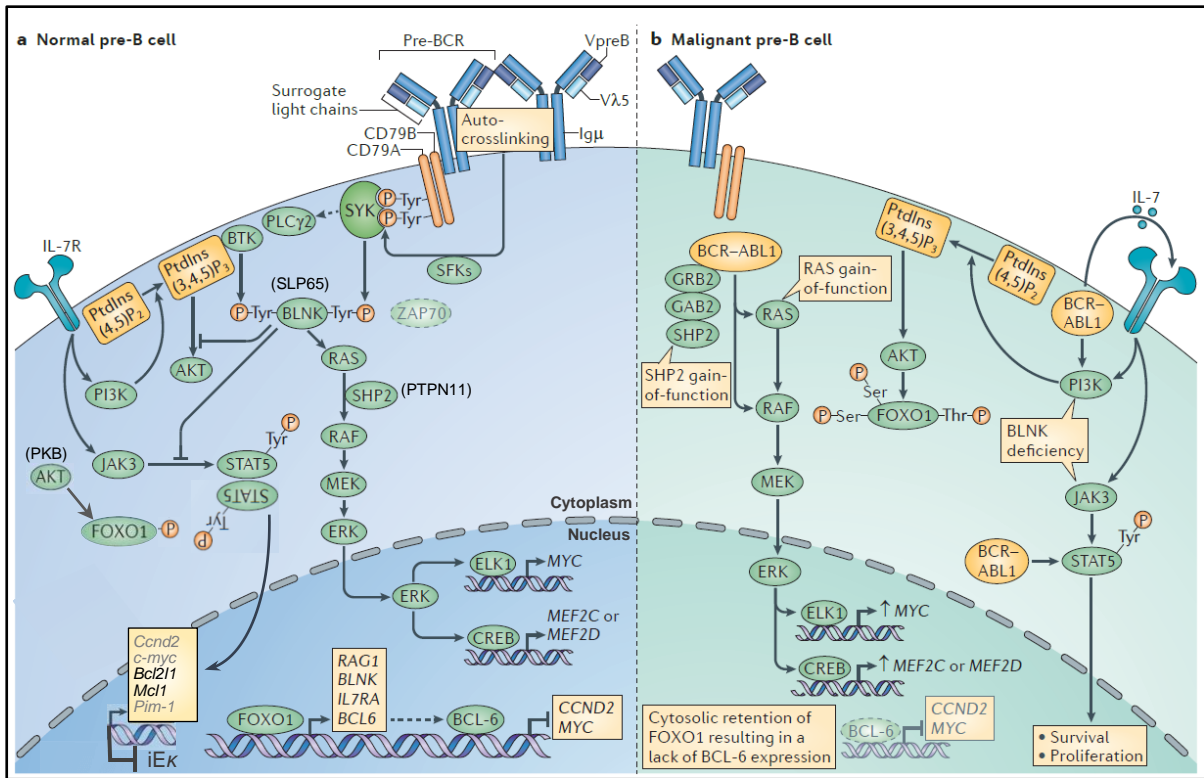


Figure 3.10 (b) : Dérégulation de la signalisation du pré-BCR dans les LAL-B.

*Nota Bene* : SHP2 est codé par *PTPN11*. Le panel de gauche (a) correspond à la Figure 2.11 (*Synergie des signaux de l'IL-7R et du pré-BCR chez la souris*) décrit dans la section 2.2.3 (*Voies de signalisation principales de la lymphopoïèse B précoce*). Figure originale issue de Rickert, 2013.

### 3.2.5.2. La dérégulation de la lymphopoïèse est la cause principale des LAL-B

Le développement leucémique des LAL-B découle de multiples atteintes géniques (primaires ou secondaires) qui perturbent principalement le développement lymphoïde B, la signalisation sous les récepteurs cytokiniques et la signalisation du pré-BCR (Figure 3.10). En effet, elles affectent des voies de signalisation primordiales à la lymphopoïèse B (J. Zhang et al., 2011) qui normalement contrôlent la prolifération et survie cellulaires, et maintiennent un équilibre entre ces processus biologiques (voir section 2.2.3). La majorité des réarrangements chromosomiques caractéristiques des LAL-B et des autres altérations génétiques récurrentes (sous-microscopiques telles que des délétions, gains d'ADN/amplification, mutations ponctuelles) ciblent ainsi :

- des facteurs de transcription (tels que PAX5, IKZF1, EBF1, RUNX1, TCF3, ETV6) qui régulent les programmes transcriptionnels de l'hématopoïèse et en particulier de la différenciation lymphoïde B,
- des composants des voies de signalisation sous l'IL-7R (tels que l'axe JAK3/STAT5) et le pré-BCR (tels que RAS et SHP2 codé par *PTPN11*). Ce sont typiquement des mutations activatrices qui favorisent la prolifération et la survie des cellules tumorales (J. Zhang et al., 2011 ; Figure 3.10).

Les altérations structurales, essentiellement des délétions partielles ou totales (anomalies du nombre de copies), ciblant les gènes codant les régulateurs transcriptionnels du développement des cellules B et de la différenciation lymphoïde B sont détectées dans plus de deux tiers des cas de LAL-B (Kuiper et al., 2007; Mullighan et al., 2007; Mullighan & Downing, 2009). Les cibles les plus fréquentes de ces altérations sont les gènes codant les facteurs de transcription PAX5, IKZF1 et EBF1 – 40% des LAL-B présentent des altérations de ces FTs avec *PAX5* en tête, étant altéré dans près d'un tiers des cas de LAL-B pédiatriques et adultes (Familiades et al., 2009; Mullighan et al., 2007) – auxquelles s'ajoutent des délétions de *LEF1*, *IKZF3*, *TCF3* et *BLNK* ainsi que d'*ERG* (Clappier et al., 2014).

### 3.2.5.3. La leucémogénèse est un processus multi-étapes

La récurrence et la caractérisation de ces altérations génétiques contribuant à la leucémogénèse et à la rechute, et l'étude de l'architecture clonale des LAL-B, ont permis de déduire un modèle en 3 étapes majeures (Figure 3.11). Brièvement, la maladie s'établit progressivement : une lésion initiatrice ne suffit pas à induire de leucémie à elle seule et requiert l'acquisition de lésions oncogéniques coopératives pertinentes causant l'acutisation leucémique, à l'exception des cas infantiles de réarrangements de *KMT2A*, présentant une faible fréquence de mutations secondaires somatiques, celles-ci étant souvent sous-clonales et fréquemment perdue à la rechute (Andersson et al., 2015).



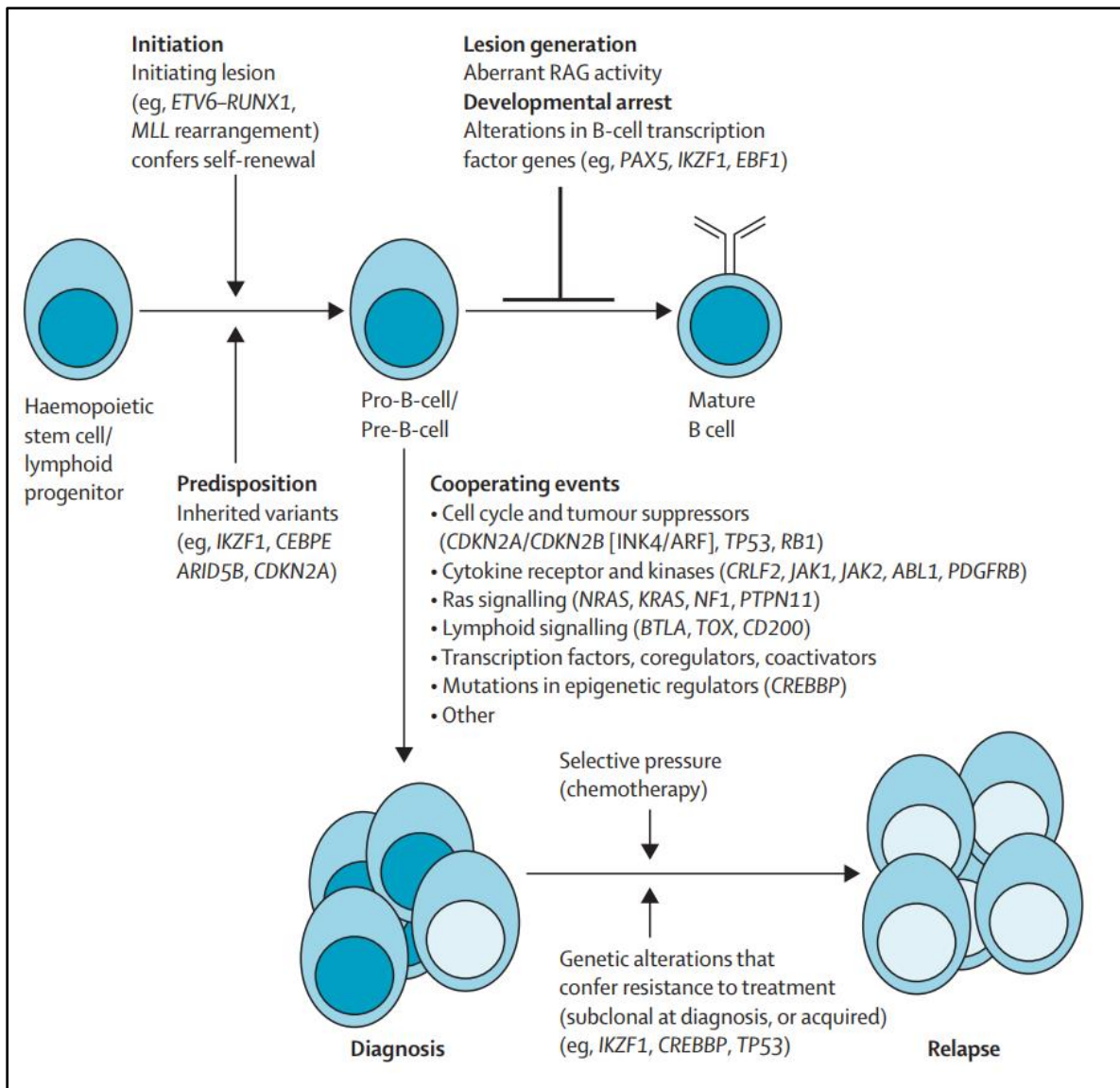


Figure 3.11 : Modèle de pathogénèse des leucémies aiguës lymphoblastiques de la lignée B au diagnostic et à la rechute, schématisant l'acquisition séquentielle des altérations génétiques qui contribuent à la leucémogénèse. Figure issue de Inaba et al., 2013.

① Une lésion génétique initiatrice est acquise dans un progéniteur lymphoïde. Il peut s'agir d'une translocation chromosomique (Tableau 6 des sous-types cytogénétiques), pouvant éventuellement survenir *in utero* en ce qui concerne t(12;21) *ETV6-RUNX1* (Lukes et al., 2018; Schäfer et al., 2018; Wiemels et al., 1999; *et voir section 3.2.6.1.*). En règle générale, les translocations chromosomiques et les réarrangements intrachromosomiques ont lieu précocement au regard de la leucémogénèse et sont considérés comme des événements initiateurs. Ces réarrangements chromosomiques sont habituellement présents dans tous les clones leucémiques au diagnostic et sont conservés à la rechute (Ma et al., 2015; Mullighan et al., 2008). Ces lésions initiatrices conduisent à une dérégulation transcriptionnelle et/ou épigénétique, et confèrent une capacité d'auto-renouvellement aberrant, menant à l'expansion du clone préleucémique. Il est plausible qu'elles n'interfèrent pas avec la différenciation lymphoïde. Ces cellules préleucémiques présentent une susceptibilité accrue à être affectées par d'autres lésions génétiques (Kantner et al., 2013 ; en accord avec le facteur réplicatif de Tomasetti & Vogelstein, 2015, plus une cellule se divise, plus elle risque d'être mutée et accumulent les dommages à l'ADN). Par ailleurs, il existe des variants héréditaires courants (d'*IKZF1*, de *CEBPε* ou d'*ARID5B*) ou plus rarement des mutations germinales hétérozygotes délétères (de *PAX5*, *ETV6* ou *TP53*) qui confèrent une prédisposition génétique aux LAL (Figures 3.11 et 3.12 ; *une section ultérieure abordera les facteurs de risques incluant les susceptibilités héréditaires au développement de LAL-B*).

② Ensuite, une altération ou plusieurs types d'altérations ciblent des facteurs de transcription essentiels pour le développement lymphoïde B (*IKZF1*, *PAX5*, *EBF1*, *ETV6*), provoquant un blocage de la différenciation lymphoïde B au stade pro-B/pré-B, typique des LAL-B (*hallmark*).

③ Enfin, des événements coopératifs additionnels – environ 6 à 8 altérations génétiques additionnelles (Inaba et al., 2013) – s'accumulent dans les clones leucémiques, accélérant la transformation tumorale.

Ces lésions génétiques secondaires récurrentes peuvent notamment être des anomalies du nombre de copies ou des mutations de séquence d'ADN. Elles perturbent de nombreux processus cellulaires clés, ciblant fréquemment des régulateurs du cycle cellulaire (*CDKN2A/CDKN2B*) et des suppresseurs de tumeur (*RB1*, *TP53*), des régulateurs de l'apoptose, des coactivateurs et autres régulateurs transcriptionnels (*ETV6*, *ERG*), des régulateurs épigénétiques (*CREBBP*, *SETD2*, *WHSC1*), des régulateurs de la signalisation lymphoïde (*BTLA*, *TOX*, *CD200*) des molécules de la voie de signalisation RAS (*NRAS*, *KRAS*, *NF1*, *PTPN11*), des récepteurs de facteurs de croissance (*PDGFRβ*), des récepteurs cytokiniques (*CRLF2*) et des kinases (*JAK1/2*, *ABL1*) (Iacobucci & Mullighan, 2017; Kuiper et al., 2007; J. Zhang et al., 2011). Nonobstant, les activations aberrantes des voies JAK/STAT/MYC et RAS/ERK/BCL6 semblent fonctionnellement divergentes dans un contexte de transformation leucémique B (Chan et al., 2020).

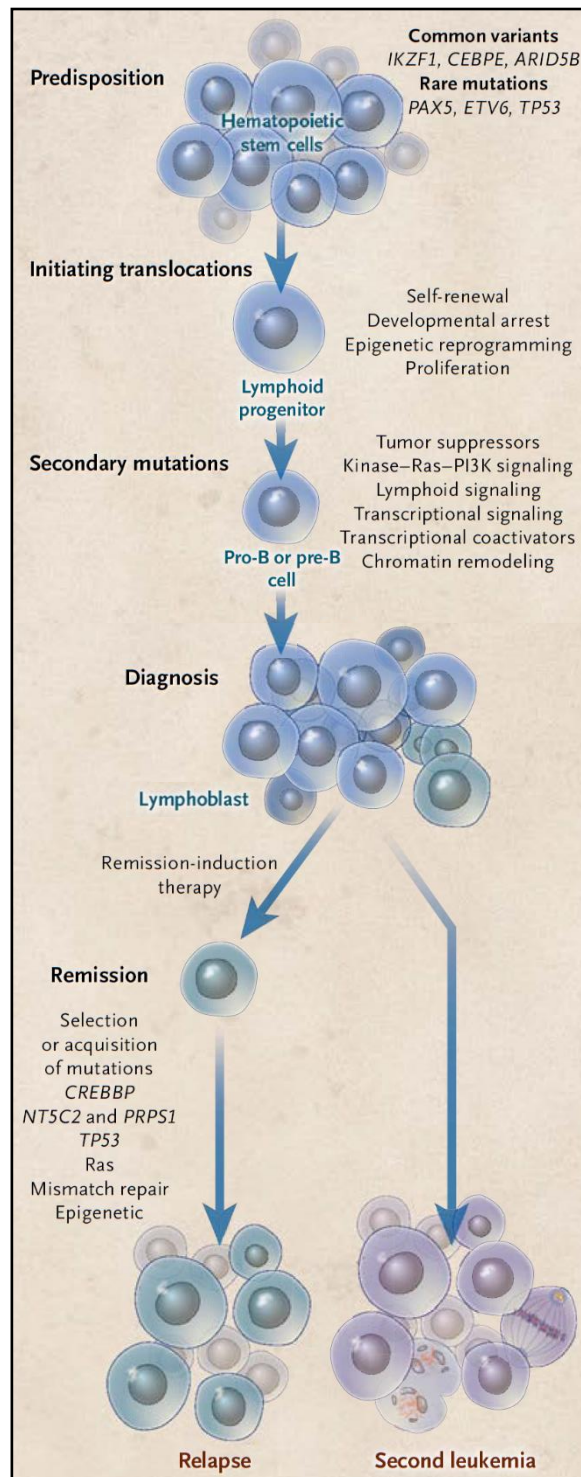


Figure 3.12 : Modèle de leucémogénèse avec emphase sur les possibilités de leucémies secondaires. Adaptation d'une figure issue de Hunger & Mullighan, 2015.

En outre, la perturbation du métabolisme des nucléosides et des modifications épigénétiques est particulièrement fréquent à la rechute (Ma et al., 2015; J. Zhang et al., 2011).

Au diagnostic, les LAL sont généralement oligoclonales ou polyclonales à l'échelle moléculaire (génétiquement). La chimiothérapie d'induction qui aboutit généralement à la rémission du patient induit une pression de sélection en faveur d'un sous-clone (clone mineur), résistant ou dormant par exemple, à l'origine de la rechute. En effet, la chimiothérapie initiale élimine les clones prédominants les plus prolifératifs (clones majeurs formant le *bulk*), mais épargnant les sous-clones ayant acquis des mutations conférant une chimiorésistance (Figure 3.12, côté gauche). Moins fréquemment, les clones de la rechute ne partagent aucune des altérations génétiques acquises par les clones du diagnostic, et établissent une seconde leucémie chez les patients portant une prédisposition génétique (Figure 3.12, côté droit ; Hunger & Mullighan, 2015).

Ce modèle de leucémogénèse B s'appuie essentiellement sur des événements génétiques et épigénétiques, c'est-à-dire intrinsèques à la cellule tumorale. Toutefois, une proportion non négligeable de LAL ne semble pas être assimilée à un des sous-types majeurs de LAL-B (mentionnés en section 3.2). En effet, pour rappel, les algorithmes de profilage transcriptionnel susmentionnés (dont Gu et al., 2019) – discriminant les signatures d'expression génique qui sont le fruit d'altérations génétiques *drivers* associées à une combinaison d'altérations secondaires et à une signalisation cellulaire aberrante – ne parviennent pas à classer la totalité des cas de LAL-B (Tableau 9). Ces LAL-B actuellement inclassables sont regroupées sous la catégorie nommée « *other* », apparaissant sur la Figure 3.06 montrant que les échantillons (représentés par des points gris clair) sont dispersés entre les différents clusters représentant des sous-types génétiques de LAL-B.

Sous-types génétiques	Nombre de cas analysés	Enfants à Risque Standard	Enfants à Risque élevé	Adolescents et Jeunes Adultes	Adultes ≥ 40 ans	Critères
<b>PAX5alt</b>	148	3.3%	9.3%	9.3%	7.7%	<i>Hierarchical gene expression profile cluster enriched with PAX5 alterations</i>
<b>PAX5 P80R</b>	44	0.4%	1.9%	3.1%	4.2%	<i>PAX5 p.Pro80Arg (P80R) mutation or clustered with PAX5 P80R subtype</i>
<b>Other</b>	125	<b>6.4%</b>	<b>4.7%</b>	<b>7.9%</b>	<b>7.1%</b>	

Tableau 9 : Prévalence des LAL-B qui n'ont pas pu être classées d'après leurs profils moléculaires en comparaison avec celles des sous-types PAX5alt et PAX5 P80R, en fonction de l'âge et du risque.

Catégories d'âge : enfants entre 0.2 et 15 ans, adolescents et jeunes adultes (AYAs) de 16 à 39 ans, et adultes de 40 ans à 79 ans. PAX5alt et PAX5 P80R sont des sous-types définis par intégration du profil d'expression génique, du caryotype et/ou des mutations génétiques. *Critères* : Une analyse prédictive de puce à ADN (PAM, Prediction Analysis of Microarrays) a été utilisée pour calculer la similarité du profil d'expression génique avec un/des sous-types de LAL-B spécifique·s. D'après les données de Gu et al., 2019.



Des mécanismes moléculaires distincts des propriétés intrinsèques des cellules leucémiques peuvent également expliquer la leucémogénèse.

### 3.2.5.4. Les mécanismes de la niche favorisent la transformation maligne

Les interactions entre les cellules leucémiques et le microenvironnement médullaire (ou niche stromale hématopoïétique) jouent un rôle dans le développement leucémique et la réponse thérapeutique. En effet, les cellules stromales de la niche hématopoïétique (cellules endothéliales, cellules stromales mésenchymateuses, ostéoblastes, adipocytes, neutrophiles, *et cetera*) peuvent contribuer à la survie et à la chimiorésistance des cellules leucémiques (Méndez-Ferrer et al., 2020 pour revue).

#### 3.2.5.4.1. Des régulateurs extrinsèques de la survie des cellules leucémiques

Par exemple, l'induction de la Protéine Kinase C- $\beta$  par les cellules stromales de la niche hématopoïétique aboutit à la stabilisation de la molécule anti-apoptotique Bcl-xL dans les cellules malignes B, constituant en somme un mécanisme de survie cellulaire et de chimiorésistance (E. Park et al., 2020).

#### 3.2.5.4.2. Des mécanismes de chimiorésistance dépendant de la niche

Les cellules stromales peuvent protéger les cellules leucémiques du stress oxydatif. La chimiothérapie repose souvent sur l'induction excessive d'espèces réactives de l'oxygène (ROS pour *reactive oxygen species*) provoquant des dommages à l'ADN qui nuisent aux cellules cancéreuses. Les cellules stromales mésenchymateuses de la niche sécrètent de la cystéine qui peut être utilisée par les cellules cancéreuses pour synthétiser du glutathion servant à la détoxification des ROS (Boutter et al., 2014).

#### 3.2.5.4.3. Le remodelage du microenvironnement immunitaire

Outre les signaux du microenvironnement, la reprogrammation de la niche médullaire hématopoïétique peut induire le remodelage du compartiment myéloïde de la moelle osseuse dès le début de la maladie (LAL-B), comme illustré par l'abondance d'un sous-type défavorable de monocytes promouvant la survie des cellules leucémiques et ainsi favorisant la rechute au détriment d'une rémission complète (Figure 3.13 ; Witkowski et al., 2020).

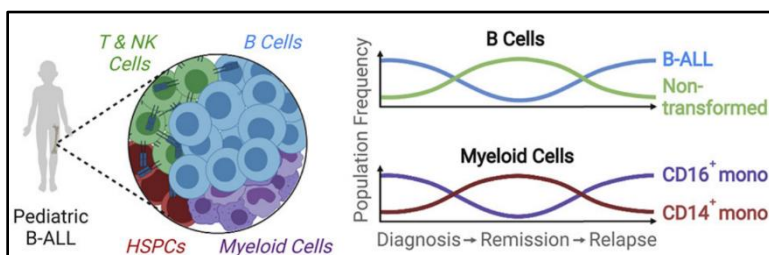


Figure 3.13 issue de Witkowski et al., 2020 (*partial graphical abstract*)



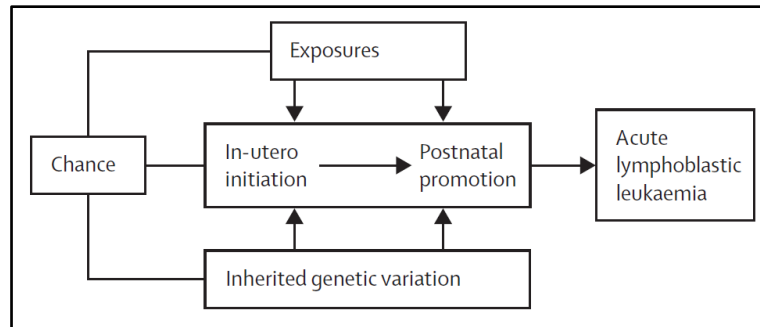


Figure 3.14 : Causalité composite des leucémies aiguës lymphoblastiques pédiatriques.

Les expositions exogènes (infections par exemple) et endogènes (par exemple, inflammation, stress oxydatif), les variants hérités et la chance (occurrence aléatoire) possèdent notamment un rôle dans la leucémogénèse (*the covert natural history*) des LAL pédiatriques. Figure issue de Inaba et al., 2013.

Tableau 10 : Allèles hérités liés au risque de développer une LAL-B pédiatrique, rapportés à partir d'études d'association pangénomiques<sup>a</sup>. Tableau issu de (Greaves, 2018).

Candidate genes (chromosome) <sup>b</sup>	ALL subset bias	OR <sup>c</sup> (95% CI)	Refs
<i>IKZF1</i> (7p12.2)	None detected	1.69 (1.58–1.81)	190
		1.69 (1.40–1.90)	191
<i>ARID5B</i> (10q21.2)	Hyperdiploid	1.65 (1.54–1.36)	190
		1.91 (1.60–2.20)	191
<i>CDKN2A</i> and <i>CDKN2B</i> (9p21.3)	None detected	1.36 (1.18–1.56)	75
		1.72 (1.50–1.97)	192
		2.41 (1.99–2.92) <sup>b</sup>	193
<i>CEBPE</i> (14q11.2)	Hyperdiploid	1.34 (1.22–1.45)	190
		1.45 (1.30–1.62)	194
<i>GATA3</i> (10p14)	<i>BCR-ABL1</i> -like	1.31 (1.21–1.41)	195
		3.85 (2.70–5.40)	196
<i>PIP4K2A</i> (10p12.2)	None detected	1.40 (1.40–1.53)	75,195
<i>TP63</i> (3q28)	<i>ETV6-RUNX1</i>	0.63 (0.52–0.75)	197
<i>PTPRJ</i> (11p11.2)	<i>ETV6-RUNX1</i>	0.72 (0.68–0.89)	197
<i>LHPP</i> (10q26.13)	None detected	1.20 (1.15–1.28)	198
<i>ELK3</i> (12q23.1)	None detected	1.19 (1.12–1.26)	198
Unknown <sup>d</sup> (8q24.1)	None detected	1.34 (1.21–1.47)	199,200
Unknown (2q22.3)	<i>ETV6-RUNX1</i>	1.32 (1.17–1.49)	199
<i>IKZF3</i> (17q12)	None detected	1.18 (1.11–1.25)	200

ALL, acute lymphoblastic leukaemia; NA, not available. <sup>a</sup>Cohort sizes in these studies were between ~400 and 3,500 patients. All patients were white and of European descent, in which the alleles exist at a relatively high frequency (10–50%). <sup>b</sup>More low risk alleles are likely to be uncovered as larger cooperative group studies are pursued and other populations are considered. Rare allelic variants, missed in previous studies, may also be identified that have a stronger impact on risk of ALL as described for *CDKN2A* (OR 2.4)<sup>193</sup>. <sup>c</sup>P values in these studies were between 10<sup>-5</sup> and 10<sup>-19</sup>. <sup>d</sup>The locus at 8q24.1 may regulate expression of *MYC*<sup>199</sup>

Référence 75. Xu, H. et al. Inherited coding variants at the *CDKN2A* locus influence susceptibility to acute lymphoblastic leukaemia in children. *Nat. Commun.* 6, 7553 (2015).  
 Référence 190. Papaemmanuil, E. et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nat. Genet.* 41, 1006–1010 (2009).  
 Référence 191. Trevino, L. R. et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. *Nat. Genet.* 41, 1001–1005 (2009).  
 Référence 192. Hungate, E. A. et al. A variant at 9p21.3 functionally implicates *CDKN2B* in paediatric B cell precursor acute lymphoblastic leukaemia aetiology. *Nat. Commun.* 7, 10635 (2016).  
 Référence 193. Vijaykrishnan, J. et al. The 9p21.3 risk of childhood acute lymphoblastic leukaemia is explained by a rare high-impact variant in *CDKN2A*. *Sci. Rep.* 5, 15065 (2015).  
 Référence 194. Wiemels, J. L. et al. A functional polymorphism in the *CEBPE* gene promoter influences acute lymphoblastic leukemia risk through interaction with the hematopoietic transcription factor Ikaros. *Leukemia* 30, 1194–1197 (2016).  
 Référence 195. Migliorini, G. et al. Variation at 10p12.2 and 10p14 influences risk of childhood B cell acute lymphoblastic leukemia and phenotype. *Blood* 122, 3298–3307 (2013).  
 Référence 196. Perez-Andreu, V. et al. Inherited *GATA3* variants are associated with Ph-like childhood acute lymphoblastic leukemia and risk of relapse. *Nat. Genet.* 45, 1494–1498 (2013).  
 Référence 197. Ellinghaus, E. et al. Identification of germline susceptibility loci in *ETV6-RUNX1*-rearranged childhood acute lymphoblastic leukemia. *Leukemia* 26, 902–909 (2012).  
 Référence 198. Vijaykrishnan, J. et al. A genome-wide association study identifies risk loci for childhood acute lymphoblastic leukemia at 10q26.13 and 12q23.1. *Leukemia* 31, 573–579 (2017).  
 Référence 199. Vijaykrishnan, J. et al. Genome-wide association study identifies susceptibility loci for B-cell childhood acute lymphoblastic leukemia. *Nat. Commun.* 9, 1340 (2018).  
 Référence 200. Wiemels, J. L. et al. GWAS in childhood acute lymphoblastic leukemia reveals novel genetic associations at chromosomes 17q12 and 8q24.21. *Nat. Commun.* 9, 286 (2018).

### 3.2.6. Mécanisme causal des LAL-B et facteurs de risque

À l'image de la leucémogénèse, le mécanisme causal des LAL-B pédiatriques s'avère **multifactoriel** (Greaves, 2018), impliquant notamment les expositions aux infections, la prédisposition génétique, le régime alimentaire et une part de (mal)chance (Figure 3.14).

#### 3.2.6.1. Susceptibilité génétique

Peu de facteurs de risque environnementaux sont associés aux LAL-B pédiatriques et ils ne concernent qu'une minorité de cas. En revanche, il existe des facteurs de susceptibilité génétique associés à un risque accru de développement de LAL-B. Les types courants de variations génétiques héréditaires comprennent les polymorphismes mononucléotidiques (*single nucleotide polymorphisms, SNP*), les insertions et les suppressions (gain ou perte de courts segments de séquence, indels) et les variations structurales (gain ou perte de grands segments de séquence ; par exemple, variations du nombre de copies, *CNV*).

Les facteurs de prédisposition génétique des LAL pédiatriques peuvent être divisés en plusieurs catégories et comprennent notamment :

- Association avec des syndromes de prédisposition au cancer
  - le syndrome de Down qui est la conséquence d'une trisomie 21 constitutive (Whitlock, 2006) ;
  - plus de 40% des LAL-B pédiatriques hypodiploïdes présentent des mutations germinales de *TP53* (souvent bialléliques) caractéristiques du syndrome de Li-Fraumeni (Holmfeldt et al., 2013).

- Variants hérités courants, à pénétrance faible à intermédiaire (Tableau 10)

Les études pangénomiques d'association (*Genome-Wide Association Studies*) permettent d'identifier un nombre croissant de polymorphismes génétiques germinaux (variants hérités) contribuant de manière non-négligeable au risque de développer une LAL pédiatrique (Tableau 10 ; pour revues (Moriyama, Relling, et al., 2015; Vijayakrishnan et al., 2018, 2019; Vijayakrishnan & Houlston, 2010). Il s'agit d'allèles à risque relativement courants. Les variants polymorphiques hérités affectent notamment *ARID5B* (en association avec les LAL-B hyperdiploïdes) et *IKZF1* (induisant la diminution de son expression) (Papaemmanuil et al., 2009; Treviño et al., 2009), identifiant également la variation de *CEBPε* associée aux LAL-T), ainsi que *CDKN2A* (Sherborne et al., 2010; H. Xu et al., 2013). Les variants hérités de *GATA3* sont associés aux LAL-B pédiatriques *Ph-like* (Perez-Andreu et al., 2013). Les autres gènes avec des polymorphismes nucléotidiques associés à un risque accru de LAL sont *CDKN2B*, *PIP4K2A* et *TP63* (Moriyama, Relling, et al., 2015). Ces facteurs de prédisposition génétique sont courants et un individu peut en porter plusieurs. Ils sont additifs plutôt que synergiques en termes de probabilité de développer une LAL.





- Rares variants germinaux délétères, à pénétrance intermédiaire à forte

De rares mutations germinales de *PAX5*, réduisant son activité (Shah et al., 2013), et de *ETV6*, induisant une perte de fonction (Moriyama, Metzger, et al., 2015; Nishii et al., 2020; Noetzli et al., 2015; M. Y. Zhang et al., 2015) prédisposent aux LAL-B familiales. Les variants germinaux pathogéniques de *TP53* sont également associés à un risque accru de LAL (Qian et al., 2018) et la plupart des mutations germinales d'*IKZF1* nuisent à sa fonction (Churchman et al., 2018).

- Variations du nombre de copies

La survenue d'une translocation Robertsonienne rob(15 ; 21) germinale ou la formation germinale d'un chromosome 21 en anneau prédisposent spécifiquement et fortement aux LAL-B à amplifications intrachromosomiques du chromosome 21 (iAMP21 ; Harrison & Schwab, 2016).

- Translocations chromosomiques *in utero*

En revanche, une proportion relativement élevée d'enfants naît avec de potentiels clones préleucémiques, comportant une translocation chromosomique considérée comme une lésion initiatrice de leucémogénèse :

- La fusion *ETV6-RUNX1* a lieu *in utero* dans 2 à 5% des nouveau-nés sains (Schäfer et al., 2018) dont 1% développeront ultérieurement des leucémies, et la fusion *TCF3-PBX1* se produirait *in utero* chez 0.6% des nouveau-nés (Hein et al., 2019).
- Des clones préleucémiques comportant l'oncogène de fusion *BCR-ABL* (p190) ou l'oncogène de fusion *KMT2A-AF4* sont également fréquemment détectés dans le sang de cordon ombilical des nouveau-nés (G. Cazzaniga et al., 2011; Kosik et al., 2017).

➔ Les clones préleucémiques restent le plus souvent silencieux, à l'état latent, et ne se convertissent pas en clones leucémiques, expliquant pourquoi la majorité de ces individus asymptomatiques (porteurs sains) ne développent pas de LAL-B.

### 3.2.6.2. Rôle causal des infections courantes

➔ Les infections courantes, virales et/ou bactériennes, sont considérées comme l'élément déclencheur le plus probable du développement des LAL-B pédiatriques (G. Cazzaniga et al., 2017; Francis et al., 2012; Heath & Hasterlik, 1990).

En effet, selon le modèle proposé par Sir Melvyn Francis Greaves, les LAL s'établiraient en deux étapes distinctes à l'échelle de l'individu (« *two-hit model* », Figure 3.15 ; Greaves, 2018).

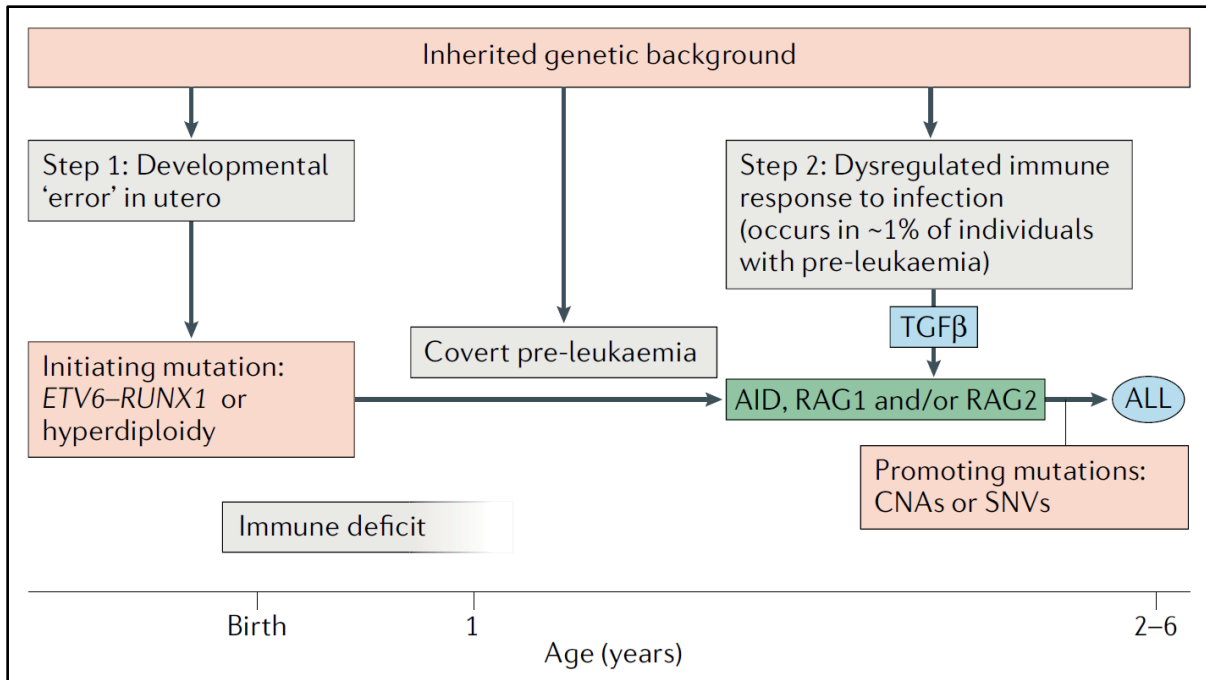


Figure 3.15 : Résumé du modèle de développement des LAL-B pédiatriques en deux étapes (*two-hit model*) de Sir Melvyn Francis Greaves

Figure issue de (Greaves, 2018).

1<sup>ère</sup> étape : initiation par une lésion prénatale

Le premier événement se produirait avant la naissance et correspondrait à la génération d'un clone préleucémique latent initiée par la survenue d'une altération génétique telle que la translocation de fusion *ETV6-RUNX1* ou une hyperdiploïdie par exemple (Taub et al., 2002) (et références susmentionnées dans la section précédente), rendant l'enfant susceptible de développer une leucémie sans que ça ne puisse lui arriver avant l'acquisition postnatale de lésions secondaires coopératives.

2<sup>nde</sup> étape : conversion par une infection postnatale

Puis dans une petite fraction (1%) de ces individus (enfants) génétiquement prédisposés, un second événement se produit après la naissance dans une population de cellules préleucémiques. Selon l'hypothèse de « l'infection retardée » proposée par le Pr. Greaves (***the Greaves delayed infection hypothesis***), les LAL-B pédiatriques surviendraient à la suite d'une réaction immunitaire aberrante en réponse à une ou plusieurs infections courantes (virales et/ou bactériennes) dont l'exposition aurait anormalement lieu tardivement. L'exposition à ces agents pathogènes courants serait ainsi dite retardée (« *delayed* ») en l'absence de suffisamment d'épisodes infectieux dans la toute-petite enfance (tout-petits de 0 à 3 ans ; (Greaves, 2018). Cette hypothèse expliquerait notamment le pic d'incidence des LAL-B pédiatriques entre 2 et 5 ans. Cette réponse immunitaire anormale déclencherait indirectement l'acquisition de mutations secondaires requises pour l'évolution clonale et la transformation maligne conduisant à l'émergence des signes cliniques des LAL.

Plusieurs preuves expérimentales étayent l'hypothèse du Professeur Greaves : les souris présentant une susceptibilité génétique aux LAL-B (exprimant *ETV6-RUNX1* ou *Pax5* à l'état hétérozygote) assimilée à un stade préleucémique, développent un phénotype de LAL-B à condition d'être exposées à des agents pathogènes courants (Martín-Lorenzo et al., 2015; Rodríguez-Hernández et al., 2019). Fonctionnellement, ces lésions initiatrices présentes dès la naissance (*Pax5*<sup>+/-</sup>, *BCR-ABL1 p190* combiné à *Pax5*<sup>+/-</sup>, *ETV6-RUNX1*) peuvent générer un compartiment progéniteur B aberrant (Kantner et al., 2013; Martín-Lorenzo et al., 2015, 2018). Mécanistiquement, les infections provoquant une réaction inflammatoire (la « tempête cytokinique »), la surproduction de cytokines pro-inflammatoires IL6/TNF $\alpha$ /IL $\beta$  en coopération avec les cellules stromales de la niche inflammée favoriserait l'émergence des cellules préleucémiques exprimant le transcrit de fusion *ETV6-RUNX1* (Beneforti et al., 2020). La niche inflammée contribuerait à exacerber l'accumulation de dommages à l'ADN (Beneforti et al., 2020), d'ores et déjà induite par un stress oxydatif endogène (excès de ROS) dans les cellules pro-B préleucémiques avec *ETV6-RUNX1* (Kantner et al., 2013). Cette instabilité génétique favoriserait alors l'accumulation de mutations délétères coopératives additionnelles entraînant *de facto* la transformation maligne des cellules préleucémiques exprimant l'altération génétique princeps.



En revanche, si l'exposition postnatale à des agents microbiens est effectivement plus précoce (tout-petits de moins de 2 ans, à la crèche par exemple, avant le pic d'incidence vers 2-5 ans), elle est considérée comme protectrice contre le risque de développer une LAL (Greaves, 2018). Fréquenter la garderie serait ainsi une action préventive (Gilham et al., 2005; Urayama et al., 2010).

### 3.2.6.3. Le mode de vie

Des facteurs de risque généralistes, applicables à tout type de cancers, interviennent au second plan dans le développement multi-étapes des LAL-B. Outre le patron temporel d'exposition aux infections courantes, **le régime ou la prise alimentaire**, une autre composante du mode de vie, pourrait influencer sur la leucémogénèse. Un régime restreint pourrait réduire les risques de développement tumoral. En effet, le jeûne intermittent pourrait être bénéfique pour bloquer l'expansion leucémique, comme démontré chez la souris (Lu et al., 2017).

De surcroît, des prédispositions génétiques aux LAL-B, comme l'expression de *PAX5* à l'état hétérozygote ou la fusion *ETV6-RUNX1*, induisent intrinsèquement une dysbiose du microbiome intestinal (remodelé et altéré) chez la souris, et l'absence de microbiote commensal (à l'instar d'un manque d'exposition aux agents infectieux précédemment énoncé) promeut le développement leucémique dans de telles prédispositions même en l'absence de stimuli infectieux (Vicente Dueñas et al., 2020). Par déduction, de telles prédispositions génétiques pourraient ainsi potentialiser la réponse immunitaire inappropriée susmentionnée. En somme, un mode de vie trop prudent et/ou déséquilibré (sous-exposition aux agents infectieux courant pendant la petite enfance, excès d'antibiotiques et/ou malnutrition induisant un appauvrissement de la flore commensale) pourrait contribuer à révéler une prédisposition génétique sous-jacente et promouvoir la transformation leucémique.

### 3.2.6.4. Le facteur (mal)chance

La causalité du cancer est étroitement liée au **hasard** (Greaves M., 2000), par exemple :

- *Événements fortuits à la conception impliquant le brassage et la recombinaison des gènes parentaux* : il s'agit de la chance, par définition aléatoire, d'acquérir *in utero* un événement de prédisposition génétique (tel qu'une translocation de fusion), d'hériter d'un ou plusieurs allèles à risque, d'hériter ou d'acquérir *de novo* une mutation germinale rendant susceptible de développer une LAL.
- *Impact moléculaire et cellulaire* : il s'agit de la chance que les dommages à l'ADN affectent un oncogène pertinent dans une cellule pertinente (cellule souche ou cellule progénitrice).
- *Exposition externe accidentelle* : il s'agit donc de la chance d'être au mauvais endroit au mauvais moment (ou de ne pas être au bon endroit au bon moment) quant à l'exposition accidentelle aux agents infectieux qui ont le potentiel de dévoiler un clone préleucémique latent et d'aggraver la maladie.



# Partie II : Objectifs et Résultats

## Objectifs de la thèse

Mon travail de thèse s'inscrit dans la continuité de mon stage de seconde année de master en filière recherche (M2R mention biologie-santé, spécialité cancérologie) effectué dans le laboratoire d'accueil U1037– le Centre de Recherche en Cancérologie de Toulouse (CRCT) – au sein de l'équipe du Professeur Éric Delabesse (équipe 16 du CRCT).

Mes travaux de thèse reflètent la thématique générale de notre équipe qui étudie l'impact des **altérations de facteurs de transcription dans les leucémies aiguës** dans un contexte de recherche translationnelle. Notre équipe focalise notamment ses recherches sur les rôles de deux facteurs de transcription essentiels de l'hématopoïèse : PAX5 et GATA2, dont les gènes sont fréquemment affectés par diverses altérations, identifiées chez des patients atteints d'hémopathies malignes.

**Mon travail de thèse s'est principalement articulé autour du rôle du facteur de transcription PAX5 au cours de la différenciation lymphoïde B, dans un contexte physiologique ou pathologique.** Je me suis particulièrement attachée à **étudier les propriétés et modes d'action de différents mutants de structure de PAX5** (l'oncoprotéine de fusion PAX5-ELN et le mutant germinal PAX5<sup>R38H</sup>) aussi bien *in vitro* qu'*in vivo*. Dans ce but, j'ai eu l'opportunité de caractériser un modèle murin original exprimant PAX5-ELN, modèle élaboré avant mon arrivée dans le laboratoire.

L'expertise acquise lors de ce travail m'a également permis de participer activement à l'analyse phénotypique et fonctionnelle d'un modèle murin original exprimant le mutant GATA2<sup>R396Q</sup>, des résultats que je ne développerai pas dans ce manuscrit de thèse (manuscrit d'article en préparation pour publication).

### 4.1. Comparaison de l'expression et des fonctions des isoformes A et B de Pax5

Je me suis appliquée à comparer le rôle physiologique des deux isoformes principales de PAX5 : PAX5A (également appelée BSAP pour B-cell activator protein) et PAX5B dans la continuité des travaux de thèse de Charlotte Cresson qui portaient sur la régulation et la fonction des isoformes PAX5A et PAX5B. Ces investigations ont été menées dans le cadre d'une collaboration avec le Dr Sophie Péron de l'équipe du Pr. Michel Cogné à l'Université de Limoges. Bien que l'isoforme PAX5A s'exprime spécifiquement dans le tissu lymphoïde B, l'isoforme PAX5B y est aussi exprimée.



Comme de nombreux facteurs de transcription essentiels de l'hématopoïèse, l'expression de PAX5 nécessite d'être finement contrôlée, ce qui justifierait la co-expression des deux isoformes, qui ne diffèrent structurellement que par leur premier exon respectif.

Nous avons démontré que les deux isoformes PAX5A et PAX5B peuvent induire un programme d'expression génique de différenciation lymphoïde B et assurent des fonctions équivalentes durant la différenciation B précoce mais présentaient des disparités à un stade cellulaire B plus tardif, et en termes de croissance cellulaire (Cresson et al., 2018).

## 4.2. Caractérisation du modèle murin *in vivo* de LAL-B multi-étapes induite par l'oncoprotéine de fusion PAX5-ELN

Les causes moléculaires de LAL-B sont identifiées en grande partie mais leurs modes d'action sont encore peu décrits *in vivo*. Pour rappel, les altérations somatiques hétérozygotes de *PAX5* surviennent dans un tiers des cas de LAL-B, parmi lesquelles les réarrangements chromosomiques de *PAX5* ont lieu dans 2 à 3% des cas de LAL-B (Coyaud, Struski, Prade, et al., 2010; Familiades et al., 2009; Mullighan et al., 2007). La création de modèles animaux exprimant le produit de ces altérations représente un moyen idéal pour élucider le rôle oncogénique des protéines de fusion PAX5-X dans la leucémogénèse, ainsi que les mécanismes moléculaires sous-jacents.

La chimère de fusion PAX5-ELN résulte d'une translocation chromosomique récurrente, notamment identifiée chez deux patients à Toulouse (Bousquet et al., 2007). Un modèle murin original exprimant le transgène *PAX5-ELN* à partir du locus *IgH* sous le contrôle de l'*enhancer* E $\mu$  a été élaboré dans le cadre d'une collaboration avec l'équipe du Dr Ahmed Khamlichi (IPBS, Toulouse).

**Mes objectifs étaient de :**

- A. Caractériser le nouveau modèle murin exprimant la protéine de fusion PAX5-ELN**
- B. Déterminer le mode d'action de la protéine de fusion PAX5-ELN**

J'ai montré que les souris transgéniques exprimant la protéine de fusion PAX5-ELN représentent un robuste modèle *in vivo* de leucémogénèse multi-étapes. Mon travail a démontré que PAX5-ELN perturbait *in vitro* et *in vivo* le développement du lignage lymphoïde B et induisait une expansion aberrante du compartiment cellulaire pro-B au stade préleucémique qui est suivi par une transformation leucémique associée à des mutations secondaires récurrentes des voies de signalisation Ras et Jak/Stat (Jamrog et al., 2018).

### 4.3. Étude fonctionnelle *in vitro* et *in vivo* de l'effet du mutant $PAX5^{R38H}$

J'ai été amenée à collaborer avec le Dr Nicolas Duployez et le Pr. Claude Preudhomme dirigeant le laboratoire d'hématologie du CHU de Lille.

Suite à la prise en charge au CHU de Lille de patients d'une même fratrie atteints de LAL-B, ils ont identifié une mutation germinale hétérozygote de  $PAX5$  (c.113G>A, p.Arg38His) déjà décrite dans une famille non-apparentée (Yazdanparast et al., 2015), établissant ainsi la récurrence de cette mutation ponctuelle germinale de  $PAX5$  prédisposant à la LAL-B.

Dans le but d'enrichir ces études de LAL-B familiales, j'ai étudié l'effet intrinsèque du mutant  $PAX5^{R38H}$ , adressé à la fois *in vitro* et *in vivo*. Mes résultats ont démontré que  $PAX5^{R38H}$  agissait comme un variant hypomorphique de  $PAX5$ , inhibant les propriétés fonctionnelles de  $PAX5$  dans la différenciation du lignage lymphoïde B. De plus, j'ai validé le rôle de  $PAX5^{R38H}$  dans la susceptibilité au développement leucémique puisque son expression est capable d'induire un phénotype de LAL-B chez la souris (Duployez et al., 2021).

### 4.4. Caractérisation du modèle murin *in vivo* exprimant $Gata2^{R396Q/+}$

Dans un second temps, j'ai participé à la poursuite des travaux de thèse du Pr. Marlène Pasquet sur les effets du variant  $GATA2^{R396Q}$ . Ce projet de recherche a été initié suite à l'identification de cette mutation germinale de  $GATA2$  dans des cas de SMD/LAM (syndromes myélodysplasiques / leucémies aiguës myéloïdes) notamment dans une famille de patients soignés à Toulouse (Pasquet et al., 2013). Pour évaluer le rôle de  $GATA2^{R396Q}$  dans la dérégulation pathologique de l'hématopoïèse, le Dr Cyril Broccardo a élaboré un modèle murin innovant *knock-in* de la mutation germinale faux-sens  $GATA2^{R396Q}$ . En étroite collaboration avec le Dr Laëtitia Largeaud (PharmD, doctorante en sciences), j'ai contribué à caractériser phénotypiquement et fonctionnellement ce nouveau modèle murin qui présente un déséquilibre quantitatif des sous-populations de progéniteurs multipotents et cellules souches hématopoïétiques (CSH) dans la moelle osseuse ainsi qu'un immunophénotype aberrant des CSH. Nous avons également mis en évidence un déficit fonctionnel des LSK incluant les CSH. Je ne développerai pas cette partie dans ce manuscrit (Largeaud et al., 2021).

## **Publications originales des travaux de thèse collaboratifs**

Mes travaux de thèse collaboratifs ont donné lieu à trois publications scientifiques originales en tant que co-première auteur (deux articles de recherche et une lettre à l'éditeur).

Cresson C, Péron S, Jamrog L *et al.*, *Oncotarget*, 2018

Jamrog L, Duchemin G *et al.*, *PNAS*, 2018

Duployez N, Jamrog LA *et al.*, *Blood*, 2021

## **PAX5A and PAX5B isoforms are both efficient to drive B cell differentiation**

**Charlotte Cresson<sup>1,\*</sup>, Sophie Péron<sup>2,\*</sup>, Laura Jamrog<sup>1,\*</sup>, Nelly Rouquié<sup>1</sup>, Nais Prade<sup>3</sup>, Marine Dubois<sup>1</sup>, Sylvie Hébrard<sup>1</sup>, Stéphanie Lagarde<sup>3</sup>, Bastien Gerby<sup>1</sup>, Stéphane J.C. Mancini<sup>4</sup>, Michel Cogné<sup>5</sup>, Eric Delabesse<sup>3</sup>, Laurent Delpy<sup>2</sup> and Cyril Broccardo<sup>1</sup>**

<sup>1</sup>Inserm, UMR1037 CRCT, F-31000, Université Toulouse III-Paul Sabatier, UMR1037 CRCT, Oncopole, F-31000 Toulouse, France

<sup>2</sup>Université de Limoges-CNRS UMR 7276, F-87025 Limoges, France

<sup>3</sup>Inserm, UMR1037 CRCT, F-31000, Université Toulouse III-Paul Sabatier, UMR1037 CRCT, Toulouse Hospital University, Oncopole, CS 53717, F-31000 Toulouse, France

<sup>4</sup>Aix Marseille Univ, CNRS, INSERM, Institut Paoli-Calmettes, CRCM, F-13009 Marseille, France

<sup>5</sup>Université de Limoges-CNRS UMR 7276, Institut Universitaire de France, F-87025 Limoges, France

\*These authors contributed equally to this work

**Correspondence to:** Cyril Broccardo, **email:** [cyril.broccardo@inserm.fr](mailto:cyril.broccardo@inserm.fr)

**Keywords:** Pax5; transcription factor; B cell development; B cells; acute lymphoblastic leukemia

**Received:** August 11, 2016

**Accepted:** July 31, 2018

**Published:** August 28, 2018

**Copyright:** Cresson et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

### **ABSTRACT**

**Pax5 is the guardian of the B cell identity since it primes or enhances the expression of B cell specific genes and concomitantly represses the expression of B cell inappropriate genes. The tight regulation of Pax5 is therefore required for an efficient B cell differentiation. A defect in its dosage can translate into immunodeficiency or malignant disorders such as leukemia or lymphoma.**

**Pax5 is expressed from two different promoters encoding two isoforms that only differ in the sequence of their first alternative exon. Very little is known regarding the role of the two isoforms during B cell differentiation and the regulation of their expression. Our work aims to characterize the mechanisms of regulation of the expression balance of these two isoforms and their implication in the B cell differentiation process using murine *ex vivo* analyses. We show that these two isoforms are differentially regulated but have equivalent function during early B cell differentiation and may have functional differences after B cell activation. The tight control of their expression may thus reflect a way to finely tune Pax5 dosage during B cell differentiation process.**

### **INTRODUCTION**

The commitment of hematopoietic stem cells to each cell lineage is strictly controlled. Specific extracellular stimuli and transcription factors play important roles in the development of B cells from hematopoietic stem cells. The transcription factors SPI1 (PU.1) and IKZF1 (Ikaros) work together during the early developmental stages, while TCF3 (E2A), Early B cell Factor 1 (EBF1) and Paired box 5 (PAX5) are crucial from pro-B cells to mature B cell stages

[1]. PAX5 is one of the 9 members (PAX1 to PAX9) of the highly conserved paired-box (PAX) domain family of transcription factors [2, 3] characterized by a conserved PAIRED DNA-binding domain at the N-terminal part of the protein and involved in the regulation of tissue homeostasis.

During hematopoiesis, PAX5 is expressed from the pro-B cell stage and has to be turned off to allow plasma-cell transition [4]. PAX5 is crucial for the maintenance of the B lymphoid lineage identity [5, 6] and for suppression of alternative lineage choices [1, 7].

PAX5 also enhances the transcription of B cell specific genes and participates in the chromatin-remodeling of the immunoglobulin heavy chain (IGH) locus, ensuring its contraction during VDJ recombination [8]. At later stages, PAX5 regulates the IGH 3' regulatory region (3'RR). The 3'RR is a 30 kb-long cis-acting regulation element of the immunoglobulin heavy chain (IGH) locus containing four enhancers in mice (hs1,2, hs3a, hs3b and hs4) with a strict B lineage specificity. They have been implicated in the late stages of B cell differentiation with a crucial role in class switch recombination (CSR) and somatic hypermutation (SHM) [9–12].

*Pax5* homozygous inactivation in mouse leads to a blockade at the pro-B cell stage [6]. *Pax5*-deficient pro-B cells are able to transdifferentiate into other cell types, such as T, natural killer, and dendritic cells [1, 13] while mature B cells can dedifferentiate into lymphoid precursor cells upon *Pax5* loss even at late stages of B cell differentiation as shown by *Pax5* conditional inactivation [14].

In vertebrates, PAX5 expression is controlled by two distinct promoters: a distal P1a and a proximal P1b [15] which initiate transcription from two alternative 5' first exons (exons 1A and 1B respectively) leading to the expression of two isoforms, PAX5A and PAX5B. The only structural difference between PAX5A and PAX5B consists in a short different N-terminal sequence. Exon 1A encodes 15 amino acids that differ totally from the 14 amino acids encoded by the exon 1B sequence [16]. Despite high sequence similarities between these two isoforms, they show different patterns of expression: PAX5B is transcribed in B cells, central nervous system and testis, while PAX5A, whose product is also named B cell-specific activator protein (BSAP) is restricted to the B cell lymphoid lineage [4]. However their respective roles in B cell differentiation have never been determined. Our work aims to detail the mechanisms that control the expression of PAX5A and PAX5B isoforms along B cell development and their effect on B cell differentiation.

## RESULTS

### ***Pax5* expression in B cell differentiation is independent of adjacent genes**

The murine *Pax5* gene encompasses a region of 392 kb of chromosome 4 from the end of its upstream neighbor gene, *Melk*, to the start of its downstream neighbor gene, *Zcchc7* (Figure 1A). *Pax5* has a reverse orientation compared to its two neighbors, from telomere to centromere (Figure 1A). The human PAX5 gene has a similar organization covering a slightly larger region of 444 kb on chromosome 9. In order to clarify the transcriptional activities within the *Pax5* locus, quantitative RT-PCR (QPCR) was performed to measure the overall expression of *Pax5* transcripts and its neighboring genes (*Melk* and *Zcchc7*), together with *Abll* as a widely expressed control

gene and *Ebfl* as a transcriptional target of *Pax5*. Their expression were measured in a series of murine B cell lines representing different stages of B cell differentiation (from the less to the most differentiated: Ba/F3, 70Z3, 38B9, 18.81, A20 and WEHI-231) along with murine primary tissues (T and B cells, Figure 1B). Since *Pax5* expression is regulated by *Ebfl*, *Pax5* expression is highly correlated to the expression of *Ebfl*, as expected (Pearson correlation,  $r^2 = 0.85$ ). In contrast, expression of *Pax5* is independent of the expression of its two neighboring genes, *Melk* and *Zcchc7* (Pearson correlation,  $r^2 = 0.40$  and  $r^2 = 0.54$  respectively), suggesting that the regulatory elements of *Pax5* are not shared by *Melk* and *Zcchc7*.

### ***Pax5* isoforms are differentially expressed during B cell differentiation**

Two major 5' isoforms of *Pax5* are expressed during B cell differentiation. *Pax5A* expression is driven by the promoter 1A and *Pax5B* by the promoter 1B [using alternative first exons (1A and 1B respectively, Figure 1A)]. We detailed the expression pattern of the two *Pax5* isoforms during murine B cell differentiation using specific primers of these two isoforms on sorted B cell subsets from bone marrow. *Pax5B* has a low expression which does not vary during B cell differentiation. In contrast, *Pax5A* expression is strongly modulated during B cell differentiation with a higher expression in immature B cells (Figure 2A, left panel).

### **Expression of *Pax5* isoforms is correlated with the histone H3 acetylation of their respective promoter**

We defined the expression pattern of *Pax5* isoforms on B cell lines. The 38B9, 18.81, A20 cell lines expressed both *Pax5* isoforms. The early blocked 70Z3 B cell line expressed only *Pax5A* isoform. The Ba/F3 cell line did not express any of the two isoforms (Figure 2A, right panel).

Using the fact that B cell lines differentially expressed *Pax5* isoforms, we detailed the transcriptional regulation within the 392 kb genomic region based on the acetylation of histone H3, a major mark of active regulatory elements [17] at low resolution over the 50 most conserved regions in vertebrates. We observed a peak of acetylation covering *Pax5* promoters A and B in pro-B cells and the B cell lines expressing at least *Pax5A* in contrast to Ba/F3 (Figure 2B). We increased the resolution of the ChIP analysis to 1 kb over this region on five B cell lines (Figure 2C). The three cell lines expressing both *Pax5* isoforms (38B9, 18.81 and A20) exhibited two acetylation peaks located upstream of exon 1A (defined as the location origin, 0 kb) and 1B (+7 kb). 70Z3 expressing only *Pax5A* isoform had peaks at 0 kb and -4kb, but none at the +7 kb while Ba/F3 displayed no acetylation peak (Figure 2C).

Decker *et al.* identified an enhancer region located in the intron 5–6 of *Pax5* [24]. We analyzed the acetylation status of this region and showed that the acetylation is weak compared to the promoter regions. Interestingly, only the cell lines expressing both isoforms have an H3 acetylation mark on the enhancer region (Figure 2D).

### Both Pax5 isoforms restore early B cell differentiation

We took advantage of the *Pax5*<sup>-/-</sup> cells which do not express any of the two isoforms and are subsequently blocked at a pro-B cell stage but still proliferating in presence of IL7. *PAX5* isoforms were individually transduced in *Pax5*<sup>-/-</sup> pro-B cells and assessed for their ability to reinitiate B cell differentiation, using the expression of *Cd19*, a specific target of PAX5, and the appearance of IgM<sup>pos</sup>IgK<sup>pos</sup> B cells as hallmarks of B cell progression. Cells infected with the empty vector (MIE, only expressing eGFP) did not induce expression of *Cd19*. In contrast, both isoforms induced the expression of surface *Cd19* in presence or absence of IL7 (Figure 3A, respectively upper and lower panels). Furthermore, both isoforms were also able to initiate the expression of IgM and IgK at the cell surface after removal of IL7 (Figure 3B, lower panel).

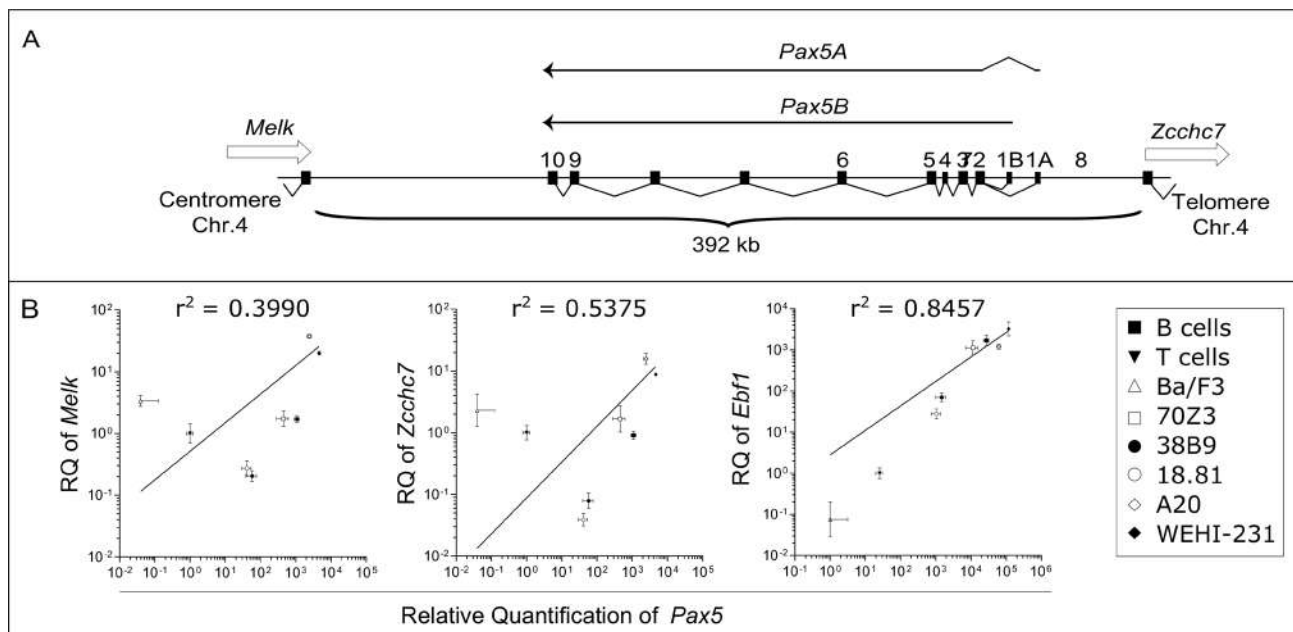
Transduced *Pax5*<sup>-/-</sup> cells were seeded at the same concentration and cultured for 3 days in presence of IL7. Although the percentage of *Cd19*<sup>pos</sup> cells is the same after

transduction with *PAX5A* or *PAX5B* (Figure 3C, left panel), the absolute number of pro-B cells transduced by *PAX5* isoforms was increased compared to the MIE with the most potent proliferation being driven by *PAX5B* (Figure 3C, right panel). Furthermore, after 10 days of culture in presence of IL7, *Pax5*<sup>-/-</sup> pro-B cells complemented with *PAX5B* have a quarter less in spontaneous cell death rate compared with *PAX5A* (Figure 3D, left panel). This difference of cell survival may explain the highest percentage of IgM<sup>pos</sup>IgK<sup>pos</sup> B cells after IL7 withdrawal with *PAX5B* complementation (Figure 3B).

We also stably transduced Ba/F3 cells with either MIE, *PAX5A* or *PAX5B* and stressed the cells by serum deprivation for 6 hours. Ba/F3 cells are less sensitive to serum deprivation when they express *PAX5B* compared to *PAX5A* whereas no difference was observed in 10% serum condition (Figure 3D, right panel). This difference is not linked to a difference of *PAX5* isoform stability since half-life of the two proteins are comparable (Figure 3E). Consequently, the alternative use of *PAX5* exon 1 has no qualitative effect on early B cell differentiation but may have an impact on cell death.

### Both Pax5 isoforms induce a B cell specific program

After sorting transduced *Pax5*<sup>-/-</sup> pro-B cells, their respective transcriptomes were analyzed (Figure 4A–4C).



**Figure 1: Expression of *Pax5* isoforms is independent of the expression of neighboring genes.** (A) Schematic organization of the genomic region of murine *Pax5* gene. *Pax5* is composed of 11 exons, the first two (exons 1A and 1B) being alternatively used to generate two isoforms (*Pax5A* and *Pax5B* respectively). *Pax5* gene is flanked by *Melk* and *Zcchc7* genes. (B) Correlation between *Melk*, *Zcchc7* or *Ebf1* expression and *Pax5* expression. Quantitative PCR (QPCR) was performed at least as triplicate on Ba/F3, 70Z3, 38B9, 18.81, A20 and WEHI-231 cell lines and on B and T cells. Relative expressions (RQ) to *Abi1* expression are expressed as mean with error bars representing RQMIN and RQMAX and constitute the acceptable error level for a 95% confidence interval according to Student's *t* test. The square of the Pearson correlation ( $r^2$ ) is indicated for each comparison.

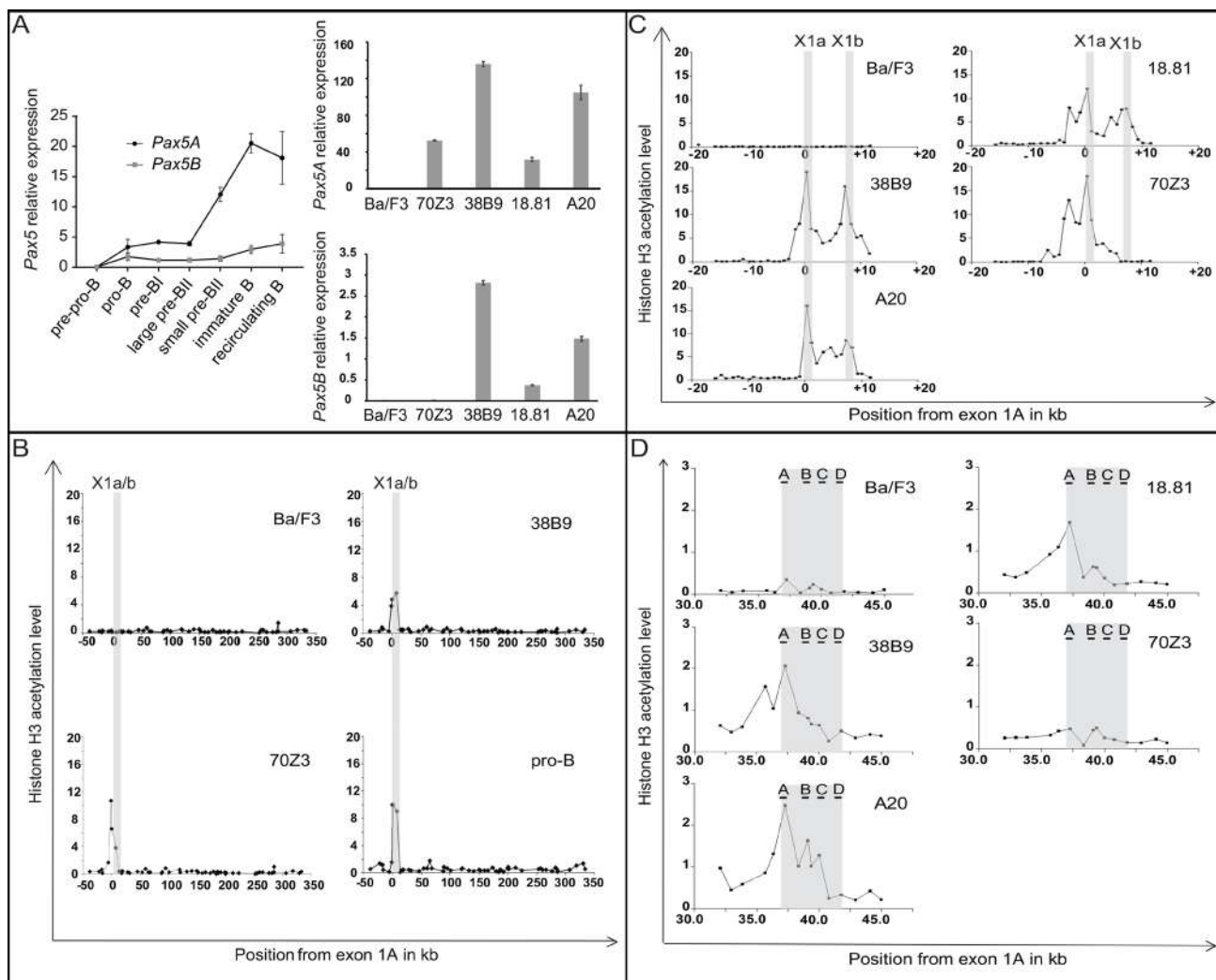


Both PAX5A and PAX5B enhance the expression of B cell specific genes such as *Cd19* and repress the expression of non-B cell specific markers such as *Cd7* (a T cell marker) or *Gp49a* (a myeloid marker) (Figure 4D). Furthermore, out of the 24,184 genes, no expression differs more than 5-fold between PAX5A and PAX5B target genes (Figure 4C). By performing a Gene Ontology (GO) analysis [18, 19] on the genes whose expression is the most discriminant, we found a significant enrichment in genes implicated in the regulation of apoptosis (Supplementary Figure 1A, 1B). These data strongly suggest that the two PAX5 isoforms globally induce B cell differentiation by activating the same targets during early B cell differentiation and differences

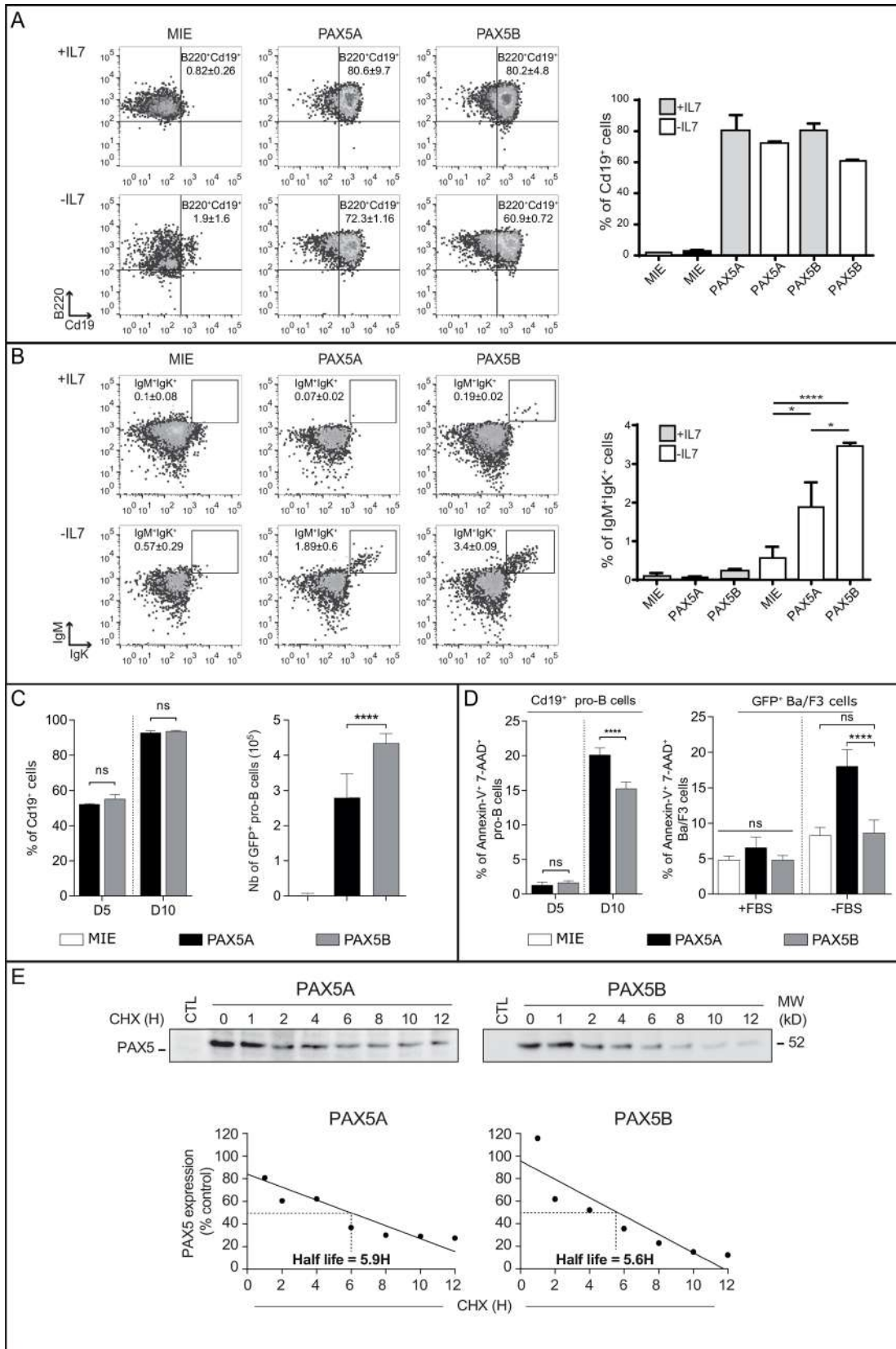
in their target gene expression may explain the differences observed in cell death rate and cell growth (Figure 3C and 3D).

### Pax5 isoforms have a distinct pattern of expression during late B cell differentiation

We also investigated the functional specificities of *PAX5* isoforms during late B cell differentiation. We quantified *Pax5* isoforms in murine sorted splenic resting B cells ( $Cd43^{neg}$   $B220^{pos}$ ), *in vitro* LPS-activated B cells ( $B220^{pos}$   $Cd138^{neg}$ ) and plasmablasts ( $B220^{pos}$   $Cd138^{pos}$ ). Overall, *Pax5* expression is decreased after B cell



**Figure 2: Correlation between *Pax5* isoforms expression and chromatin acetylation of the *Pax5* locus.** (A) QPCR of *Pax5A* and *Pax5B* isoforms on B cells during differentiation and in various cell lines. Expression of *Pax5* isoforms during medullar murine B cell differentiation was quantified on B cell subsets. Expression relative to *Ab11* are presented as mean and standard deviation (left panel); Expression of *Pax5A* (upper right panel) or *Pax5B* (lower right panel) on Ba/F3,70Z3, 38B9, 18.81, A20 cell lines and pro-B cells from murine bone marrow are expressed relative to *Gapdh* expression as mean and standard deviation. (B–D) Chromatin Immunoprecipitation (ChIP) analysis of *Pax5* locus on Ba/F3, 38B9, 70Z3 cell lines and murine pro-B sorted cells: QPCR was performed using primers covering the most conserved region (B) or located every 1 kb throughout promoter (C) or enhancer (D) regions. The increase over background was calculated by comparison with immunoprecipitates obtained using rabbit IgG antibody. Peaks of H3 acetylation are labeled according to their positions relative to the first nucleotide of *Pax5* exon 1A.



**Figure 3: Both PAX5A and PAX5B can restore B cell differentiation program in complemented *Pax5*<sup>-/-</sup> pro-B cells.** (A–B) Representative FACS analysis and statistics of differentiation of three independent infections of *Pax5*<sup>-/-</sup> pro-B cells. Cells were gated on FSC/SSC criteria, 7-AAD negative population (living cells) and eGFP-positive cells (infected cells) (not shown). Gated cells were analyzed for (A) the presence of B220 and Cd19 (the percentage of Cd19<sup>+</sup> cells is indicated on the right panel graph) or (B) IgM and Igk



membrane markers (the percentage of IgM<sup>+</sup>/Igk<sup>+</sup> cells is indicated on the right panel graph) in two different conditions: with (+IL7, upper line) or without IL7 (-IL7, lower line) for three days. The first column show *Pax5*<sup>-/-</sup> cells infected with the eGFP-expressing vector (MIE), the second column with *PAX5A* retroviral construct and the third column with the *PAX5B* retroviral construct. The mean and the standard deviation of percentage of each population compared to the parental gate are indicated, \**p* < 0.05, \*\*\*\* *p* < 0.001. (C) Mean and standard deviation of percentage of Cd19<sup>+</sup> cells (left panel); Absolute number of GFP<sup>+</sup> *Pax5*<sup>-/-</sup> pro-B complemented with eGFP (MIE), *PAX5A* or *PAX5B* after 3 days of culture for three independent infections, \*\*\*\* *p* < 0.001 (right panel). (D) Percentage of 7-AAD<sup>pos</sup> AnnexinV<sup>pos</sup> cells after 5 or 10 days in presence of IL7 (left panel) for three independent infections with *PAX5A*- or *PAX5B*-expressing vectors. \*\*\*\* *p* < 0.001. (E) Quantification of *PAX5A* and *PAX5B* proteins by Western-blot using an anti-Pax5 antibody after cycloheximide treatment for the indicated period of time on Ba/F3 infected cells demonstrating a similar half-life.

activation (Figure 5A). In details, *Pax5A* was the most expressed isoform and downregulated in activated B cells and plasmablasts. In contrast, *Pax5B* was significantly upregulated in activated B cells and almost undetectable in plasmablasts. Therefore, the expression patterns of the two isoforms are distinct during late B cell differentiation.

### **Pax5B is inactive in plasma cells**

The differential regulation of *Pax5* isoforms at late B cell differentiation stages may reflect different properties of these isoforms. To assess this possibility, we evaluated the transactivation of *Cd19* at an early B cell stage using transduced Ba/F3; both *PAX5A* and *PAX5B* can efficiently transactivate *Cd19* (Figure 5B). We also evaluated the transactivation potential of *PAX5A* and *PAX5B* at later stage of B cell differentiation using S194, a plasmacytoma cell line, or in a non-lymphoid context using 293T, a human kidney cell line, using a luciferase reporter system containing the *Cd19* promoter region. Compared to AID, a negative control, both *PAX5A* and *PAX5B* activate the *Cd19* promoter (Figure 5C). The activity of *PAX5B* is significantly lower than *PAX5A*. In S194, only *PAX5A* displays transcriptional activity suggesting that *PAX5B* transactivation capacity on *Cd19* promoter is altered in those cells (Figure 5C). To characterize the binding of Pax5 on the *Cd19* promoter at late stages of the B cell differentiation, we precipitated chromatin (ChIP) using an anti-Pax5 antibody in splenic B cells (Figure 5D). We detected a specific binding of Pax5 to the *Cd19* promoter sequence in resting B cells compared to the amplification of an irrelevant sequence (*Upfl* intronic sequence) devoid of Pax5 binding sites (Figure 5D, left panel). After 4 days of LPS stimulation, we could detect the same enrichment of Pax5 on *Cd19* promoter (Figure 5D, right panel).

### **Neither PAX5A nor PAX5B are able to silence the 3'RR in plasma cells**

The 3'RR is an important region regulating the expression of the heavy immunoglobulin chain at late B cell differentiation stages. It has been reported that Pax5 could bind hypersensitive sites (hs)1,2 (3'αE) and hs4 elements ([20, 21]). A *Pax5*-specific ChIP using an antibody that does not discriminate between the two isoforms, identified a 3.8-fold enrichment at the hs4 element site in resting B cells (Figure 5E, left panel). Upon

activation, Pax5 is enriched both at the hs4 and to a much lesser extend to hs1,2 sites (6-fold enrichment for hs4 and a 2.4-fold for hs1,2; Figure 5E, right panel).

We then analyzed whether both *PAX5* isoforms have a similar transactivation activity on the 3'RR in S194 plasma cells (Figure 5F). *PAX5* isoforms are physiologically downregulated in plasma cells, in contrast the 3'RR activity is increased. We evaluated the effect of the *PAX5* isoforms on the 3'RR activity using a reporter assay depending on the *VH* promoter (pVH) active in the B cell lineage. Its transcription level is upregulated by the addition of a synthetic 3'RR containing the four enhancers in their palindromic structure (hs3a, hs1.2, hs3b, hs4). We did not observe any significant change upon the addition of both *PAX5* isoforms, suggesting that 3'RR is insensitive to Pax5 (Figure 5F) despite the accumulation of Pax5 at the 3'RR after activation of B cells (Figure 5E, right panel).

## **DISCUSSION**

*PAX5* is a major regulator of B cell differentiation promoting the expression of B cell specific genes and repressing commitment to other cell lineages. Consequently, to allow an efficient B cell differentiation, the expression of *PAX5* has to be tightly controlled. An alteration of *PAX5* dosage induced B cell malignancies: its overexpression, secondary to the *IGH-PAX5* rearrangement, is associated with B cell lymphomas [16] and loss of heterozygosity is associated with B cell progenitor acute lymphoblastic leukemia [22, 23]. The regulation of *PAX5* expression, in vertebrates, is complex with two 5' isoforms, *PAX5A* and *PAX5B*, initiated from two different promoters. These isoforms differ only by the sequence of their first coding exon encoding no specific protein motif. *PAX5B* isoform is related to *PAX2*, since the sequence encoded by their first exon are 65% homologous and is present as early as in the Fugu (Table 1). The *PAX5A* isoform appears later, with mammals and by extension B cells.

Very little is known about the differential regulation of these two isoforms and their functional differences. The aim of this study was to determine if the presence of these two isoforms translate a functional difference or are merely a differential regulation of expression. In B cells, *PAX5* isoforms are differentially expressed suggesting that different sets of transcription factors enhance or repress the activity of these promoters. ChIP analyses on cell lines and pro-B cells confirm a correlation of histone H3

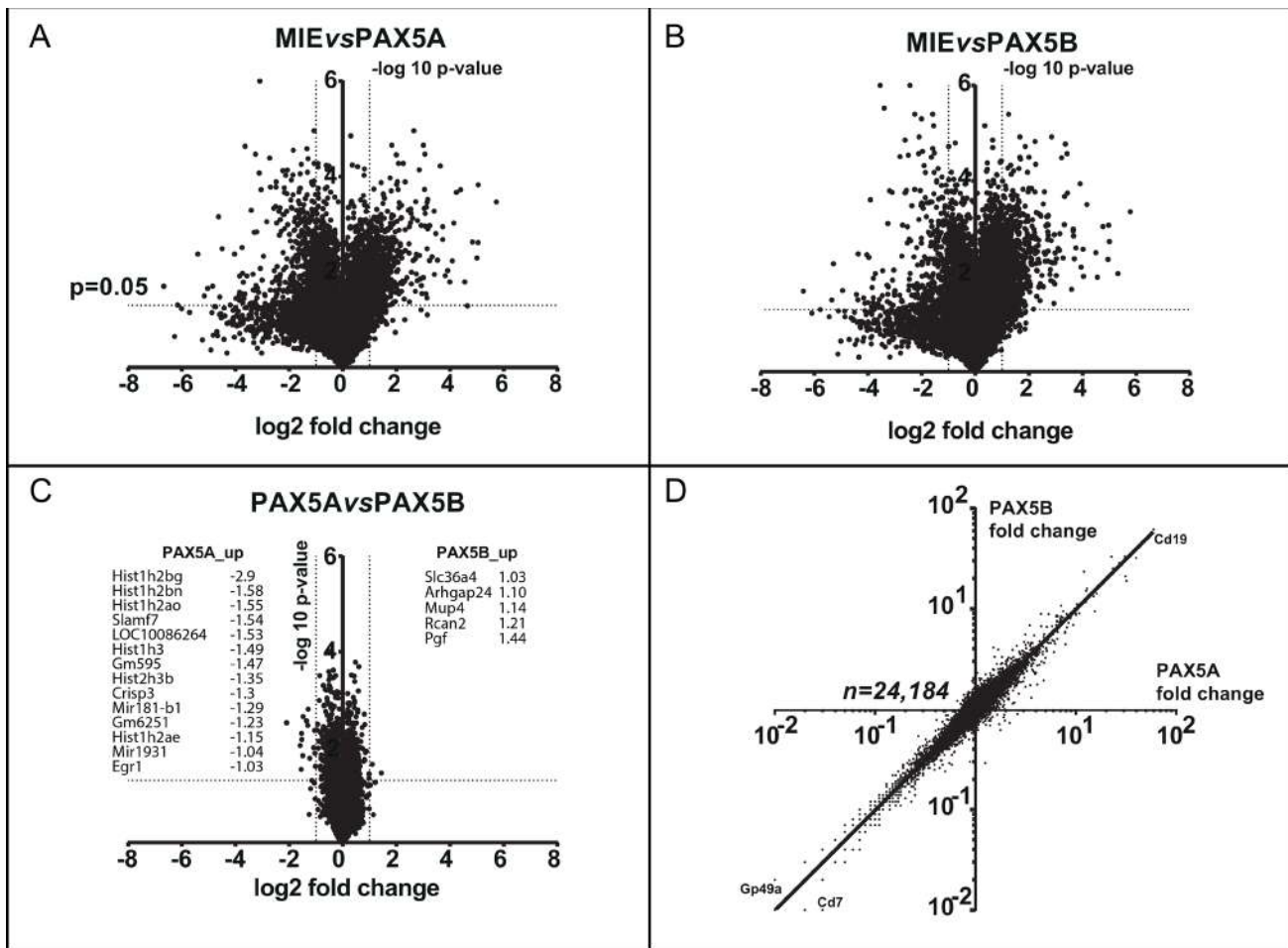
acetylation on both promoter regions and expression of the two isoforms. Interestingly, we also detected histone H3 acetylation on the enhancer region [24] only when the two isoforms are expressed.

Functionally, PAX5A and PAX5B can act similarly in early B cell differentiation. Both are able to complement the invalidation of *Pax5* in *Pax5*<sup>-/-</sup> cells and resume *ex vivo* the B cell differentiation until the Cd19<sup>pos</sup>/mIgK<sup>pos</sup> stage demonstrating a similar capacity to transactivate or repress target genes during this process. This is shown by the transcriptome since no gene expression differs more than five times in *Pax5*<sup>-/-</sup> complemented by PAX5A or PAX5B. Nonetheless, we detected differences between the two isoforms. PAX5B confers a growth advantage of the rescued *Pax5*<sup>-/-</sup> cells compared to PAX5A-rescued cells. The physiological up-regulation of *Pax5B* during B cell activation may prevent apoptosis of activated B cells since

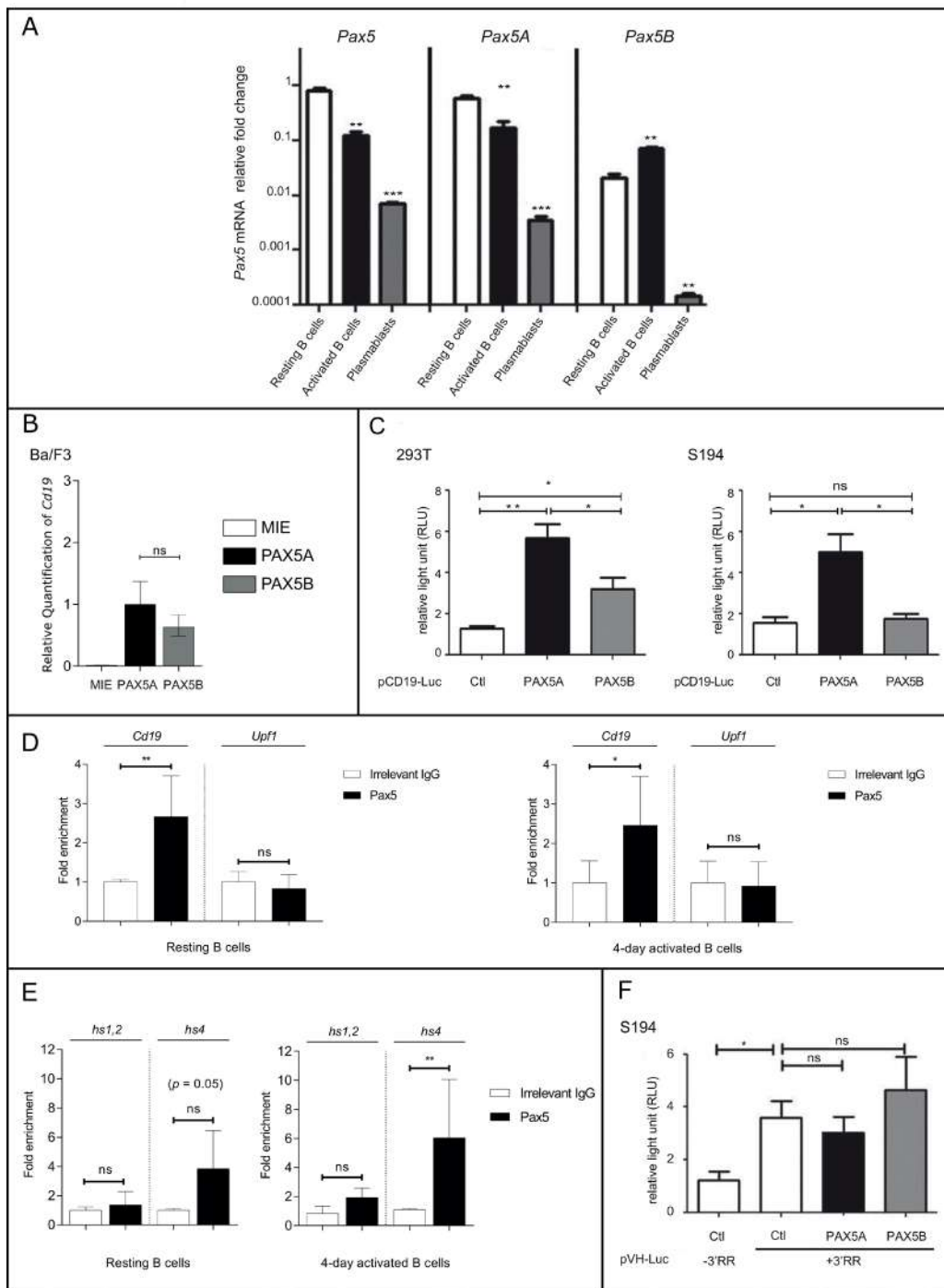
suppression of PAX5B increases susceptibility of human acute lymphocytic leukemia REH cells to apoptotic death [25]. Interestingly GO term analysis on discriminant genes for which the expression is at least 1.5 fold higher for PAX5A or PAX5B condition shows that they are involved in cell death pathways.

As described by Revilla-i-Domingo *et al.*, although Pax5 interacts with a similarly high number of genomic binding sites defining 8,000 target genes in pro-B and mature B cells, only 13 target genes are commonly activated and 18 target genes are commonly repressed in pro-B and mature B cells (cited from [26]) suggesting that Pax5A and Pax5B could have different transcriptional activities in early and late B cell differentiation stages.

At later stages of B cell differentiation, we observed that *Pax5B* and *Pax5A* are expressed simultaneously in primary B cells at different levels. After B cell activation,



**Figure 4: PAX5A and PAX5B drive a comparable transcriptomic early B cell program.** (A–C) Volcano plots presenting the fold change in log<sub>2</sub> on x axis and the log<sub>10</sub> *p*-value of the false discovery rate permutation analysis on y axis of the transcriptome comparing eGFP expressing vector (MIE) to PAX5A (A), MIE to PAX5B (B) and PAX5A to PAX5B (C) showing the similar effect of PAX5A and PAX5B on early B cell differentiation restoration. For C, gene with a difference of expression >2 fold between PAX5A and PAX5B are listed. (D) Scatter graph with gene expression fold changes between PAX5A and MIE on x-axis and PAX5B and MIE on y-axis. The regression line shows the relationship between PAX5A and PAX5B impact on target gene expression. As example, *Cd19* up-regulated gene and *Gp49a* and *Cd7* down-regulated genes are indicated.



**Figure 5: Pax5A and Pax5B show different activities during late B cell stages.** (A) Expression by QPCR of total *Pax5* or specific *Pax5* isoforms in resting murine B cells, activated B cells and plasmablasts. Results are expressed as log scale with mean  $\pm$  SEM of independent experiments. (B) Relative quantification of *Cd19* transactivation in Ba/F3 after infection with either eGFP-expressing retrovirus (MIE) or MIE containing *PAX5A* or *PAX5B* (C) *PAX5A* or *PAX5B* trans-activity was determined using a Luciferase reporter (pCD19-Luc) assay in 293T cells (left panel) or S194 cells (right panel). Results are expressed as mean  $\pm$  SEM of independent experiments. (D) ChIP was performed on *Cd19* promoter (pCD19) and *Upf1* intron 9 (as a *PAX5* non-target) and using an irrelevant IgG or an anti-*PAX5* antibody. Results are expressed as mean of fold change compared to IgG condition  $\pm$  SEM of independent experiments. (E) ChIP was performed on the hypersensitive sites (hs)1,2 and hs4 elements of the 3'RR in resting splenic B cells or activated B cells (left and right panel respectively) using an irrelevant IgG or a anti-*PAX5* antibody. Results are expressed as mean  $\pm$  SEM of independent experiments. (F) A *Luciferase* expression B cell specific vector (pVH-luc) containing the IGH 3'RR minilocus (+3'RR) or not (-3'RR) was used in a reporter assay to determine the regulatory function of *PAX5A* and *PAX5B* in S194 cells. Results are expressed as mean of fold change compared to IgG condition  $\pm$  SEM of independent experiments. Ctl: negative control, \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0,001$  (unpaired two-tailed *t* test).

**Table 1: Comparative analysis of exon 1A and exon 1B coding sequences during evolution**

Species	PAX5 coding exon 1A	PAX5 coding exon 1B
Human	MDLEKNYPTPRTSRT	MEIHCKHDPFASMH
Chimpanzee	MDLEKNYPTPRTSRT	MEIHCKHDPFASMH
Mouse	MDLEKNYPTPRTIRT	MEIHCKHDPFASMH
Rat	MDLEKNYPTPRTIRT	MEIHCKHDPFASMH
Rabbit	MDLEKNYPTPRTGRT	MEIHCKHDPFASMH
Dog	MDLEKNYPTPRSGRT	MEIHCKHDPFASMH
Pig	MDLEKNYPTPRTGRT	MEIHCKHDPFASMH
Cow	MDLEKNYPTPRTGRT	MEIHCKHDPFASMH
Xenopus	No hits found	MEIHCKHDPFAAMH
Zebrafish	No hits found	MEIHCKHDPFAAMH
Fugu	No hits found	MEIHCKHDPFAAMH
C. elegans	No hits found	No hits found
Fruitfly	No hits found	No hits found
Lamprey	No hits found	No hits found

The table shows the alignment of PAX5 exon 1A and exon 1B primary sequence in different species. The discrepancies between species are boxed.

they exhibit a distinct pattern of expression: *Pax5A* decreases in activated B cells and plasmablasts while *Pax5B* increases in activated B cells and is virtually absent in plasmablasts.

PAX5A is able to transactivate *Cd19* promoter in both early (Ba/F3) and late B cells (S194) and in a non-lymphoid 293T whereas PAX5B is able to transactivate *Cd19* only at an early stage of B cell differentiation and in a non-lymphoid context. These results point out that PAX5B isoform cannot transactivate the *Cd19* promoter in a late lymphoid context (S194) or may function in a different context as transcriptional repressor as proposed by Robichaud *et al.* The authors have shown that PAX5B negatively regulates *CD19* expression in REH cell line since the specific suppression of *PAX5B* expression using ribozymes leads to an increase of *CD19* expression [25].

PAX5 has been described as a negative regulator of the 3'RR activity. Nonetheless, we observed an increased binding of PAX5 to the 3'RR upon B cell activation. The accumulation of PAX5 binding on the 3'RR in activated B cells does not seem to reflect a role of PAX5 in the 3'RR activity regulation but rather a function in chromatin remodeling permitting accessibility to the *IGH* locus. Recently, it has been shown that PAX5 recruits PTIP, a molecule involved in class switch recombination [27, 28] which induces H3K4me3 and stabilizes the PAX5 binding on the *IGH* locus [28]. Alternatively, this could be explained by the fact that the activation of B cells results in the induction or increase of the expression of other factors counteracting PAX5 activity.

PAX5 isoforms are physiologically downregulated in plasma cells in contrast to the 3'RR activity which is strongly increased. We investigated whether these

isoforms could modify the 3'RR activity. Our data showed that 3'RR is insensitive to a PAX5A or PAX5B regulation in the S194 plasmacytoma cell line, in agreement to the observation that an hs1,2-dependent transgene is not inhibited whereas Pax5 expression is still detected in activated B cells [29]. Consequently, repression of the 3'RR activity by Pax5 [20, 30, 31] is likely more complex.

Our results reveal only discrete differences of action between PAX5A and PAX5B functions and demonstrate that these two isoforms may be interchangeable in early B cell differentiation. Their very tight regulation may reflect different binding of alternative transcription factors according to the cellular context (testis, brain, B cells) leading to a fine control of the *Pax5* dosage which is required for a zero-default B cell differentiation.

## MATERIALS AND METHODS

### Cell line culture

In this study, we used several B cell lines showing characteristics of various stages of B cell differentiation: Ba/F3 are pro-B cells (DSMZ ACC 300) [32], 38B9 are pro-B cells [33, 34], 70Z/3 are pre-B lymphoblast (ATCC® TIB-158™) [35], A20 is a BALB/c B cell lymphoma line derived from a spontaneous reticulum cell neoplasm found in an old BALB/cAnN mouse (ATCC® TIB-208™) [36], 18.81 is a Abelson-virus-transformed mouse lymphoid cell line with pre-B cell characteristic [37], WEHI-231 are immature B cells (ATCC® CRL-1702™) [38], S194 are mouse myeloma B cells, this line was derived from an IgA secreting mineral oil induced BALB/c myeloma.



(ATCC® TIB-19™) [39]. Murine B cell lines S194, 38B9, A20, 18.81, and WEHI-231 were cultured in RPMI (Gibco), 10% FBS (fetal bovine serum, PAN Biotech), 2 mM L-glutamine, and 1 mM sodium pyruvate (Invitrogen). The same medium supplemented with 0.05 mM  $\beta$ -mercaptoethanol or 10 ng/ml recombinant murine IL3 (PeproTech) was used respectively to culture the murine cell line 70Z3 or Ba/F3. The Phoenix retrovirus producer line (Orbigen) was cultured in high glucose Dulbecco's Modified Eagle's Medium (DMEM; Gibco), 10% FBS (PAN Biotech). Murine OP9 bone marrow stromal cells were cultured in MEM $\alpha$  (Gibco), 20% FBS (PAN Biotech) when in monoculture. All the cells were cultured in presence of penicillin (100U/ml) and streptomycin (100U/ml) at 37° C, in 5% CO<sub>2</sub> atmosphere.

## Mice

*Pax5*<sup>±</sup> mice were kindly provided by Pr M. Busslinger [6]. *Pax5*<sup>-/-</sup> and *Pax5*<sup>+/+</sup> (wild-type) mice were obtained by breeding two *Pax5*<sup>±</sup> mice. All mice were maintained on the C57BL/6 genetic background. The *Pax5* genotype was determined using REDExtract-N-Amp Tissue PCR Kit according to the manufacturer's instructions (Sigma-Aldrich) with the following oligonucleotides: 5'-AGG ACA TGG AGG AGT GAA TCA G-3' (primer1); 5'-ACC CGA AGC TGC CTG GAG-3' (primer2); and 5'-AGG CGA TTA AGT TGG GTA ACG-3' (primer 3). All animal experiments were carried out according to valid project licenses, which were approved and regularly controlled by French Authorities.

## Purification and culture of B cell subsets

*Pax5*-deficient pro-B cells were harvested from *Pax5*<sup>-/-</sup> embryonic liver at E17.5. Cells were amplified on  $\gamma$ -irradiated (30 Gy) OP9-derived stromal cells in Iscove's Modified Dulbecco's Medium (IMDM; Gibco), supplemented with 5% FBS (Stemcell technologies), 100 U/ml murine IL7 (PeproTech), 0.05 mM  $\beta$ -mercaptoethanol (Sigma-Aldrich), 2 mM L-glutamine (Invitrogen), penicillin (100 U/ml) and streptomycin (100 U/ml). Wild-type pro-B cells were flushed from bone marrow. After red blood cell lysis, B cell fraction was enriched by B220 magnetic microbead sorting (Myltenyi Biotech). B cell fraction was subsequently incubated with B220, Kit (Cd117) and Igk antibodies and pro-B cells (B220<sup>pos</sup> kit<sup>pos</sup> and Igk<sup>neg</sup>) were sorted using a FACS-Vantage sorter (BD Biosciences). Splenic resting B cells were purified by negative selection using anti-CD43 microbeads (Myltenyi Biotech) and stimulated (0.5 × 10<sup>6</sup> cells/ml) by LPS (20  $\mu$ g/ml; Sigma) for 2 to 4 days in RPMI supplemented with 10% FBS, sodium pyruvate, non-essential amino acids,  $\beta$ -mercaptoethanol, penicillin (100 U/ml) and streptomycin (100 U/ml) (Invitrogen).

B cell subsets were sorted on a MoFlo Astrios sorter (Beckman Coulter) after bone marrow flush of C57BL/6 8 to 12 week-old mice as follows: pre-pro-B (B220<sup>+</sup>Cd19<sup>-</sup>), pro-B (B220<sup>+</sup>Cd19<sup>+</sup>IgL-IgD<sup>-</sup>Cd117<sup>+</sup>BP1<sup>-</sup>), pre-BI (B220<sup>+</sup>Cd19<sup>+</sup>IgL-IgD<sup>-</sup>Cd117<sup>+</sup>BP1<sup>+</sup>), large pre-BII (B220<sup>+</sup>Cd19<sup>+</sup>IgL-IgD<sup>-</sup>Cd117<sup>+</sup>, large FSC), small pre-BII (B220<sup>+</sup>Cd19<sup>+</sup>IgL-IgD<sup>-</sup>Cd117<sup>+</sup>, small FSC), immature B (B220<sup>+</sup>Cd19<sup>+</sup>IgL<sup>+</sup>IgD<sup>-</sup>), recirculating B (B220<sup>+</sup>Cd19<sup>+</sup>IgL<sup>+</sup>IgD<sup>+</sup>) and RNA extraction were performed on fresh samples.

For LPS stimulation B cells (B220<sup>pos</sup> Cd138<sup>neg</sup>) and plasmablasts (B220<sup>pos</sup> Cd138<sup>pos</sup>) were sorted 3 to 4 days after stimulation.

## Quantitative RT-PCR

RNA was isolated using the AllPrep™ DNA/RNA Micro Kit (QIAGEN) and cDNA was synthesized using SuperScript® VILO™ cDNA Synthesis Kit (Invitrogen™) according to the manufacturer's instructions. Quantitative PCR were performed on a LightCycler®480 II System (Roche). Quantitative SYBR Green PCR was performed to quantify relative mouse Cd19 cDNA expression in transduced Ba/F3 using LightCycler®480 SYBR Green I Master (Roche Diagnostics GmbH) according to the manufacturer's instructions, and primers 10  $\mu$ M primers for *Cd19* (forward: 5'-AAGGAACAGGTCCTCTGGGA-3' and reverse: 5'-GATTCAAACTGCTCCCCCGA-3') and 10  $\mu$ M primers for *Ab11* (forward: 5'-GGAGAA GGTCTACGAGCTCAT-3' and reverse: 5'-ATCTGAGATA CTGGATTCCTGGA-3'). Relative gene expression in cell lines or sorted bone marrow B cells was quantified using LightCycler®480 Probes Master (Roche Diagnostics GmbH) and the following TaqMan gene expression assays from Applied Biosystems™ Life Technologies: Mm00802038\_m1 for *Ab11*, Mm00487680\_m1 for *Melk*, Mm00435501\_m1 for *Pax5* (both isoforms), Mm00621713\_m1 for *Zcchc7*, Mm00395519\_m1 for *Ebf1*, Mm00435494 for specific expression of *Pax5A*. Primers/TaqMan probe mix specific for PAX5B transcripts was synthesized by Applied Biosystems (5'FAM-ATG CAT AGA CAT GGA GGA GTG AAT CAG CTT G-Tamra3' as probe; 5'-CAC TGT AAG CAC GAC CCG TT-3' as forward primer; 5'-AGT GGC CGT CCA TTC ACA AA-3' as reverse primer). All QPCR programs were carried out as follows in a 20  $\mu$ l volume: 5 min at 95° C, followed by 45 cycles of 10 s at 95° C, 10 s at annealing temperature of 60° C and 10 s at 72° C. All signals were quantified using either the  $\Delta$ Ct method or the Relative Quantification ( $\Delta\Delta$ Ct) study both including a normalization to the  $\Delta$ Ct values of *Ab11* gene expression levels. Data were analyzed using the LC480 software (Roche Diagnostics) followed by statistical analyses.

## ChIP experiment

Chromatin ImmunoPrecipitation (ChIP) assays were performed essentially as described previously [40]

using anti-acetylated histone H3 or nonspecific rabbit immunoglobulin G (IgG) (06-599 and 12-370, respectively; Upstate Biotechnology, Lake Placid, NY). Oligonucleotide sequences are available on Supplementary Table 1. Murine splenic B cells were cultured ( $10^6$  cells/ml) in presence of LPS (40  $\mu$ g/ml) for 4 days. At day 3, dead cells were eliminated by Ficoll centrifugation (Lympholyte-Mammal, Cedarlane) and living cells were stimulated for an additional 24 h period. Chromatin preparations were performed on  $5 \times 10^7$  to  $10^8$  LPS-stimulated B cells using standard fixation and sonication methods as described [41]. In all cases, 200 ng of chromatin was immunoprecipitated (ON, 4° C) with 1 $\mu$ g of anti-PAX5 mAb (mouse anti-Pax-5 mAb (A-11), Santa Cruz Biotechnology) or an irrelevant IgG antibody (purified mouse unlabeled IgG1, SouthernBiotech, ref.0102-01). Oligonucleotides used are given in Supplementary Table 1.

### Retroviral particles production

Full-length human *PAX5A* and *PAX5B* cDNA were amplified, sequenced and inserted into the retroviral vector pMSCV-IRES-eGFP (MIE), a bicistronic vector allowing the expression of eGFP under the control of the long terminal repeat (LTR) promoter. Retroviral supernatants were produced by Phoenix retrovirus producer line (Orbigen) using Lipofectamine 2000 (Invitrogen) reagent for transfection and proceeding according to the manufacturer's instructions. After overnight incubation of the transfection mix, cells were treated with 10 mM sodium butyrate (Millipore) and washed with PBS before pro-B cell medium was added. Viral supernatants were harvested after 24 h of incubation at 32° C, passed through a 0.45  $\mu$ m filter, aliquoted and frozen at -80° C.

### Pro-B culture, infection and differentiation

The *in vitro* pro-B cell differentiation/proliferation protocol used was adapted from Rolink *et al.* [42]. The pro-B *Pax5*<sup>-/-</sup> cells were isolated from *Pax5*<sup>-/-</sup> embryo fetal liver and cocultivated with irradiated OP9 cells in 24-well plates in Iscove's Modified Dulbecco's Medium (IMDM; Gibco), supplemented with 5% FBS (Stemcell Technologies), 2 ng/ml murine IL7 (Peprotech), 0.05 mM  $\beta$ -mercaptoethanol (Sigma-Aldrich), 2 mM L-glutamine (Invitrogen), penicillin (100 U/ml) and streptomycin (100 U/ml). Experiments were performed after an initial pro-B cell expansion on irradiated OP9 cells in the presence of 2 ng/ml IL7. Every 3 days, pro-B cells were harvested and propagated on fresh OP9 stromal cells.

The pro-B cells, cultured on irradiated OP9 cells in the presence of 2 ng/ml of IL7, are infected during exponential growth. Retroviral supernatant was diluted 1:3, supplemented with 4  $\mu$ g/ml polybrene (Sigma) and 2 ng/ml of IL7 (Peprotech); spinoculation was performed at 1000  $\times$  g

at 32° C for 90 min. After centrifugation, 1ml of fresh medium supplemented with IL7 was added and the cells were incubated overnight at 37° C. After infected pro-B cell expansion for few days, cells were washed out of IL-7 and plated in IMDM, 5% FBS, 0.05 mM  $\beta$ -mercaptoethanol and 2 mM L-glutamine on irradiated OP9-derived stromal cells with or without IL7. At day 2.5, single-cell suspensions were stained using standard protocols for flow cytometry using the antibodies listed above. Dead cells were excluded from FACS analysis as 7-amino-actinomycin D (7-AAD; BD Biosciences #559763)) positive (BD Biosciences). Stained cells were analyzed with a FACS Calibur flow cytometer (BD). Data were analyzed with the BD FACSDiva™ software (BD Biosciences) or FlowJo (TreeStar, Ashland, OR). All antibodies are rat monoclonal antibodies from BD biosciences: anti-B220 (clone RA3-6B2) conjugated with pacific blue, anti-CD19 (clone 1D3) conjugated with Alexa Fluor 700, anti-mouse Igk light chain (clone 187.1) conjugated with Phycoerythrin-Cyanin7, anti-IgM (clone R6-60.2) conjugated with Peridinin chlorophyll protein-cyanin5.5.

Ba/F3 cells were similarly transduced with empty MIE, MIE-PAX5A or MIE-PAX5B and then GFP-positive cells were isolated by fluorescence-activated cell sorting (FACS) in MoFlo Astrios (Beckman Coulter).

### Cycloheximide treatment and Western blot

Stability assay was performed in the presence of cycloheximide (100  $\mu$ g/mL, Sigma) for the indicated periods of time. The day before,  $1.2 \times 10^6$  transduced Ba/F3 cells were incubated overnight at 37° C, 5% CO<sub>2</sub>. At each time point, treated cells were washed once with ice-cold Phosphate buffer saline (PBS) and then cell pellets were lysed in 40  $\mu$ l of lysis buffer composed of 150 mM NaCl, 1 mM EDTA, 50 mM Tris, 1% Triton and protease inhibitor cocktail (cOmplete™, Roche). Protein concentrations of cell lysates were determined using the BCA1-Kit for Protein Determination (Sigma-Aldrich) to load 30  $\mu$ g in every well. Samples were subjected to 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred to PVDF membranes. Blots were blocked 1 hour in the presence of 5% milk and incubated with 0.3  $\mu$ g/mL of primary antibody overnight at +4° C (mouse anti-Pax-5 mAb (A-11), Santa Cruz Biotechnology). After incubation with secondary antibodies conjugated to horseradish peroxidase (Cell Signaling Technology), the immunoreactive bands were visualized by using the enhanced chemiluminescence lighting system (Amersham™ ECL™ Prime Western Blotting Detection Reagent, GE Healthcare) and data was acquired with Chemi-Smart 5000 and Bio-1D software (Vilber Lourmat). Bands were quantified by using ImageJ software.

## Flow cytometry and cell growth

Transduced pro-B cells and Ba/F3 cells were stained with AlexaFluor700-conjugated anti-CD19 (1D3) antibody (BD Pharmingen). For cell-viability assays, cells were stained with Annexin-V-PE and 7-AAD (BD Biosciences #559763). Flow cytometric data were acquired on a Fortessa X20 or LSR II flow cytometer (BD Biosciences), and data were analyzed using FACSDiva (BD Biosciences) and FlowJo software (TreeStar).

## Transcriptomic analysis

After 5 days in culture with 2 ng/ml of IL7, infected cells were sorted according to their positivity for GFP fluorescence with a FACSARIA II (BD Biosciences) cell sorter. Total RNA from GFP<sup>pos</sup> sorted pro-B infected cells were extracted using the Trizol method according to manufacturer's instruction (Invitrogen). RNA quality was assessed using the RNA 6000 Nano Assay on the Agilent 2400 Bioanalyser (Agilent Technologies, Massy, FRANCE). For each condition (PAX5A, PAX5B and MIE), three independent infection experiments were performed. RNA samples were purified and prepared according to the manufacturer's protocol with the Affymetrix's GeneChip Whole Transcript Sense Target Labeling Assay Kit (Affymetrix, UK) and hybridized on Affymetrix GeneChip Mouse Gene 2.0 ST arrays. Probe-signal intensities were normalized and summarized by the Robust Multiarray Average (RMA) method using Affymetrix's Expression Console software.

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus [43] and are accessible through GEO Series accession number GSE104890 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE104890>).

For the volcano plots, permutation analyses were performed with dChip software [44] allowing the calculation of false discovery rate  $p$ -value and the fold difference volcano plots were generated using GraphPad prism software.

The gene ontology analysis has been performed with the AutoCompare ZE software [45] using Gene Ontology database [18, 19],  $p$ -values have been calculated with the ZE test.

## Luciferase reporter assay

Experiments were performed by cell transfections with 2 $\mu$ g DNA at 1:7:1 (pcDNA3-PAX5 vector: reporter vector: GFP-CMV) vector ratios. GFP-CMV was used to evaluate transfection efficiency. The S194 line was cultured in RPMI as above and for each condition  $2 \times 10^6$  cells were electroporated according to the manufacturer's instructions (Lonza). The 293T line was cultured in DMEM (with 10% FBS, sodium pyruvate, nonessential amino acids, penicillin and streptomycin) and subconfluent cells in 6-well

dishes were transfected by lipofection with Exgen500 (Euromedex). Firefly luciferase-CD19-2(A ins) [46] was kindly provided by Pr M. Busslinger was used to measure PAX5 activity on the PAX5 binding site of the *Cd19* promoter. PAX5 activity on the 3'RR was evaluated with firefly luciferase-pV<sub>H</sub> enhanced or not by the combination of core enhancers of 3'RR [47]. After 48 h transfection, cells were collected and split for GFP fluorescence analysis by flow cytometry and luminescence detection using the Luciferase Reporter Assay System and the 20/20<sup>n</sup> Luminometer (Promega). For each sample, luminescence was normalized to the measured GFP fluorescence.

## Statistical analysis

Experiments have been performed at least as three independent conditions. All statistical analyses were performed using Excel (Microsoft) and GraphPad Prism software, version 7 (GraphPad Software). Statistical differences were determined using a 2-tailed Student's  $t$  test or RM one-way ANOVA with Tukey's correction or one-way ANOVA with Sidak's correction for multiple comparisons or Pearson correlation, as mentioned in legends. All data are presented as mean  $\pm$  SD or SEM. A  $p$ -value of less than 0.05 was considered statistically significant, the null hypothesis being rejected at the 0.05 level (\*\* $p < 0.0005$ , \* $p < 0.005$ ,  $p < 0.05$ ).

## ACKNOWLEDGMENTS

The authors thank "the Pôle Technologique of CRCT – Plateaux Cytométrie et Bioinformatique (INSERM U1037), Toulouse" for technical assistance. We also thank the Anexplo-Génotoul platform CREFRE INSERM - UMS 006 animal facility.

## CONFLICTS OF INTEREST

The authors declare no financial or commercial conflict of interest.

## GRANT SUPPORT

This work has been supported by INCa R11190BB-INCA 2011-131-PAX5, by ARC grant ARC SFI20101201888 and the FEDER grant CITTIL. CC, SP and NR have been supported by INCa grant R11190BB-INCA 2011-131-PAX5. The team is supported by association Laurette FUGAIN, la société française des cancers de l'enfant (SFCE), The association Capucine, "111 des Arts", la Ligue régionale contre le cancer and the "région Occitanie".

MC, ED, LD and CB designed the experiments. CC, SP, LJ, NR, NP, MD, BG, SH and SL performed the experiments. SM designed the FACS analysis strategy and performed the preliminary experiments. LD, SP, ED and CB wrote the manuscript.



## REFERENCES

1. Nutt SL, Heavey B, Rolink AG, Busslinger M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature*. 1999; 401:556–62. <https://doi.org/10.1038/44076>.
2. Lang D, Powell SK, Plummer RS, Young KP, Ruggeri BA. PAX genes: roles in development, pathophysiology, and cancer. *Biochem Pharmacol*. 2007; 73:1–14. <https://doi.org/10.1016/j.bcp.2006.06.024>.
3. Stapleton P, Weith A, Urbanek P, Kozmik Z, Busslinger M. Chromosomal localization of seven PAX genes and cloning of a novel family member, PAX-9. *Nat Genet*. 1993; 3:292–8. <https://doi.org/10.1038/ng0493-292>.
4. Adams B, Dorfner P, Aguzzi A, Kozmik Z, Urbanek P, Maurer-Fogy I, Busslinger M. Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. *Genes Dev*. 1992; 6:1589–607.
5. Nutt SL, Thevenin C, Busslinger M. Essential functions of Pax-5 (BSAP) in pro-B cell development. *Immunobiology*. 1997; 198:227–35. [https://doi.org/10.1016/S0171-2985\(97\)80043-5](https://doi.org/10.1016/S0171-2985(97)80043-5).
6. Urbanek P, Wang ZQ, Fetka I, Wagner EF, Busslinger M. Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell*. 1994; 79:901–12.
7. 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; 24:269–81. <https://doi.org/10.1016/j.immuni.2006.01.012>.
8. Nutt SL, Urbanek P, Rolink A, Busslinger M. Essential functions of Pax5 (BSAP) in pro-B cell development: difference between fetal and adult B lymphopoiesis and reduced V-to-DJ recombination at the IgH locus. *Genes Dev*. 1997; 11:476–91.
9. Dunnick WA, Collins JT, Shi J, Westfield G, Fontaine C, Hakimpour P, Papavasiliou FN. Switch recombination and somatic hypermutation are controlled by the heavy chain 3' enhancer region. *J Exp Med*. 2009; 206:2613–23. <https://doi.org/10.1084/jem.20091280>.
10. Rouaud P, Vincent-Fabert C, Saintamand A, Fiancette R, Marquet M, Robert I, Reina-San-Martin B, Pinaud E, Cogne M, Denizot Y. The IgH 3' regulatory region controls somatic hypermutation in germinal center B cells. *J Exp Med*. 2013; 210:1501–7. <https://doi.org/10.1084/jem.20130072>.
11. Saintamand A, Rouaud P, Saad F, Rios G, Cogne M, Denizot Y. Elucidation of IgH 3' region regulatory role during class switch recombination via germline deletion. *Nat Commun*. 2015; 6:7084. <https://doi.org/10.1038/ncomms8084>.
12. Vincent-Fabert C, Fiancette R, Pinaud E, Truffinet V, Cogne N, Cogne M, Denizot Y. Genomic deletion of the whole IgH 3' regulatory region (hs3a, hs1,2, hs3b, and hs4) dramatically affects class switch recombination and Ig secretion to all isotypes. *Blood*. 2010; 116:1895–8. <https://doi.org/10.1182/blood-2010-01-264689>.
13. Schaniel C, Bruno L, Melchers F, Rolink AG. Multiple hematopoietic cell lineages develop *in vivo* from transplanted Pax5-deficient pre-B I-cell clones. *Blood*. 2002; 99:472–8.
14. Cobaleda C, Jochum W, Busslinger M. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature*. 2007; 449:473–7. <https://doi.org/10.1038/nature06159>.
15. Palmisano WA, Crume KP, Grimes MJ, Winters SA, Toyota M, Esteller M, Joste N, Baylin SB, Belinsky SA. Aberrant promoter methylation of the transcription factor genes PAX5 alpha and beta in human cancers. *Cancer Res*. 2003; 63:4620–5.
16. Busslinger M, Klix N, Pfeffer P, Graninger PG, Kozmik Z. Deregulation of PAX-5 by translocation of the Emu enhancer of the IgH locus adjacent to two alternative PAX-5 promoters in a diffuse large-cell lymphoma. *Proc Natl Acad Sci U S A*. 1996; 93:6129–34.
17. Delabesse E, Ogilvy S, Chapman MA, Piltz SG, Gottgens B, Green AR. Transcriptional regulation of the SCL locus: identification of an enhancer that targets the primitive erythroid lineage *in vivo*. *Mol Cell Biol*. 2005; 25:5215–25. <https://doi.org/10.1128/MCB.25.12.5215-5225.2005>.
18. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000; 25:25–9. <https://doi.org/10.1038/75556>.
19. The Gene Ontology Consortium. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res*. 2017; 45: D331–D8. <https://doi.org/10.1093/nar/gkw1108>.
20. Michaelson JS, Singh M, Snapper CM, Sha WC, Baltimore D, Birshtein BK. Regulation of 3' IgH enhancers by a common set of factors, including kappa B-binding proteins. *J Immunol*. 1996; 156:2828–39.
21. Chatterjee S, Ju Z, Hassan R, Volpi SA, Emelyanov AV, Birshtein BK. Dynamic changes in binding of immunoglobulin heavy chain 3' regulatory region to protein factors during class switching. *J Biol Chem*. 2011; 286:29303–12. <https://doi.org/10.1074/jbc.M111.243543>.
22. Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, Girtman K, Mathew S, Ma J, Pounds SB, Su X, Pui CH, Relling MV, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*. 2007; 446:758–64. <https://doi.org/10.1038/nature05690>.
23. Familiades J, Bousquet M, Lafage-Pochitaloff M, Bene MC, Beldjord K, De Vos J, Dastugue N, Coyaud E, Struski S, Quelen C, Prade-Houdellier N, Dobbstein S, Cayuela JM, et al. PAX5 mutations occur frequently in



- adult B-cell progenitor acute lymphoblastic leukemia and PAX5 haploinsufficiency is associated with BCR-ABL1 and TCF3-PBX1 fusion genes: a GRAALL study. *Leukemia*. 2009; 23:1989–98. <https://doi.org/10.1038/leu.2009.135>.
24. Decker T, Pasca di Magliano M, McManus S, Sun Q, Bonifer C, Tagoh H, Busslinger M. Stepwise activation of enhancer and promoter regions of the B cell commitment gene Pax5 in early lymphopoiesis. *Immunity*. 2009; 30:508–20. <https://doi.org/10.1016/j.immuni.2009.01.012>.
  25. Robichaud GA, Perreault JP, Ouellette RJ. Development of an isoform-specific gene suppression system: the study of the human Pax-5B transcriptional element. *Nucleic Acids Res*. 2008; 36:4609–20. <https://doi.org/10.1093/nar/gkn432>.
  26. Revilla IDR, Bilic I, Vilagos B, Tagoh H, Ebert A, Tamir IM, Smeenk L, Trupke J, Sommer A, Jaritz M, Busslinger M. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J*. 2012; 31:3130–46. <https://doi.org/10.1038/emboj.2012.155>.
  27. Daniel JA, Santos MA, Wang Z, Zang C, Schwab KR, Jankovic M, Filsuf D, Chen HT, Gazumyan A, Yamane A, Cho YW, Sun HW, Ge K, et al. PTIP promotes chromatin changes critical for immunoglobulin class switch recombination. *Science*. 2010; 329:917–23. <https://doi.org/10.1126/science.1187942>.
  28. Schwab KR, Patel SR, Dressler GR. Role of PTIP in class switch recombination and long-range chromatin interactions at the immunoglobulin heavy chain locus. *Mol Cell Biol*. 2011; 31:1503–11. <https://doi.org/10.1128/MCB.00990-10>.
  29. Andersson T, Neurath MF, Grant PA, Pettersson S. Physiological activation of the IgH 3' enhancer in B lineage cells is not blocked by Pax-5. *Eur J Immunol*. 1996; 26:2499–507. <https://doi.org/10.1002/eji.1830261034>.
  30. Neurath MF, Strober W, Wakatsuki Y. The murine Ig 3' alpha enhancer is a target site with repressor function for the B cell lineage-specific transcription factor BSAP (NF-HB, S alpha-BP). *J Immunol*. 1994; 153:730–42.
  31. Singh M, Birshstein BK. NF-HB (BSAP) is a repressor of the murine immunoglobulin heavy-chain 3' alpha enhancer at early stages of B-cell differentiation. *Mol Cell Biol*. 1993; 13:3611–22.
  32. Palacios R, Henson G, Steinmetz M, McKearn JP. Interleukin-3 supports growth of mouse pre-B-cell clones *in vitro*. *Nature*. 1984; 309:126–31.
  33. Alt F, Rosenberg N, Lewis S, Thomas E, Baltimore D. Organization and reorganization of immunoglobulin genes in A-MULV-transformed cells: rearrangement of heavy but not light chain genes. *Cell*. 1981; 27:381–90.
  34. Alt FW, Yancopoulos GD, Blackwell TK, Wood C, Thomas E, Boss M, Coffman R, Rosenberg N, Tonegawa S, Baltimore D. Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J*. 1984; 3:1209–19.
  35. Paige CJ, Kincade PW, Ralph P. Murine B cell leukemia line with inducible surface immunoglobulin expression. *J Immunol*. 1978; 121:641–7.
  36. Kim KJ, Kanellopoulos-Langevin C, Merwin RM, Sachs DH, Asofsky R. Establishment and characterization of BALB/c lymphoma lines with B cell properties. *J Immunol*. 1979; 122:549–54.
  37. Rosenberg N, Baltimore D. The effect of helper virus on Abelson virus-induced transformation of lymphoid cells. *J Exp Med*. 1978; 147:1126–41.
  38. Boyd AW, Goding JW, Schrader JW. The regulation of growth and differentiation of a murine B cell lymphoma. I. Lipopolysaccharide-induced differentiation. *J Immunol*. 1981; 126:2461–5.
  39. Horibata K, Harris AW. Mouse myelomas and lymphomas in culture. *Exp Cell Res*. 1970; 60:61–77.
  40. Forsberg EC, Downs KM, Christensen HM, Im H, Nuzzi PA, Bresnick EH. Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. *Proc Natl Acad Sci U S A*. 2000; 97:14494–9. <https://doi.org/10.1073/pnas.97.26.14494>.
  41. Iacovoni JS, Caron P, Lassadi I, Nicolas E, Massip L, Trouche D, Legube G. High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. *EMBO J*. 2010; 29:1446–57. <https://doi.org/10.1038/emboj.2010.38>.
  42. Rolink AG. B-cell development and pre-B-1 cell plasticity *in vitro*. *Methods Mol Biol*. 2004; 271:271–81. <https://doi.org/10.1385/1-59259-796-3:271>.
  43. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. 2002; 30:207–10.
  44. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci U S A*. 2001; 98:31–6. <https://doi.org/10.1073/pnas.011404098>.
  45. Ycart B, Pont F, Fournie JJ. Curbing false discovery rates in interpretation of genome-wide expression profiles. *J Biomed Inform*. 2014; 47:58–61. <https://doi.org/10.1016/j.jbi.2013.09.004>.
  46. Czerny T, Busslinger M. DNA-binding and transactivation properties of Pax-6: three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5). *Mol Cell Biol*. 1995; 15:2858–71.
  47. Chauveau C, Decourt C, Cogne M. Insertion of the IgH locus 3' regulatory palindrome in expression vectors warrants sure and efficient expression in stable B cell transfectants. *Gene*. 1998; 222:279–85.

# PAX5-ELN oncoprotein promotes multistep B-cell acute lymphoblastic leukemia in mice

Laura Jamrog<sup>a,1</sup>, Guillaume Chemin<sup>b,1</sup>, Vincent Fregona<sup>a</sup>, Lucie Coster<sup>c</sup>, Marlène Pasquet<sup>d</sup>, Chloé Oudinet<sup>b</sup>, Nelly Rouquié<sup>a</sup>, Naïs Prade<sup>a,c</sup>, Stéphanie Lagarde<sup>a,c</sup>, Charlotte Cresson<sup>a</sup>, Sylvie Hébrard<sup>a</sup>, Ngoc Sa Nguyen Huu<sup>b</sup>, Marina Bousquet<sup>e</sup>, Cathy Quelen<sup>e</sup>, Pierre Brousset<sup>e</sup>, Stéphane J. C. Mancini<sup>f</sup>, Eric Delabesse<sup>a,c</sup>, Ahmed Amine Khamlichi<sup>b,2</sup>, Bastien Gerby<sup>a,2,3</sup>, and Cyril Broccardo<sup>a,2,3</sup>

<sup>a</sup>Centre de Recherches en Cancérologie de Toulouse (CRCT), Université de Toulouse, Institut National de la Santé et de la Recherche Médicale (INSERM), Université Toulouse III Paul Sabatier (UPS), 31037 Toulouse, France; <sup>b</sup>Institut de Pharmacologie et de Biologie Structurale (IPBS), Université de Toulouse, Centre National de la Recherche Scientifique (CNRS), UPS, 31077 Toulouse, France; <sup>c</sup>Laboratoire d'Hématologie, Centre Hospitalier Universitaire (CHU) de Toulouse, 31000 Toulouse, France; <sup>d</sup>Department of Pediatric Hematology, CHU de Toulouse, 31000 Toulouse, France; <sup>e</sup>CRCT, INSERM, UPS, ERL5294 CNRS, Laboratoire d'Excellence Toulouse Cancer (TOUCAN), 31037 Toulouse, France; and <sup>f</sup>Aix Marseille University, CNRS, INSERM, Institut Paoli-Calmettes (IPC), Centre de Recherche en Cancérologie de Marseille (CRCM), 13009 Marseille, France

Edited by Brian J. Druker, Oregon Health & Science University, Portland, OR, and approved August 29, 2018 (received for review December 14, 2017)

**PAX5** is a well-known haploinsufficient tumor suppressor gene in human B-cell precursor acute lymphoblastic leukemia (B-ALL) and is involved in various chromosomal translocations that fuse a part of PAX5 with other partners. However, the role of PAX5 fusion proteins in B-ALL initiation and transformation is ill-known. We previously reported a new recurrent t(7;9)(q11;p13) chromosomal translocation in human B-ALL that juxtaposed PAX5 to the coding sequence of elastin (ELN). To study the function of the resulting PAX5-ELN fusion protein in B-ALL development, we generated a knockin mouse model in which the PAX5-ELN transgene is expressed specifically in B cells. PAX5-ELN-expressing mice efficiently developed B-ALL with an incidence of 80%. Leukemic transformation was associated with recurrent secondary mutations on *Ptpn11*, *Kras*, *Pax5*, and *Jak3* genes affecting key signaling pathways required for cell proliferation. Our functional studies demonstrate that PAX5-ELN affected B-cell development in vitro and in vivo featuring an aberrant expansion of the pro-B cell compartment at the preleukemic stage. Finally, our molecular and computational approaches identified PAX5-ELN-regulated gene candidates that establish the molecular bases of the preleukemic state to drive B-ALL initiation. Hence, our study provides a new in vivo model of human B-ALL and strongly implicates PAX5 fusion proteins as potent oncoproteins in leukemia development.

B-cell acute lymphoblastic leukemia | engineered mouse models | PAX5 fusion proteins | leukemia initiation | oncogenic transformation

**B**-cell precursor acute lymphoblastic leukemia (B-ALL) is the most common pediatric cancer. B-ALL is characterized by a blockade of B-cell differentiation combined with an uncontrolled proliferation of blastic cells. Current chemotherapy is efficient at inducing long-term remission in childhood B-ALL, but the most common cause of treatment failure remains relapse that occurs in 15 to 20% of patients (1). The prognosis is even worse in adult B-ALL, as only 30% of adults achieve long-term disease-free survival (2).

B-cell development is initiated by the entry of hematopoietic progenitors into the B-cell lineage transcription program and the concomitant sequential rearrangement of Ig genes through V(D)J recombination, ultimately leading to the generation of immunocompetent plasma cells. B-cell development can be dissected into pre-pro-B, pro-B, pre-B, immature B, and mature B-cell populations corresponding to different stages of differentiation (3). PAX5 is critical from early stages of B-cell development up to mature B cells (4). B-cell differentiation is completely blocked at the pro-B stage in *Pax5* knockout mice, revealing its importance for early B lymphogenesis (5). Indeed, PAX5 plays a critical role in B-cell lineage commitment by activating the transcription of B cell-specific genes such as *CD19* and *BLK* and suppressing alternative lineage choices (6–8).

PAX5 is the main target of genetic alterations in B-ALL. Heterozygous deletions and loss-of-function mutations of PAX5

found in more than one-third of human B-ALL (9–11). These alterations result in loss of PAX5 expression and impairment of DNA-binding activity and/or transcriptional activity of PAX5. PAX5 is also rearranged in 2.6% of pediatric B-ALL cases, being fused to various fusion partners (9, 12–14). PAX5 translocations have been associated with a blockade of B-cell differentiation, as illustrated by PAX5-ETV6 and PAX5-FOXP1, which fuse the PAX5 paired domain to ETV6 and FOXP1 transcription factors, respectively (15).

We previously reported the molecular characterization of a new chromosomal t(7;9)(q11;p13) translocation in two cases of adult B-ALL. This translocation juxtaposed the 5' region of PAX5 and almost the entire sequence of elastin (ELN) (13). The resulting PAX5-ELN fusion protein had conserved the nuclear localization sequence (NLS) and the DNA-binding paired-box domain of PAX5, and could therefore act as a constitutive repressor of the residual wild-type PAX5. Indeed, similar to PAX5-ETV6 and PAX5-FOXP1, transient expression of PAX5-ELN displayed a dominant-negative

## Significance

Engineered mouse models of acute leukemia are critical to understanding the biological mechanisms by which a primary oncogene induces disease. While PAX5 fusion proteins are considered primary oncogenic events in B-ALL, their role in leukemia development is ill-known due to the lack of animal models. This report provides a novel and accurate in vivo model for B-ALL induced by PAX5-ELN fusion protein that establishes a preleukemic phase and recapitulates the key features of human disease, including acquired mutations in genes of the JAK/STAT and RAS/MAPK pathways. This study is of general interest, as it allows a better understanding of the biological mechanism by which an oncoprotein perturbs normal B-cell development and leads to pathological B-ALL.

Author contributions: L.J., G.C., V.F., L.C., M.P., C.O., N.R., N.P., S.L., C.C., S.H., N.S.N.H., M.B., C.Q., P.B., S.J.C.M., E.D., A.A.K., B.G., and C.B. designed research; L.J., G.C., V.F., L.C., M.P., C.O., N.R., N.P., S.L., C.C., S.H., N.S.N.H., M.B., C.Q., B.G., and C.B. performed research; M.P. provided pediatric B-ALL samples; L.J., G.C., V.F., L.C., M.P., C.O., N.R., N.P., S.L., C.C., S.H., N.S.N.H., M.B., C.Q., S.J.C.M., E.D., A.A.K., B.G., and C.B. analyzed data; and A.A.K., B.G., and C.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Published under the PNAS license.

<sup>1</sup>L.J. and G.C. contributed equally to this work.

<sup>2</sup>A.A.K., B.G., and C.B. contributed equally to this work.

<sup>3</sup>To whom correspondence may be addressed. Email: bastien.gerby@inserm.fr or cyril.broccardo@inserm.fr.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1721678115/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1721678115/-DCSupplemental).

effect on PAX5 by down-regulating the transcription of PAX5 target genes (11, 13, 16, 17). However, the generality of the dominant-negative effect was recently questioned *in vivo* (15).

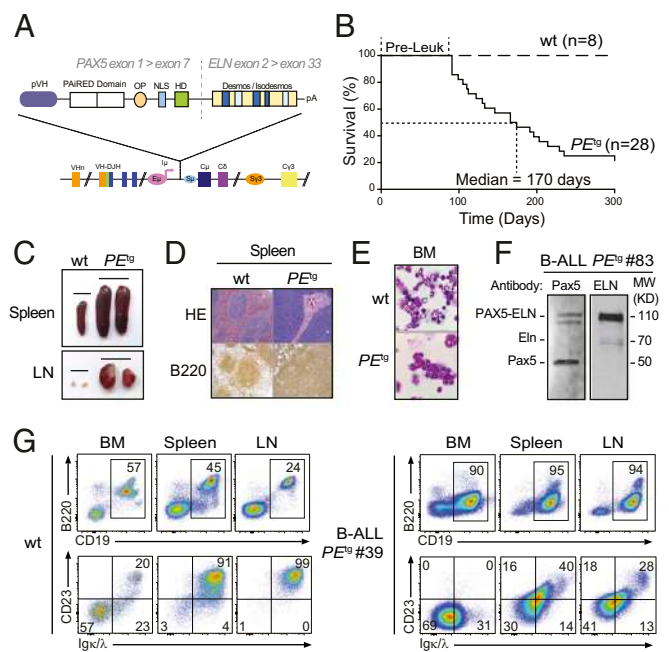
A detailed and dynamic analysis of the PAX5-ELN effect on B-ALL initiation and transformation requires the availability of a mouse model that allows a B-cell restricted expression of PAX5-ELN. Here, we demonstrate, by using a knockin (KI) mouse line, that PAX5-ELN is an initiating oncogenic event that perturbs normal B-cell development by aberrantly expanding pro-B cells at the earliest steps of leukemogenesis. Nonetheless, complete B-ALL transformation induced by PAX5-ELN involved the acquisition of additional mutations and the deregulated expression of multiple genes. Our study provides novel insights into the functional role of PAX5-ELN as a potent oncoprotein in B-ALL development, and should help to elucidate the molecular mechanisms that underlie B-ALL initiation and transformation.

## Results

**Constitutive Expression of Human PAX5-ELN Leads to B-ALL Development.** To investigate the B-leukemic potential of PAX5-ELN, we developed a KI mouse model in which the cDNA encoding the human PAX5-ELN fusion protein was expressed in B-cell progenitors. This was achieved by inserting the human cDNA at the *IgH* locus under the control of a VH promoter (PVH) and the endogenous E $\mu$  enhancer whose activity is triggered early in B-cell development (18). To avoid transcriptional readthrough from upstream promoters at different developmental stages, a pause/polyadenylation site (19) was added upstream of the ectopic PVH promoter (Fig. 1A and *SI Appendix*, Fig. S1). Unless otherwise indicated, the experiments have been performed on heterozygous mice.

Mutant mice that constitutively expressed PAX5-ELN (hereafter  $PE^{tg}$  mice) efficiently developed leukemia with a penetrance of 80% at 300 d (Fig. 1B). Leukemic development was associated with a splenomegaly (Fig. 1C, *Upper*) characterized by a massive infiltration of B220<sup>+</sup> cells and a dramatic perturbation of the spleen architecture (Fig. 1D). Moreover,  $PE^{tg}$  mice developed a lymphadenopathy (Fig. 1C, *Lower*) and exhibited blast cells in the bone marrow (BM) (Fig. 1E). Importantly, although the expression of PAX5-ELN was driven by *IgH* regulatory sequences, immunoblot analysis of protein extracts with a PAX5 paired domain-specific antibody revealed that the abundance of PAX5-ELN was not higher than that of endogenous PAX5 (Fig. 1F). This observation indicates that the reported effects on  $PE^{tg}$  mice cannot be ascribed to high expression levels of PAX5-ELN fusion protein. Immunophenotypic characterization showed that BM, spleen, and lymph nodes (LNs) of leukemic  $PE^{tg}$  mice contained an aberrant proportion of CD19<sup>+</sup> B cells associated with an aberrant and variable expression of CD23 and Ig $\kappa/\lambda$  markers (Fig. 1G and *SI Appendix*, Fig. S2).

**Clonal Transformation and Collaborating Events with PAX5-ELN Oncoprotein.** The *IgH* variable region can broadly be divided into the V<sub>H</sub> domain, including the distal V<sub>H558</sub> and the proximal V<sub>H7183</sub> gene families, and the D<sub>H</sub>J<sub>H</sub> domain, comprising a dozen D<sub>H</sub> segments followed by four J<sub>H</sub> segments (Fig. 2A, *Upper*) (20). Assembly of the *IgH* variable region involves two recombination steps: first D<sub>H</sub> to J<sub>H</sub>, followed by V<sub>H</sub> to D<sub>H</sub>J<sub>H</sub>. To determine the rearrangement status of the *IgH* locus in leukemic cells, we performed a qPCR-based V(D)J recombination assay (21) on genomic DNA purified from blasts of five independent B-ALL mice (Fig. 2A, *Upper*). The data reveal prominent D<sub>H</sub>J<sub>H</sub> and V<sub>H</sub>D<sub>H</sub>J<sub>H</sub> rearrangements, involving both proximal and distal VH segments, in all B-ALL samples (Fig. 2A, *Lower*). These results indicate that PAX5-ELN fusion protein induces the clonal transformation of a B-cell progenitor that has already rearranged its *IgH* locus, and support the notion that PAX5-ELN acts as a potent B-ALL oncoprotein.

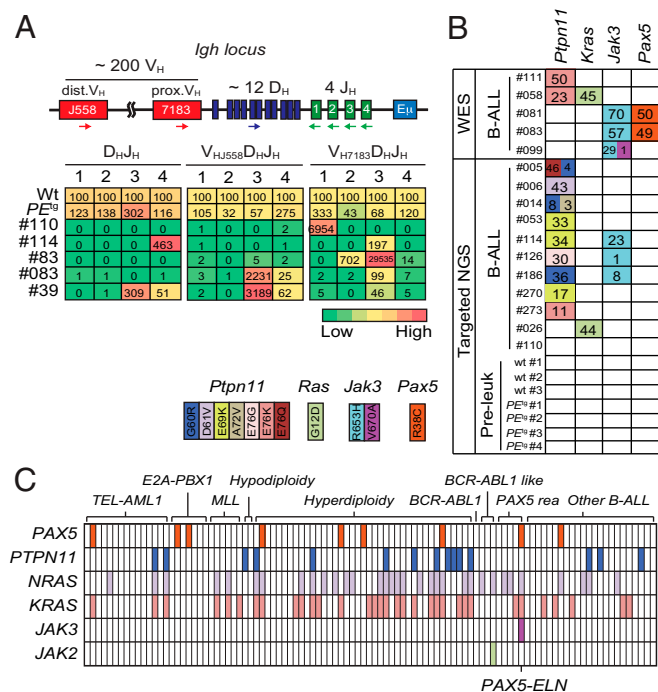


**Fig. 1.** Human PAX5-ELN expression induces efficient B-ALL development. (A) Generation of knockin mouse model expressing PAX5-ELN fusion protein. The sequence encoding the human PAX5-ELN fusion protein was inserted downstream of E $\mu$  enhancer. Desmos, Desmosine; HD, homeodomain; OP, octapeptide; pVH, VH gene promoter. (B) Kaplan-Meier curves of the time to leukemia for cohorts of  $PE^{tg}$  mice ( $n = 28$ ). WT mice ( $n = 8$ ) were used as controls. Pre-Leuk, preleukemic time. (C–E) PAX5-ELN induces B-ALL development characterized by leukemic cell invasion in the bone marrow, spleen, and lymph nodes. (C) Pictures of spleens (*Upper*) and LNs (*Lower*) from WT and leukemic  $PE^{tg}$  mice are shown. (D) Staining with hematoxylin and eosin (HE) and immunohistochemistry of B220 are shown of spleens from WT and leukemic  $PE^{tg}$  mice. (E) Pictures of May-Grünwald–Giemsa–stained cytosin of BM cells from WT and leukemic  $PE^{tg}$  mice. (F) Protein extract of leukemic cells from a B-ALL  $PE^{tg}$  mouse (no. 83) was subjected to immunoblotting with anti-N-terminal Pax5 (N19) antibody for the detection of PAX5-ELN and endogenous Pax5 and with anti-ELN antibody for the detection of PAX5-ELN and endogenous Eln. (G) Total cells from the BM, spleen, and LNs of WT (*Left*) and leukemic  $PE^{tg}$  (*Right*) mouse (no. 39) were immunophenotyped using the B220, CD19, CD23, and Ig $\kappa/\lambda$  markers.

Our observations on B-ALL transformation delay (>90 d) (Fig. 1B) led us to suspect potential acquisition of secondary mutations. To further identify these additional genetic alterations that potentially cooperate with PAX5-ELN, we performed whole murine exome sequencing in five  $PE^{tg}$  leukemias and identified recurrent somatic mutations in *Ptpn11*, *Kras*, *Jak3*, and *Pax5* genes. These mutations were also found after specific sequencing of these loci in an additional 11  $PE^{tg}$  leukemias (Fig. 2B). These findings strengthen the notion that aberrant activation of JAK/STAT and/or RAS/MAPK signaling pathways is required for an overt B-cell leukemia transformation in our model.

To address the critical question of the recurrence of these somatic mutations in human B-ALL, we performed the targeted sequencing of several exons of PAX5, PTPN11, NRAS, KRAS, JAK3, and JAK2 genes in a cohort of 101 pediatric B-ALL patients (Fig. 2C and *Dataset S1*). Importantly, we detected recurrent mutations in PAX5 (9/101), PTPN11 (15/101), NRAS (33/101), and KRAS (33/101) genes in all of the different B-ALL oncogenic subtypes (Fig. 2C). Interestingly, the patient cohort contained five PAX5-rearranged B-ALL patients, including one with the PAX5-ELN translocation that exhibited, just as in our mouse line, PAX5, NRAS, KRAS, and JAK3 mutations (Fig. 2C and *Dataset S1*). Together, the above data establish that our PAX5-ELN mouse model recapitulates the multistep pathogenesis of human B-ALL.



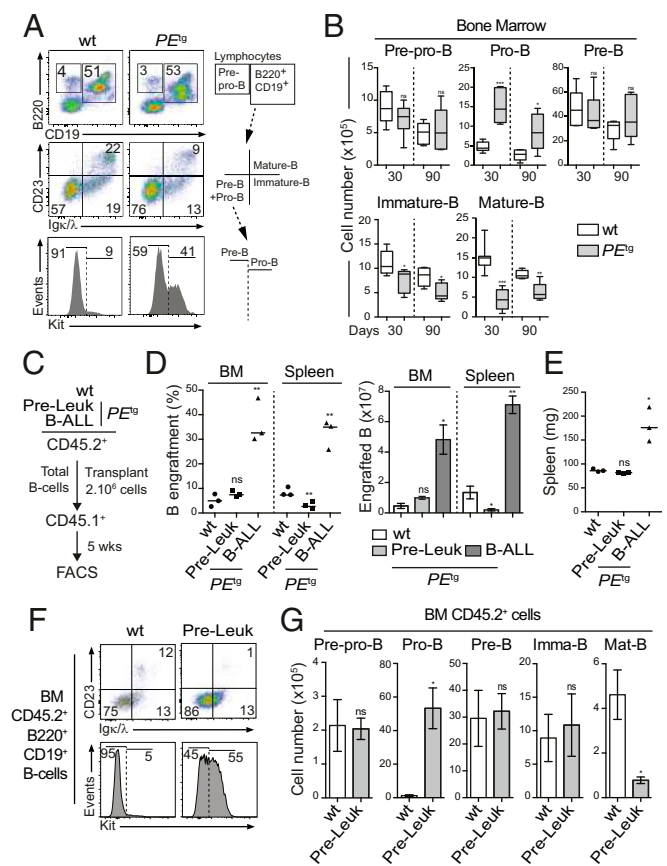


**Fig. 2.** Clonal selection in PAX5-ELN-induced B-ALL. (A) Schematic of the mouse *IgH* locus (Upper). The locus is composed of ~200 V<sub>H</sub> (red boxes), ~12 D<sub>H</sub> (blue), and 4 J<sub>H</sub> (green) segments that undergo recombination in B-cell precursors to produce a functional V<sub>H</sub>D<sub>H</sub>J<sub>H</sub> unit. Genomic DNAs were prepared from purified PE<sup>tg</sup> pro-B cells and five leukemic PE<sup>tg</sup> mice and subjected to quantitative PCR to quantify D<sub>H</sub>J<sub>H</sub> and V<sub>H</sub>D<sub>H</sub>J<sub>H</sub> rearrangements using primers that bind the indicated gene segments. dV<sub>H</sub>, distal V<sub>H</sub>; pV<sub>H</sub>, proximal V<sub>H</sub>. PCR of the H55 element downstream of the 3' regulatory region was performed for normalization of DNA input. PCR was performed in triplicate. Quantification of D<sub>H</sub>J<sub>H</sub> and V<sub>H</sub>D<sub>H</sub>J<sub>H</sub> rearrangements is represented as a heat map (Lower). DNA from purified WT pro-B cells was used as a control and normalization (100% of the signal) for each gene rearrangement ranked in each column. (B) Recurrent mutations in B-ALL blasts induced by PAX5-ELN. Whole-exome sequencing (WES) was performed on BM cells from five leukemic PE<sup>tg</sup> mice. Mutations found on *Ptpn11*, *Kras*, *Pax5*, and *Jak3* genes were verified and screened for recurrence by targeted next-generation sequencing (NGS) on BM cells from 11 leukemic, 3 WT, and 4 preleukemic 30-d-old PE<sup>tg</sup> mice. Each row represents a leukemia sample, and each column represents a genetic alteration. Colors indicate the position of the mutation, and numbers represent the variant allele frequency of each mutation. (C) Mutations on *PTPN11*, *KRAS*, *NRAS*, *JAK3*, *JAK2*, and *PAX5* genes were screened for recurrence on 101 B-ALL patient samples. Each column represents a B-ALL sample classified according to the indicated oncogenic subtypes, and each row represents a genetic alteration. Colored boxes indicate the presence of a mutation, listed in Dataset S1.

**PAX5-ELN Oncoprotein Perturbs Early B-Cell Progenitor Cell Differentiation at Preleukemic Stage.** The survival curve of PE<sup>tg</sup> mice revealed a leukemia onset starting at 90 d after birth (Fig. 1B). This leukemia development latency allowed us to explore the effect of PAX5-ELN oncoprotein on the early steps of leukemogenesis. Cytological examination of the preleukemic PE<sup>tg</sup> mice confirmed the absence of blasts in the BM (SI Appendix, Fig. S3A) and any obvious perturbation of the main cell lineages (SI Appendix, Fig. S3B). Moreover, the absence of detection of *Ptpn11*, *Kras*, *Pax5*, and *Jak3* mutations in the BM of preleukemic PE<sup>tg</sup> mice (Fig. 2B) allowed us to precisely address the role of PAX5-ELN in leukemia initiation, before the onset of clonal transformation induced by the acquisition of additional cooperating oncogenic events.

To identify the earliest cell types that are affected by PAX5-ELN oncoprotein, we analyzed B-cell progenitors (SI Appendix, Fig. S3C) in the BM of PE<sup>tg</sup> mice during early and late preleukemic stages, namely at 30 and 90 d after birth, respectively.

We observed that PAX5-ELN significantly induced a three- to fourfold expansion of the pro-B population at 30 and 90 d, whereas pre-pro-B and pre-B populations were not affected (Fig. 3A and B). Interestingly, 30-d-old PE<sup>tg</sup> pro-B cells exhibited a polyclonal profile for D<sub>H</sub>J<sub>H</sub> and V<sub>H</sub>D<sub>H</sub>J<sub>H</sub> gene rearrangements that was roughly comparable to their WT counterparts (Fig. 2A), indicating that PAX5-ELN-induced pro-B cell expansion was not associated with a clonal selection at the preleukemic stage. In addition, this pro-B cell expansion was associated with a reduction of the immature and circulating B-cell populations in the BM (Fig. 3A and B), indicating that PAX5-ELN partially blocked B-cell differentiation at the preleukemic stage. Accordingly, we observed a moderate but significant reduction of spleen size in preleukemic



**Fig. 3.** Pro-B cells are expanded by PAX5-ELN at a preleukemic phase. (A and B) Preleukemic PE<sup>tg</sup> pro-B cells exhibit an aberrant expansion potential in vivo. Immunophenotyping of BM cells from 30-d-old WT and PE<sup>tg</sup> mice was performed using B220, CD19, CD23, Igκ/λ, and Kit markers (A, Left), and a gating strategy was applied to discriminate B-cell subpopulations (A, Right and SI Appendix, Fig. S2). Absolute numbers of pre-pro-B, pro-B, pre-B, immature B, and mature B cells from the BM of 30- and 90-d-old WT and PE<sup>tg</sup> mice were calculated (B; n = 5 to 8 mice per condition). (C-E) Preleukemic and leukemic PE<sup>tg</sup> B cells exhibit different engraftment potentials. Total B cells from the BM of preleukemic and B-ALL PE<sup>tg</sup> mice (CD45.2<sup>+</sup>) were purified and i.v. transplanted in recipient mice (CD45.1<sup>+</sup>, n = 3, 2.10<sup>6</sup> cells per mouse) pretreated with 30 mg/kg busulfan (C). The same number of total B cells from the BM of WT mice was transplanted in parallel as control (C; n = 3). The proportion (D, Left) and absolute numbers (D, Right) of donor-derived (CD45.2<sup>+</sup>) B cells in the recipient BM were calculated 5 wk after transplantation. Spleen weights of engrafted mice were compared (E). (F and G) Preleukemic PE<sup>tg</sup> pro-B cells expand after transplantation. Immunophenotyping of engrafted B cells (CD45.2<sup>+</sup>B220<sup>+</sup>CD19<sup>+</sup>) in the BM was performed (F), and absolute numbers of donor-derived (CD45.2<sup>+</sup>) preleukemic pre-pro-B, pre-B, pro-B, immature B, and mature B cells in the recipient BM were calculated (G). Error bars represent standard deviations, \*\*\*P < 0.0005, \*\*P < 0.005, and \*P < 0.05; ns, nonsignificant.

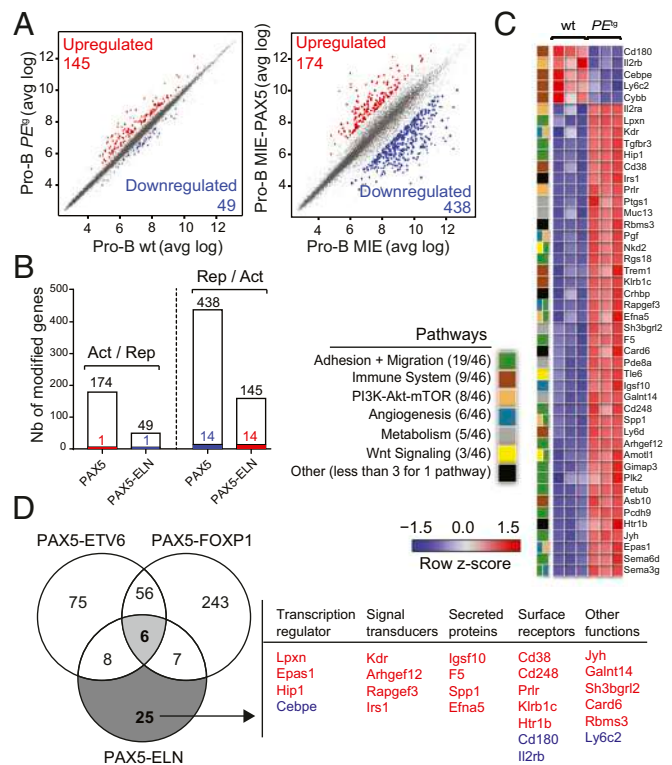
$PE^{tg}$  mice compared with age-matched WT mice (*SI Appendix, Fig. S4A*), suggesting that PAX5-ELN retained B-cell loading in peripheral lymphoid organs. Indeed, while the overall distribution of B-cell populations was unaffected by PAX5-ELN in the spleen (*SI Appendix, Fig. S4B*), the absolute number of immature and mature splenic B cells was significantly reduced in preleukemic  $PE^{tg}$  mice (*SI Appendix, Fig. S4C*). The above results indicate that PAX5-ELN oncoprotein expands preleukemic pro-B cells at the expense of subsequent stages of maturation in vivo. Finally, we confirmed that PAX5-ELN altered the intrinsic capacity of differentiation of pro-B cells using in vitro functional approaches, and had a restricted expression and effect on the B-cell lineage (*SI Appendix, Text and Figs. S5 and S6*).

**PAX5-ELN Oncoprotein Confers an Aberrant Expansion Potential to Preleukemic Pro-B Cells.** We first evaluated the effect of PAX5-ELN on B-cell progenitor turnover using a total BM transplantation assay. The data strongly suggested that PAX5-ELN induced an aberrant expansion potential to preleukemic pro-B/pre-B progenitors before B-ALL transformation (*SI Appendix, Text and Fig. S7 A–D*). To strengthen this notion, we compared the reconstitution potential of WT, preleukemic, and leukemic  $PE^{tg}$  B cells. Equal numbers of total B cells purified from the BM of WT, preleukemic, and leukemic  $PE^{tg}$  mice were transplanted into congenic mice (Fig. 3C). Five weeks after transplantation, engraftment efficiencies of WT and preleukemic  $PE^{tg}$  B cells were similarly low (Fig. 3D), and recipient mice did not exhibit splenomegaly (Fig. 3E). Interestingly, the analysis of engrafted B-cell compartments in the BM of recipient mice revealed an aberrant proportion of  $PE^{tg}$  donor-derived pro-B cells specifically ( $65 \pm 15$ -fold expansion) (Fig. 3F and G and *SI Appendix, Fig. S7E*). This striking expansion was not associated with the acquisition of *Ptpn11*, *Kras*, *Pax5*, or *Jak3* mutations (*SI Appendix, Fig. S7F*). In addition, consistent with our results in steady-state conditions, we observed a moderate but significant diminution of the engraftment efficiency of preleukemic  $PE^{tg}$  B cells in the spleen compared with WT controls (Fig. 3D). This was due to a reduction of the number of donor-derived immature and mature B cells (*SI Appendix, Fig. S7 G and H*).

Thus, our transplantation experiments with preleukemic B cells provide additional evidence to support that PAX5-ELN induces an aberrant self-renewal activity to normal pro-B cells before clonal and malignant transformation. On the other hand, recipient mice transplanted with clonal leukemic  $PE^{tg}$  cells exhibited a high level of engraftment in the BM (Fig. 3D) associated with an important blast invasion in the spleen (Fig. 3D and E), and with the presence of *Ptpn11* mutations (*SI Appendix, Fig. S7F*), characteristic of the spreading of a clonal population.

Combined, our functional approaches establish a sharp difference in the reconstitution potential between clonal and nonclonal B-cell populations in our model. Furthermore, they indicate that the pro-B cell compartment is abnormally expanded by PAX5-ELN oncoprotein at the preleukemic stage, and therefore represents the likely cellular target of leukemia initiation.

**Gene Regulation by PAX5-ELN in Pro-B Cells.** Based on transient transfection assays, PAX5 fusion proteins including PAX5-ELN, PAX5-ETV6, and PAX5-FOXP1 were thought to act as dominant-negatives by affecting the transcriptional activity of WT PAX5 (11, 13, 16, 17). However, recent evidence in vivo revealed that PAX5-ETV6 or PAX5-FOXP1 fusion proteins marginally modified the expression of PAX5 target genes (15, 22). To address this hypothesis in our in vivo model, we first compared the gene expression profiles of purified pro-B cells from preleukemic  $PE^{tg}$  mice and age-matched WT mice (*SI Appendix, Fig. S8A*). Transcriptome analysis identified 145 up-regulated and 49 down-regulated genes with an expression difference of more than 1.5-fold and an adjusted  $P$  value of  $<0.05$  in PAX5-ELN-expressing pro-B cells (Fig. 4A and *Dataset S2*). In



**Fig. 4.** (A) Scatter plot of gene expression differences between in vivo purified WT and  $PE^{tg}$  preleukemic pro-B cells (Left) and between ex vivo E17.5 fetal liver  $Pax5^{-/-}$  pro-B cells transduced with either MIE-PAX5 or MIE retroviral vectors (Right) based on three independent microarray experiments. The normalized expression data of individual coding genes (indicated by dots) were plotted as the average log ratio (avg log). Up- and down-regulated genes with an expression difference of  $>1.5$ - and  $>3$ -fold, and an adjusted  $P$  value of  $<0.05$ , are colored in red or blue, respectively. (B) Absence of a general dominant-negative effect of PAX5-ELN on ex vivo PAX5-regulated genes in pro-B cells. Comparison of PAX5-activated and PAX5-ELN-repressed genes (Left) and of PAX5-repressed and PAX5-ELN-activated genes (Right) in preleukemic pro-B cells. Overlap indicates that one gene was activated by PAX5 and repressed by PAX5-ELN and 14 genes of the PAX5-repressed genes were activated by PAX5-ELN as represented by colored bars. (C) Heat map displaying the differential expression of PAX5-ELN-activated (red) and -repressed (blue) genes in WT ( $n = 3$ ) and  $PE^{tg}$  ( $n = 3$ ) pro-B cells. The 46 PAX5-ELN-modified genes were selected on the basis of an expression difference of  $>2$ -fold ( $P < 0.05$ ) and for encoding a protein implicated in one of the indicated pathways. The expression value of each gene is visualized according to the indicated scale. The pathway annotation is shown (Left). (D) Venn diagram indicating the overlap between PAX5-ETV6-, PAX5-FOXP1- (15), and PAX5-ELN-modified genes in preleukemic pro-B cells, selected for an expression difference of  $>3$ -, 3-, and 2-fold, respectively (Left). PAX5-ELN-modified genes that were included neither in PAX5-ETV6 nor in PAX5-FOXP1 signatures are listed according to their biological functions as transcriptional regulators, signal transducers, secreting proteins, and surface receptors. PAX5-ELN-activated and -repressed genes are indicated in red and blue, respectively.

parallel, we established a list of ex vivo PAX5-modified genes by comparing the gene expression profiles of  $Pax5^{-/-}$  embryonic liver (E17.5) pro-B cells retrovirally transduced with either MIE or MIE-PAX5. This strategy led to the identification of 174 PAX5-up-regulated and 438 PAX5-repressed genes (Fig. 4A and *Dataset S3*). Interestingly, the gene set enrichment analysis of a well-established list of in vivo PAX5-regulated genes, arising from in vivo gene expression profiles of murine  $Pax5^{+/+}$  and  $Pax5^{-/-}$  pro-B cells (22), with our new ex vivo PAX5 gene signature demonstrates a global similarity between the two approaches (*SI Appendix, Fig. S8B*). We determined the overlap of

our ex vivo PAX5-activated genes with PAX5-ELN-repressed genes. This analysis revealed that only one gene was activated by PAX5 and repressed by PAX5-ELN. Conversely, only 14 of the PAX5-repressed genes were activated by PAX5-ELN in pro-B cells (Fig. 4B and Dataset S44). Closely similar results were observed when we used the well-established list of in vivo PAX5-regulated genes (22) (SI Appendix, Fig. S8C and Dataset S4B), further validating our approach. In addition, we observed by qPCR that ectopic expression of PAX5-ELN in preleukemic pro-B cells was not associated with the down-regulation of endogenous *Pax5* and its two common target genes *CD19* and *CD79a* (SI Appendix, Fig. S8D). Finally, we confirmed that the majority of the genes modified by PAX5-ETV6 and PAX5-FOXP1 in pro-B cells (15) did not overlap with our PAX5-regulated genes (SI Appendix, Fig. S8E and F and Dataset S4C and D). Together, similar to PAX5-ETV6 and PAX5-FOXP1 (15), our results indicate that, in vivo, PAX5-ELN does not generally antagonize the normal functions of PAX5 in preleukemic pro-B cells.

To gain insight into the molecular regulation of PAX5-ELN in the early steps of B-ALL development, we focused on the genes presenting a twofold change of expression when comparing preleukemic *PE*<sup>tg</sup> with WT pro-B cells. Forty-one activated and five repressed PAX5-ELN genes were identified in preleukemic *PE*<sup>tg</sup> pro-B cells (Fig. 4C and Dataset S2). To identify shared molecular programs induced by PAX5-ETV6, PAX5-FOXP1, and PAX5-ELN, we overlapped the lists of their targets in preleukemic pro-B cells (15) (Fig. 4D). Notably, we found only six common genes, including genes encoding for a signal transducer (*Gimap3*), a secreted protein (*Sema3g*), and surface receptors (*Il2ra* and *Trem1*) (SI Appendix, Fig. S9A). In contrast, we identified 25 PAX5-ELN-modified genes, which were modified neither by PAX5-ETV6 nor by PAX5-FOXP1 (Fig. 4D). Interestingly, these genes encoded four signal transducers, four secreted proteins, and seven surface receptors but also the four transcriptional regulators *Lpxn*, *Epas1*, *Hip1*, and *Cebpe* (Fig. 4D).

The above results suggest that PAX5-ELN predominantly regulates an independent molecular program in preleukemic pro-B cells. To provide additional support to this notion, we categorized PAX5-ELN-modified genes according to pathways and identified activated (act) and repressed (rep) genes coding for proteins involved in the following processes: adhesion and migration (19 act), immune system (5 act, 4 rep), PI3K-mTOR signaling (7 act, 1 rep), angiogenesis (6 act), metabolism (5 act), and Wnt signaling (3 act) (Fig. 4C). Most of these pathways are strongly deregulated in PAX5-ETV6-induced B-ALL cells compared with their preleukemic counterparts (15), suggesting that PAX5-ELN activates at the preleukemic stage a molecular program required for B-ALL development in *PAX5-ETV6*<sup>tg</sup> mice. Interestingly, we identified six genes (*Epas1*, *Arhgef12*, *Fetub*, *Rapgef3*, *Galnt14*, and *Cd248*) activated by PAX5-ELN in preleukemic pro-B cells that were shown to be specifically activated in B-ALL *PAX5-ETV6* cells (SI Appendix, Fig. S9B).

In conclusion, we have identified regulated PAX5-ELN genes in preleukemic pro-B cells with important functions in several signaling pathways and gene candidates that establish the molecular bases to drive B-ALL development.

## Discussion

Genome-wide profiling identified the *PAX5* gene as the most frequent target of somatic mutation in human B-ALL (11). Nonetheless, while *PAX5* acts as a main tumor suppressor gene in B-ALL, B-cell development is normal in heterozygous *Pax5*<sup>+/-</sup> mice (5, 23), unless primed by chemical or retroviral mutagenesis (24), suggesting oncogenic cooperation for B-ALL transformation. Indeed, the tumor suppressor functions of *Pax5* in B leukemogenesis were revealed under constitutive activation of STAT5, JAK1, and JAK3 (24, 25). *PAX5* is also involved in reciprocal translocations leading to the fusion of its N-terminal

domain with the C-terminal sequence of a second transcription factor such as ETV6 and FOXP1 (11, 26). In contrast with heterozygous *PAX5* deletions, which are considered to have a secondary role during late stages of leukemogenesis, it is generally assumed that *PAX5* fusion proteins act as primary oncogenic events altering normal B-cell development in the early steps of the disease (14). A wide diversity of fusion partners has been described (9, 14). Remarkably, they all conserve the N-terminal DNA-binding region of *PAX5* and the NLS but lack the potent C-terminal transcriptional regulatory domains, suggesting that the *PAX5* part of the chimeric protein is required for leukemogenesis.

The loss of *PAX5* transcriptional domains in the chimeric proteins led to the notion that the fusion proteins might act as constitutive repressors, as shown in various reporter assays (11, 13, 16, 27). Consequently, it was inferred that B-ALL might arise as a result of an antagonizing effect of the fusion proteins on *PAX5* activity. The situation may be more complex, as illustrated by the recent finding that *PAX5-ETV6* and *PAX5-FOXP1* do not actually modify the expression of *PAX5* target genes but rather implicate pre-B cell receptor, migration, and adhesion signaling pathways (15). *PAX5-ELN* rearrangement is unusual because it does not involve a transcription factor but an extracellular matrix protein (13). Furthermore, *PAX5-ELN* is one of the chimeric proteins that conserves the largest part of *PAX5*. Although *PAX5-ELN* blocked the transcription of *PAX5* target genes in transactivation assays (13), our results indicate that *PAX5-ELN* was able to repress only one of the *PAX5*-activated genes and to activate a small subset of the *PAX5*-repressed genes, strongly suggesting that *PAX5-ELN* does not generally antagonize the normal function of *PAX5* in preleukemic pro-B cells. Thus, in line with *PAX5-ETV6* and *PAX5-FOXP1* models (15), our in vivo data do not support the generality of a dominant-negative mechanism for *PAX5* fusion proteins.

Several hypotheses could be put forward to account for this discrepancy. One possible explanation could be merely quantitative. In heterozygous mice, thus mimicking the situation in patients, the concentration of the fusion protein may be below the threshold required to exert a dominant-negative effect, whereas this threshold may be readily achieved in transient reporter assays. As mentioned above, expression of a small subset of *PAX5*-regulated genes was impacted by *PAX5-ELN*, suggesting that there are instances in vivo where the dominant-negative effect is exerted by the oncoprotein. Alternatively, though not mutually exclusively, fusion proteins may be subjected to different regulatory mechanisms (half-life, posttranslational modifications, interacting partners, etc.) in vivo from the time lapse of transient expression. In this context, it should be noted that although the expression of *PAX5-ELN* cDNA was driven by the powerful *IgH* regulatory elements, *PAX5-ELN* levels were not higher than endogenous *PAX5* levels, excluding an inappropriate ectopic expression of the fusion protein in our model. Importantly, although the insertion of *Pax5* into the *IgH* locus led to T lymphomas (28), the weak expression of *PAX5-ELN* in thymocytes did not perturb T-cell development, underlying the restricted effect of *PAX5-ELN* expression on the B-cell lineage.

Interestingly, while *PAX5-ETV6* and *PAX5-FOXP1* arrested lymphopoiesis at the pro-B/pre-B cell transition, they did not induce leukemia on their own in the corresponding mutant mice (15), suggesting the requirement of cooperating mutations for B-ALL transformation. Specifically, in contrast to the *PAX5-ETV6* mouse model that required oncogenic cooperation (i.e., loss of the *Cdkn2a/b* tumor suppressor locus) for leukemia development (15), *PAX5-ELN* efficiently induced B-ALL in mice. In this context, most deregulated pathways in preleukemic *PAX5-ELN* pro-B cells were also strongly deregulated in *PAX5-ETV6*-induced B-ALL cells. Further analysis revealed that *PAX5-ELN* predominantly regulated an independent molecular



program in preleukemic pro-B cells, in which activation of six genes, also found activated in B-ALL PAX5-ETV6 cells (15), may explain the stronger potential of PAX5-ELN to induce B-ALL compared with PAX5-ETV6.

B cell-specific, constitutive expression of PAX5-ELN led to clonal transformation associated with the acquisition of additional mutations in key components of the JAK/STAT and RAS/MAPK pathways. These pathways represent two common targets of somatic mutations in other oncogene-induced B-ALL mice (29, 30) and in human B-ALL (11), including PAX5-ELN B-ALL cases (31, 32). Our analysis of a B-ALL patient cohort revealed the presence of *PAX5*, *PTPN11*, *NRAS*, *KRAS*, and *JAK3* mutations within different B-ALL oncogenic subtypes including PAX5-rearranged leukemias and a PAX5-ELN B-ALL case.

This study clearly demonstrated the effect of PAX5-ELN on leukemia initiation at the preleukemic stage, before malignant transformation. The partial blockade of differentiation induced by PAX5-ELN was associated with an aberrant expansion of pro-B cells, which was also revealed by transplantation assays. Importantly, this proliferative advantage was not associated with an obvious clonal selection and therefore suggests that PAX5-ELN confers self-renewal properties to pro-B cells, a situation that is reminiscent of the E2A-PBX1 mouse model (30). This observation supports the view that aberrant self-renewal activity of lymphoid progenitors induced by a primary oncogene can be an initiating event in leukemia development. This establishes a preleukemic stage setting genetic instability and accumulated genetic alterations that cooperate to lead to fully transformed B-ALL (33). Emerging evidence suggests that primary genetic alteration can convert normal committed progenitors into preleukemic stem cells by reprogramming aberrant self-renewal

properties both in lymphoid and myeloid lineages (33–35). Importantly, recent findings proposed that long-lasting preleukemic stem cells were hidden within the bulk of leukemic cells, and served as a reservoir for disease relapse (36, 37). An important challenge is to identify and target these leukemic initiating cells that represent an extremely rare subpopulation in patients. In this respect, the availability of the PAX5-ELN mouse model should be valuable. Additionally, by recapitulating the different steps of the human disease, the model represents a major opportunity to develop new therapeutic strategies on both B-ALL initiation and transformation, and a robust and reproducible tool for preclinical studies of drug screening and development.

## Experimental Procedures

Generation of the PAX5-ELN knockin mouse model, FACS analysis, Ig rearrangement assay, whole-exome sequencing and specific resequencing, transplantation assay, microarray experiment, retroviral transduction, co-culture of pro-B cells, RT-PCR, Western blot, and statistical analysis are described in *SI Appendix*.

**ACKNOWLEDGMENTS.** We acknowledge the cytometry and cell-sorting facility of the CRCT (INSERM U1037) and the IPBS animal facility for technical assistance. We are grateful to Manon Farcé and Laetitia Ligat from the facilities of the CRCT for assistance with flow cytometry and microscopy, respectively. We thank the Anexplo/Genotoul platforms for technical assistance (UMS006 and histology department). This research was supported by Institut National du Cancer (INCa) R11190BB-INCA 2011-131-PAX5, ARC Grant SFI20101201888, and FEDER Grant CITTIL. G.C., N.R., C.C., and N.S.N.H. were supported by INCa Grant R11190BB-INCA 2011-131-PAX5. The team is supported by association Laurette FUGAIN, Ligue nationale contre le cancer, “111 des arts,” Association Capucine, Société Française des Cancres de l’Enfant, and la région Occitanie. Work in the A.A.K. lab is supported by INCa, ANR, and Fondation ARC.

- Pui CH, Carroll WL, Meshinchi S, Arcenci RJ (2011) Biology, risk stratification, and therapy of pediatric acute leukemias: An update. *J Clin Oncol* 29:551–565.
- Armstrong SA, Look AT (2005) Molecular genetics of acute lymphoblastic leukemia. *J Clin Oncol* 23:6306–6315.
- Hardy RR, et al. (2000) B-cell commitment, development and selection. *Immunol Rev* 175:23–32.
- Adams B, et al. (1992) Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. *Genes Dev* 6:1589–1607.
- Urbánek P, Wang ZQ, Fetka I, Wagner EF, Busslinger M (1994) Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell* 79:901–912.
- Souabni A, Cobaleda C, Schebesta M, Busslinger M (2002) Pax5 promotes B lymphopoiesis and blocks T cell development by repressing Notch1. *Immunity* 17:781–793.
- Nera KP, et al. (2006) Loss of Pax5 promotes plasma cell differentiation. *Immunity* 24:283–293.
- Nutt SL, Heavey B, Rolink AG, Busslinger M (2015) Pillars article: Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature*. 401: 556–562. *J Immunol* 195:766–772.
- Nebral K, et al. (2009) Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia. *Leukemia* 23:134–143.
- Familiades J, et al. (2009) PAX5 mutations occur frequently in adult B-cell progenitor acute lymphoblastic leukemia and PAX5 haploinsufficiency is associated with BCR-ABL1 and TCF3-PBX1 fusion genes: A GRAALL study. *Leukemia* 23:1989–1998.
- Mullighan CG, et al. (2007) Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 446:758–764.
- Medvedovic J, Ebert A, Tagoh H, Busslinger M (2011) Pax5: A master regulator of B cell development and leukemogenesis. *Adv Immunol* 111:179–206.
- Bousquet M, et al. (2007) A novel PAX5-ELN fusion protein identified in B-cell acute lymphoblastic leukemia acts as a dominant negative on wild-type PAX5. *Blood* 109:3417–3423.
- Coyaud E, et al. (2010) Wide diversity of PAX5 alterations in B-ALL: A Groupe Francophone de Cyto-genétique Hematologique study. *Blood* 115:3089–3097.
- Smeenk L, et al. (2017) Molecular role of the PAX5-ETV6 oncoprotein in promoting B-cell acute lymphoblastic leukemia. *EMBO J* 36:718–735.
- Kurahashi S, et al. (2011) PAX5-PML acts as a dual dominant-negative form of both PAX5 and PML. *Oncogene* 30:1822–1830.
- Kawamata N, Pennella MA, Woo JL, Berk AJ, Koeffler HP (2012) Dominant-negative mechanism of leukemogenic PAX5 fusions. *Oncogene* 31:966–977.
- Kemp DJ, Harris AW, Cory S, Adams JM (1980) Expression of the immunoglobulin C mu gene in mouse T and B lymphoid and myeloid cell lines. *Proc Natl Acad Sci USA* 77:2876–2880.
- Haddad D, et al. (2011) Sense transcription through the 5' region is essential for immunoglobulin class switch recombination. *EMBO J* 30:1608–1620.
- Jung D, Giallourakis C, Mostoslavsky R, Alt FW (2006) Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol* 24:541–570.
- Braikia FZ, Chemin G, Moutahir M, Khamlichi AA (2014) Quantification of V(D)J recombination by real-time quantitative PCR. *Immunol Lett* 162:119–123.
- Revilla-I-Domingo R, et al. (2012) The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J* 31:3130–3146.
- Nutt SL, Heavey B, Rolink AG, Busslinger M (1999) Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* 401:556–562.
- Dang J, et al. (2015) PAX5 is a tumor suppressor in mouse mutagenesis models of acute lymphoblastic leukemia. *Blood* 125:3609–3617.
- Heltemes-Harris LM, et al. (2011) Ebf1 or Pax5 haploinsufficiency synergizes with STAT5 activation to initiate acute lymphoblastic leukemia. *J Exp Med* 208:1135–1149.
- Cazzaniga G, et al. (2001) The paired box domain gene PAX5 is fused to ETV6/TEL in an acute lymphoblastic leukemia case. *Cancer Res* 61:4666–4670.
- Fazio G, et al. (2013) PAX5/ETV6 alters the gene expression profile of precursor B cells with opposite dominant effect on endogenous PAX5. *Leukemia* 27:992–995.
- Souabni A, Jochum W, Busslinger M (2007) Oncogenic role of Pax5 in the T-lymphoid lineage upon ectopic expression from the immunoglobulin heavy-chain locus. *Blood* 109:281–289.
- van der Weyden L, et al. (2015) Somatic drivers of B-ALL in a model of ETV6-RUNX1; Pax5(+/-) leukemia. *BMC Cancer* 15:585.
- Duque-Afonso J, et al. (2015) Comparative genomics reveals multistep pathogenesis of E2A-PBX1 acute lymphoblastic leukemia. *J Clin Invest* 125:3667–3680.
- Mullighan CG, et al. (2009) JAK mutations in high-risk childhood acute lymphoblastic leukemia. *Proc Natl Acad Sci USA* 106:9414–9418.
- Denk D, Bradtke J, König M, Strehl S (2014) PAX5 fusion genes in t(7;9)(q11.2;p13) leukemia: A case report and review of the literature. *Mol Cytogenet* 7:13.
- Hong D, et al. (2008) Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. *Science* 319:336–339.
- Gerby B, et al. (2014) SCL, LMO1 and Notch1 reprogram thymocytes into self-renewing cells. *PLoS Genet* 10:e1004768.
- Krivtsov AV, et al. (2006) Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* 442:818–822.
- Shlush LI, et al.; HALT Pan-Leukemia Gene Panel Consortium (2014) Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature* 506:328–333, and erratum (2014) 508:420.
- Mullighan CG, et al. (2008) Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science* 322:1377–1380.

# PAX5-ELN oncoprotein promotes multistep B-cell acute lymphoblastic leukemia in mice

Jamrog et al.

## Supporting Information

### Supplemental text

#### PAX5-ELN oncoprotein alters the functional activity of pro-B cells *in vitro*.

Our *in vivo* experiments bring evidence that PAX5-ELN oncoprotein expands pre-leukemic pro-B cells at the expense of subsequent stages of maturation (**SI Appendix, Fig. 3A-B**), suggesting that PAX5-ELN perturbs their B-cell differentiation potential. To directly address this hypothesis at the functional level, we purified and expanded pro-B cells from 30 days old wt and pre-leukemic  $PE^{tg}$  mice (**SI Appendix, Fig. S5A**). Expanded wt and  $PE^{tg}$  pro-B cells were then cultured in undifferentiated (IL7+) or under B-cell differentiation (IL7-) conditions as described previously (1) (**SI Appendix, Fig. S5A**). Our results demonstrated that IL7 withdrawal allowed for an efficient differentiation of wt pro-B cells in pre-B ( $CD2^{+}Igk^{-}$ ) and immature-B ( $CD2^{+}Igk^{+}$ ) cells (**SI Appendix, Fig. S5B**). In contrast,  $PE^{tg}$  pro-B cells were not able to differentiate in the absence of IL7 (**SI Appendix, Fig. S5B**), further corroborating the partial blockade of differentiation observed *in vivo* (**SI Appendix, Fig. 3A-B**).

Pax5 is a critical B cell lineage commitment factor that restricts the developmental program of lymphoid progenitors to the B cell pathway. Precisely, Pax5 controls definitive B-cell commitment at the transition to pro-B cell stage (2, 3) and it is known that the loss of Pax5 affects the onset of fetal B cell development. Indeed, B lymphopoiesis is arrested prior to the appearance of  $CD19^{+}$  progenitors in the fetal liver of *Pax5*-deficient embryos (4). We therefore compared the ability of PAX5 and PAX5-ELN to rescue the blockade of B-cell commitment in *Pax5*<sup>-/-</sup> pro-B cells. To this end, fetal liver cells from *Pax5*<sup>-/-</sup> mice were retrovirally transduced with either MIE-PAX5, MIE-PAX5-ELN (MIE-PE) or with the empty (MIE) vectors and were cultured in undifferentiated (IL7+) or under B-cell differentiation (IL7-) conditions (**SI Appendix, Fig. S5C**). As expected, MIE-PAX5 induced CD19 expression in *Pax5*<sup>-/-</sup> pro-B cells in the presence of IL7 (**SI Appendix, Fig. S5D**). Interestingly, this induction was not associated with a downregulation of the Kit marker (**SI Appendix, Fig. S5D**), indicating that PAX5 activated its target gene *CD19* without inducing B-cell differentiation in the presence of IL7.



However, the withdrawal of IL7 revealed the ability of PAX5 to rescue the differentiation of *Pax5*<sup>-/-</sup> pro-B cells as shown by the down-regulation of Kit (**SI Appendix, Fig. S5D**). In stark contrast, PAX5-ELN modified neither CD19 nor Kit expression in the absence or presence of IL7 (**SI Appendix, Fig. S5D**).

The data demonstrate that PAX5-ELN fails to rescue B-cell development of *Pax5*-deficient pro-B cells.

### **PAX5-ELN oncoprotein has a restricted effect on the B-cell lineage**

B-cells are generated from HSCs and develop in the BM through several differentiation stages before they migrate to peripheral lymphoid organs. B-cell development can be divided into five cell compartments according to the expression of stage-specific cell-surface markers and the rearrangement status of their *Ig* loci (**SI Appendix, Fig. S3C**). We first quantified *PAX5-ELN* transcript levels in the five B-cell compartments and observed a significant induction of *PAX5-ELN* expression from the pro-B stage onwards (**SI Appendix, Fig. S6A**), which correlated with the E $\mu$ -mediated activation of the ectopic PV<sub>H</sub> promoter used to drive transgene expression.

To address the potential B cell-restricted expression and effect of PAX-ELN, we quantified *PAX5-ELN* transcript levels in thymocytes and analyzed T-cell development in the thymus of *PE*<sup>tg</sup> mice (**SI Appendix, Fig. S6B-D**). While *PAX5-ELN* transcripts were detected in *PE*<sup>tg</sup> thymocytes, their levels were manifold lower than those in *PE*<sup>tg</sup> BM B-cells (**SI Appendix, Fig. S6B**). In addition, the T-cell compartment was unaffected in mutant mice (**SI Appendix, Fig. S6C-D**), confirming the highly restricted effect of PAX5-ELN expression on the B-cell lineage.

PAX5 is also involved in V<sub>H</sub> to D<sub>H</sub>J<sub>H</sub> recombination step. Indeed, distal V<sub>H</sub> to D<sub>H</sub>J<sub>H</sub> recombination involving V<sub>H</sub>J558 genes are 100-fold reduced in *Pax5*-deficient pro-B cells (4). Here, our results indicated that pro-B cells expressing PAX5-ELN oncoprotein efficiently rearranged both their distal and proximal V<sub>H</sub>D<sub>H</sub>J<sub>H</sub> genes (**Fig. 2A**). Moreover, allelic competition between wt and mutant *IgH* loci enabled us to show that IgM expression almost exclusively originated from the wt allele in *PE*<sup>tg</sup> mice (**SI Appendix, Fig. S6E**), excluding that PAX5-ELN perturbs B-cell development by interfering with the Pax5-dependent regulation of *IgH* rearrangements.

Combined, the above data allow us to conclude that PAX5-ELN oncoprotein perturbs pro-B cell functional activity *in vitro* essentially by preventing their cell differentiation.

### **Effect of PAX5-ELN on B-cell progenitor turnover *in vivo*.**

To evaluate the effect of PAX5-ELN oncoprotein on B-cell progenitor turnover *in vivo*, BM cells from 30 days old wt and pre-leukemic  $PE^{tg}$  mice were transplanted into myeloablated recipient mice and B cell reconstitution was assessed after one month (**SI Appendix, Fig. S7A**). Analysis of BM showed similar engraftment efficiencies between wt and pre-leukemic  $PE^{tg}$  donor cells (**SI Appendix, Fig. S7B**). However, analysis of B-cell lineage development in donor-derived wt and  $PE^{tg}$  BM cells revealed that PAX5-ELN induced an aberrant reconstitution of pro-B and pre-B cell compartments (**SI Appendix, Fig. S7C**). Indeed, we observed a 4 to 5-fold expansion of donor-derived  $PE^{tg}$  pro-B and pre-B cells as compared to wt controls, while the other B-cell compartments were not affected (**SI Appendix, Fig. S7C**). In addition, we verified that T-cell reconstitution of donor-derived  $PE^{tg}$  thymocytes was not affected (**SI Appendix, Fig. S7D**), confirming the restricted effect of PAX5-ELN on the early B-cell compartment.

Together, these observations suggest that PAX5-ELN oncoprotein induce an aberrant expansion potential to a pro-B/pre-B progenitor before B-ALL transformation.

## Supplemental experimental procedures

### Generation of the PAX5-ELN knock-in mouse model

The vector (pAp- $\mu$ ) used to insert the human PAX5-ELN cDNA has been previously described in detail (5). The human cDNA was first inserted between a mouse PV<sub>H</sub> promoter and SV40 polyA signal in a modified pBluescript vector, and the whole unit (PVH-cDNA-SV40 polyA) was eluted as a ClaI fragment and inserted between the pause/polyA signal and *Neo<sup>R</sup>* cassette of pAp- $\mu$ . An HSV *tk* gene was inserted into Sall site for negative selection. The ES cell line CK35 (kindly provided by C. Kress, Institut Pasteur, Paris) was transfected by electroporation, and selected using G418 (300  $\mu$ g/ml) and gancyclovir (2  $\mu$ M). Recombinant clones were identified by PCR using a forward primer in the *Neo<sup>R</sup>* cassette and a reverse primer downstream of the 3' arm of the targeting vector. Two ES clones showing homologous recombination were injected into C57Bl/6 blastocysts, the male chimeras were then mated with C57Bl/6 females. Germ-line transmission of the mutation was checked by PCR. Homozygous mutant mice were mated with EIIa-cre transgenic mice. The progeny was checked by PCR for CRE-mediated deletion of *Neo<sup>R</sup>* cassette. The experiments on mice have been carried out according to the CNRS and Inserm Ethical guidelines and approved by the Regional Ethical Committee.

### FACS analysis, antibodies and cell sorting

Single cell suspensions were prepared from bone marrow (BM), spleen and lymph node (LN) of wt and *PE<sup>tg</sup>* mice of the indicated ages. Immunostaining was performed using antibodies used for flow cytometry obtained from Pharmingen (BD Biosciences) : V500-labelled anti-B220 (RA3-6B2), AF700-labelled anti-CD19 (1D3), FITC-labelled anti-Ig $\kappa$  (187.1), FITC-labelled anti-Ig $\lambda$  (R26-46), PeCy7-labelled anti-CD23 (B3B4), BV711-labelled anti-c-Kit (2B8), PE-labelled anti-CD2 (RM2-5), APC-Cy7-labelled anti-CD45.1 (A20) and V450-labelled anti-CD45.2 (104). Dead cells were excluded by 7-AAD staining. FACS analysis was performed on a Fortessa cytometer (BD Biosciences) using DIVA (BD Biosciences) and FlowJo softwares.

BM cells from wt or pre-leukemic *PE<sup>tg</sup>* mice were flushed and B-cell fraction were enriched by B220 magnetic microbead sorting (Miltenyi Biotec). B-cell fraction was subsequently incubated with CD19, B220, CD23, Ig $\kappa$ , Ig $\lambda$  and c-Kit antibodies. Cell sorting of B cell populations was performed on a MoFlo Astrios sorter (Beckman Coulter).

### Immunoglobulin rearrangement assay

Pro-B cells were enriched from the BM of 30 days old wt or pre-leukemic  $PE^{tg}$  mice by using B220 magnetic microbeads and LS columns (Miltenyi) and purified by sorting the  $B220^{+}Ig\kappa\lambda^{-}CD23^{-}Kit^{+}$  fraction. B-ALL cells were obtained from the BM of leukemic  $PE^{tg}$  mice. Genomic DNAs were prepared by standard techniques and subjected to an *IgH* V(D)J recombination assay by quantitative PCR using forward primers binding to the distal  $V_{HJ558}$ , the proximal  $V_{H7183}$  or the  $D_H$  segments and reverse primers that pair 3' of each  $J_H$  segment. PCR of HS5 element downstream of the 3' regulatory region was performed for normalization of the DNA input. Primers are listed in **SI Appendix, Table S5**.

### Whole-exome sequencing and specific resequencing

Whole-exome sequencing was performed using 50 ng genomic DNA extracted using a Wizard Genomic DNA Purification Kit (Promega, Madison, USA). Enrichment for coding exons was performed using a transposon-based method (SureSelect QXT Clinical research Exome, Agilent Technologies). The library was quantified by quantitative PCR (Kapa Biosystems, Wilmington, USA) and normalized at 4 nM. Sequencing was performed using NextSeq500 High Output kit V2 flowcell on a NextSeq 500 (Illumina) with a 150-bp paired-end mode. Sequence analysis was performed using NextGene software (SoftGenetics, State College, USA). For mutations resequencing, a library was generated using PCR from 7 specific mouse regions identified by whole exome sequencing. The library was sequenced on a MiSeq sequencer (Illumina, San Diego, USA) using a Miseq Reagent kit V2 (paired-end sequencing 2 x 150 cycles). Alignment and variant calling were performed using NextGene software (SoftGenetics, State College, USA). Targets and primers are listed in **SI Appendix, Table S5**.

### Human B-ALL mutation analysis

101 bone marrow samples (99 pediatric and 2 additional adult cases with a PAX5 rearrangement) treated for de novo B-ALL in the Toulouse hospital were analyzed regarding the mutations identified in the whole exome analysis of leukemic PAX5-ELN mice. Cytogenetic analysis was performed using R- or G-banding and classified according to WHO 2016 classification. JAK2 (exon 16), JAK3 (exons 13 and 15), KRAS (exons 2 and 3), NRAS (exons 2 and 3), PAX5 (exons 1 to 10) and PTPN11 (exons 3 and 11) were amplified by multiplex PCR to generate

specific libraries using primers listed in the supplementary data. The library was sequenced using a MiSeq sequencer (Illumina, San Diego, USA) and Miseq Reagent kit V2 (paired-end sequencing 2 x 150 cycles). Alignment and variant calling were performed using NextGene software (SoftGenetics, State College, USA). The study on human B-ALL patients was conducted in accordance with the Declaration of Helsinki and approved by the national institutional review board and ethical committee. Written informed consent was obtained from the parents. Targets and primers are listed in **SI Appendix, Table S5**.

#### Transplantation assay

Total BM cells and total purified B-cells from wt, pre-leukemic and B-ALL  $PE^{tg}$  mice (CD45.2<sup>+</sup>) were transplanted intravenously into 6-8 weeks old recipient mice (CD45.2<sup>+</sup>) pre-treated, 24h before transplantation, with 30 mg/kg of Busulfan (Busilvex, Pierre Fabre). Chimerism was analyzed by flow cytometry (FACS) and was illustrated by the percentage of donor-derived cells (% CD45.2<sup>+</sup>) found in the recipient BM and thymus.

#### Microarray experiments

$Pax5^{-/-}$  embryonic liver (E17.5) pro-B cells were transduced with either MIE or MIE-PAX5 as described previously. Transduced cells were purified according to their positivity for GFP fluorescence with a FACS Aria II (BD Biosciences) cell sorter. For each condition, three independent transductions were performed. Total RNAs were extracted using the Trizol method according to manufacturer's instruction (Invitrogen). RNA quality was assessed using the RNA 6000 Nano Assay on the Agilent 2400 Bioanalyser (Agilent Technologies, Massy, France). RNA samples were purified and prepared according to the manufacturer's protocol with the Affymetrix's GeneChip Whole Transcript Sense Target Labeling Assay Kit (Affymetrix, UK) and hybridized on Affymetrix GeneChip Mouse Gene 2.0 ST arrays. Pro-B cells were purified from the BM of 30 days old wt (n=3) and pre-leukemic  $PE^{tg}$  (n=3) mice by cell sorting. RNAs extraction and preparation were performed as described previously and hybridized on Clariom D GeneChip Mouse Transcriptome array 1.0.

#### Bioinformatics analysis

Probe-signal intensities were normalized and summarized by the robust multi-array average

(RMA) method by using Affymetrix's Expression Console software. Data were then processed with AltAnalyzer software (6). Folds changes of genes were calculated as the ratio of the signal values of the experimental groups compared with the control groups and principal component analysis (PCA) was calculated considering the varying genes of each sample. Gene expression between *in vivo* purified wt and  $PE^{tg}$  pre-leukemic pro-B cells and between *ex vivo* E17.5 fetal liver  $Pax5^{-/-}$  pro-B cells transduced with either MIE-PAX5 or MIE retroviral vectors are considered significantly modified when expression difference is upper 1.5 and 3 fold changes, respectively, and an adjusted  $P$ -value of  $< 0.05$ . Heatmap was established by  $\log^2$  genes expression confrontation for each gene of different groups, and clusturised hierarchically. The biological processes were represented as GO Term annotations provided by the data libraries from AltAnalyzer software. Gene enrichment analysis was performed using GSEA software from the Broad institute (7). Our *ex vivo* PAX5-modified genes were ranked based on  $\log^2$  ratio of classes and compared to a well-established gene set of *in vivo* PAX5-regulated genes in pro-B cells extracted from Revilla IDR, *et al.* (2012), Table S5 (8). PAX5-ELN-modified genes in pre-leukemic pro-B cells were overlapped with gene sets of PAX5-ETV6- and PAX5-FOXP1-regulated genes in pre-leukemic pro-B cells and in B-ALL cells extracted from Smeenk L, *et al.* (2017), Tables EV1 and EV4 (9).

#### Retroviral production and transduction

Full-length human  $PAX5$  and  $PAX5-ELN$  cDNA were amplified, sequenced and inserted into the retroviral vector pMSCV-IRES-EGFP (MIE). Retroviral supernatants were produced using the Phoenix retrovirus producer line (Orbigen) and Lipofectamine 2000 (Invitrogen) reagent for transfection and proceeding according to the manufacturer's instructions. After overnight incubation of the transfection mix, cells were treated with 10 mM sodium butyrate (Millipore) and washed with PBS before pro-B-cell medium was added. Viral supernatants were harvested after 24 h of incubation at  $32^{\circ}\text{C}$ , passed on a  $0.45\ \mu\text{m}$  filter, aliquoted and frozen at  $-80^{\circ}\text{C}$ .  $Pax5^{-/-}$  deficient pro-B-cells were harvested from  $Pax5^{-/-}$  embryonic liver at E17.5, amplified on irradiated OP9 stromal cells in the presence of 2 ng/mL IL7 and infected during exponential growth. Retroviral supernatant was diluted at 1:3, supplemented with  $4\ \mu\text{g/ml}$  polybrene (Sigma) and 2 ng/mL of IL7. Spinoculation was performed at 1,000 g at  $32^{\circ}\text{C}$  for 90 min. After centrifugation, 1mL of fresh medium supplemented with IL7 is added and the cells are incubated

overnight at 37°C. After infected pro-B-cell expansion for few days, cells are washed out of IL-7 and plated in IMDM, 5% FCS, 0.05 mM β-mercaptoethanol and 2 mM L-glutamine on irradiated OP9-derived stromal cells with or without IL7. After 3 days, single-cell suspensions were stained using standard protocols and subjected to flow cytometry.

#### Cytology and immunohistochemistry

BM cells from wt, pre-leukemic and leukemic  $PE^{tg}$  mice were cytopinned and stained with May Grünwald-Giemsa coloration. Splens from wt or leukemic  $PE^{tg}$  mice were fixed with PBS and 4% paraformaldehyde overnight at 4°C, processed and embedded in paraffin. Hematoxylin and eosin (H&E) staining was performed on 5µm paraffin sections according to common methods. 5µm sections were immunostained with biotinylated anti-B220 (Serotech) antibody and secondary antibody conjugated with the streptavidin, and visualized with the video-microscope Cell Observer (Zeiss).

#### RT-qPCR

RNA was isolated using Trizol method and cDNA was synthesized using SuperScript® VILO™ cDNA Synthesis Kit (Invitrogen™) according to the manufacturer's instructions. Quantitative SYBR Green PCR was performed on a LightCycler®480 II System (Roche) to quantify human *PAX5-ELN* and murine *Pax5*, *CD19* and *CD79a* cDNAs expression in purified pre-proB, pro-B, pre-B and immature-B cells from the BM of 30 days old wt and pre-leukemic  $PE^{tg}$  mice using LightCycler®480 SYBR Green I Master (Roche Diagnostics GmbH) according to the manufacturer's instructions. Targets and primers are listed in **SI Appendix, Table S5**. All PCR were carried out as follows in a 20 µl volume: 5 min at 95°C, followed by 45 cycles of 10 s at 95°C, 10 s at annealing temperature of 60°C and 10 s at 72°C. Quantification was performed using the  $\Delta C_t$  method with normalization to the *Abl1* gene expression levels. Data were analyzed using the LC480 software (Roche Diagnostics).

#### Western blot

Protein extraction from splenocytes of 30 days old wt and pre-leukemic  $PE^{tg}$  mice was performed using RIPA buffer (20 mM Tris pH 8.0, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% Sodium Desoxycholate). The concentrations of protein extracts were determined using Bicinchoninic

Acid Kit (Sigma). Proteins were then separated on sodium dodecyl sulfate–PAGE (SDS-PAGE) and transferred on poly-vinylidene difluoride (PVDF) membranes (Millipore Corporation). Immunodetection was revealed using ECL Plus (Amersham Biosciences). The membranes were blocked and the antibodies were diluted in 5% dried milk in TBST (10 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.05% Tween-20). The goat anti-PAX5 (N19, sc-1975) and the rabbit anti-ELN (H300, sc25736) antibodies were obtained from Santa Cruz and diluted 1/500. The secondary antibodies against rabbit IgG(H+L)-HRP conjugate and goat IgG(H+L)-HRP conjugate were obtained from Cell Signaling (Danvers, MA, USA) and diluted at 1/5000.

#### Culture of BM pro-B cells

Purified pro-B cells were amplified on 30 Gy irradiated OP9 stromal cells in Iscove's Modified Dulbecco's Medium (IMDM; Gibco), supplemented with 5 % Fetal Bovine Serum (FBS) (Stemcell technologies), 2 ng/mL murine IL7 (Peprotech), 0.05 mM  $\beta$ -mercaptoethanol (Sigma-Aldrich), 2 mM L-glutamine (Invitrogen), penicillin (100 units/ml) and streptomycin (100 units/ml). Experiments were performed after an initial pro-B-cell expansion on irradiated OP9 cells in the presence of 2 ng/ml IL7. Every 3 days, pro-B-cells were harvested and propagated on fresh OP9 stromal cells.

#### Statistical analysis

Student's *t* test was used for comparison of quantitative variable ( $***p < 0.0005$ ,  $**p < 0.005$ ,  $*p < 0.05$ ), assuming normality and equal distribution of variance between the different groups analyzed. Survival in mouse experiments was represented with Kaplan-Meier curves (Prism GraphPad).



## Supplemental Figure legends

**Fig. S1. (A-B)** The germline transmission of the *PAX5-ELN* insertion and the genotype of mice were determined by PCR using primers that specifically amplify the wild type or the mutant allele.

**Fig. S2.** Representative immunophenotype of B-ALL induced by *PAX5-ELN* in mouse #83. Total cells from BM, spleen and LN of wt and leukemic *PE<sup>tg</sup>* mice were immunophenotyped using the B220, CD19, CD23 and Igκ/λ markers.

**Fig. S3. (A-B)** *PAX5-ELN* does not affect the hematopoietic compartment distribution during the pre-leukemic phase. Representative pictures of May Grünwald-Giemsa-stained cytospin of BM cells from 30 days old wt and pre-leukemic *PE<sup>tg</sup>* mice (A). Cells from lymphoid, erythroid and myeloid lineages were counted (n=300) and the proportion of cell lineages was compared between 30 days old wt and pre-leukemic *PE<sup>tg</sup>* BM (B). (C) Schematic diagram of B-cell differentiation with representation of cell surface markers and rearrangement status of *Ig* genes that correlate with cell differentiation stages.

**Fig. S4. (A-C)** *PAX5-ELN* allows for normal B-cell maturation but reduces the number of immature-B and mature-B cell population. The size of the wt and *PE<sup>tg</sup>* spleens was compared from 30 and 90 days old mice (A). Immunophenotype of splenocytes from 30 days old wt and *PE<sup>tg</sup>* mice was performed using B220, CD19, CD23 and Igκ/λ markers (B) and absolute numbers of immature-B and mature-B cells were calculated (C, n=5 to 8 per condition).

**Fig. S5. (A-B)** *PAX5-ELN* prevents B-cell differentiation of pre-leukemic pro-B cells. Pro-B cells were purified from the BM of 30 days old wt and *PE<sup>tg</sup>* mice and expanded *in vitro* in undifferentiated condition on OP9 stromal cells in the presence of IL7 (OP9 IL7+). B-cell differentiation was induced or not (OP9 IL7+) by culturing expanded pro-B cells in the absence of IL7 (OP9 IL7-) during 3 days (A). Igκ and CD2 expression was assessed by flow cytometry. Representative FACS profiles are shown (B) (n=4). (C-D) *PAX5-ELN* does not rescue the B-cell commitment blockade induced by the absence of Pax5. E17.5 Fetal liver cells from *Pax5<sup>-/-</sup>* embryos were transduced with either MIE-PAX5 or MIE-PAX5-ELN (MIE-PE) or MIE

retroviral vectors. Transduced cells were then cultured in B-cell differentiation condition (OP9 IL7-) or not (OP9 IL7+) during 3 days (C). The proportion of Kit and CD19 positive cells was calculated by FACS for each condition (D) (n=3).

**Fig. S6.** (A) Pre-pro-B, pre-B, proB and immature-B cells were purified from the BM of 30 days old pre-leukemic  $PE^{tg}$  mice by cell sorting and mRNA levels of human  $PAX5-ELN$  in each population were determined by qRT-PCR and normalized to murine  $Abl1$ . (B) mRNA levels of human  $PAX5-ELN$  (left panel) and murine  $Pax5$  (right panel) were determined by qRT-PCR in thymocytes and purified total B-cells from two 30 days old wt and pre-leukemic  $PE^{tg}$  mice and were normalized to murine  $Abl1$ . (C-D)  $PAX5-ELN$  does not affect T cell development at the pre-leukemic stage. Immunophenotype of thymocytes from 30 days old wt and  $PE^{tg}$  mice was performed using CD4 and CD8 markers and a gating strategy was applied to discriminate double negative (DN), double positive (DP), simple positive SP4 and SP8 subpopulations (C). Absolute numbers of DN, DP, SP4 and SP8 thymocytes were calculated for each genotype (D, n=4 mice per condition). (E) Allelic competition assay showing that the insertion of  $PAX5-ELN$  cDNA in the  $IgH$  locus does not impair  $V_HD_HJ_H$  gene recombination in *trans*, and leads to normal B cell development driven essentially by IgM from the wt allele. The mutant allele originates from 129Sv and the wt allele from C57BL/6 following breeding of  $PE^{tg}$  mice with C57BL/6 wt mice. IgM allotypes were determined by FACS using monoclonal antibodies specific for IgMa (129S2/Sv) and IgMb (C57BL/6) allotypes.

**Fig. S7.** (A-D) BM transplantation reveals pro-B and pre-B cell expansion induced by  $PAX5-ELN$  but does not perturb thymic reconstitution. Pre-leukemic total BM  $PE^{tg}$  cells ( $CD45.2^+$ ) were intravenously transplanted in recipient mice ( $CD45.1^+$ , n=5,  $5 \cdot 10^6$  cells/mouse) pre-treated with 30mg/kg Busulfan (A). The same number of wt BM cells were transplanted in parallel as control (A, n=2). The percentage of donor-derived cells ( $CD45.2^+$ ) from the BM of recipient mice were analysed by FACS 1 month after transplantation (B). Absolute numbers of donor-derived ( $CD45.2^+$ ) pre-pro-B, pre-B, pro-B, immature-B and mature-B cells in the recipient BM were calculated (C). Absolute number of donor-derived DN, DP, SP4 and SP8 thymocytes were calculated in the thymus of recipient mice (D). (E) Fold expansions of donor-derived ( $CD45.2^+$ ) pre-pro-B, pre-B, pro-B, immature-B and mature-B cells in the BM of recipient mice were

calculated after transplantation of total B-cells from wt and pre-leukemic  $PE^{tg}$  mice described in Fig. 3C. (F) The presence of *Ptpn11*, *Kras*, *Pax5* and *Jak3* mutations were analysed by targeted NSG in the BM of recipient mice transplanted with total B-cells from wt and pre-leukemic and B-ALL  $PE^{tg}$  mice (n=3 for each condition). Each column represents a leukemia sample and each row a genetic alteration. (G-H) Immunophenotype of engrafted B-cells ( $CD45.2^+B220^+CD19^+$ ) in the spleen was monitored after transplantation of total B-cells from wt and pre-leukemic  $PE^{tg}$  mice described in Fig. 3C (G). Absolute number of engrafted immature-B and mature-B cells in the spleen of recipient mice were calculated (H).

**Fig. S8.** (A) Principal component analysis of varying genes of wt (blue, n=3) and  $PE^{tg}$  (red, n=3) pro-B cells. Only the varying genes of each sample were considered to calculate the principal components. (B) Gene set enrichment analysis (GSEA) of the well-established list of *in vivo* PAX5 regulated genes (8) with our new *ex vivo* PAX5 gene signature in pro-B cells revealing that many *in vivo* PAX5 target genes were similarly regulated in *ex vivo* PAX5 transduced pro-B cells. (C) Absence of a general dominant-negative effect of PAX5-ELN on established *in vivo* PAX5 regulated genes in pro-B cells (8). Comparison of PAX5-activated and PAX5-ELN-repressed genes (*left panel*) and of PAX5-repressed and PAX5-ELN-activated genes (*right panel*) in pre-leukemic pro-B cells. Overlap indicates that two genes were activated by PAX5 and repressed by PAX5-ELN and 13 genes of the PAX5-repressed genes were activated by PAX5-ELN as represented by colored bars. (D) Pre-leukemic pro-B cells were purified from the BM of three 30 days old wt and three pre-leukemic  $PE^{tg}$  mice and mRNA levels of human *PAX5-ELN*, murine *Pax5* and the *Pax5* target genes *CD19* and *CD79a* were determined by qRT-PCR and normalized to murine *Abi1*. (E-F) PAX5-ETV6 and PAX5-FOXP1 do not antagonize normal PAX5 gene signature. Comparison of PAX5-activated and PAX5-ETV6-repressed genes (E, *left panel*) and of PAX5-repressed and PAX5-ETV6-activated genes (E, *right panel*) in pro-B cells (9). Overlap indicates that 11 genes were activated by PAX5 and repressed by PAX5-ETV6 and 16 genes of the PAX5-repressed genes were activated by PAX5-ETV6 as represented by colored bars (E). Comparison of PAX5-activated and PAX5-FOXP1-repressed genes (F, *left panel*) and of PAX5-repressed and PAX5-ETV6-activated genes (F, *right panel*) in pro-B cells (9). Overlap indicates that 19 genes were activated by PAX5 and repressed by PAX5-FOXP1 and 17 genes of the PAX5-repressed genes were activated by PAX5-FOXP1 as represented by colored bars (F).

**Fig. S9.** (A) PAX5-ELN-modified genes that were included in PAX5-ETV6 and in PAX5-FOXP1 signatures were listed according to their biological functions as signal transducers, secreting proteins and surface receptors. (B) Venn diagram indicating the overlap between PAX5-ETV6-modified genes in B-ALL (fold change >5) and in pre-leukemic pro-B (fold change >3) cells and with PAX5-ELN-modified genes in pre-leukemic pro-B cells (fold change >2) (*left panel*). PAX5-ELN-modified genes that were included in B-ALL but not in pre-leukemic PAX5-ETV6 signatures (9) were listed (*right panel*).

**Table S1.** Targeted next generation sequencing (NGS) of the indicated exons of *PTPN11*, *KRAS*, *NRAS*, *JAK3*, *JAK2* and *PAX5* genes were performed on 101 B-ALL patient samples. Pink boxes identify the mutation found and its corresponding variant allele frequency (VAF). NM indicates the absence of mutation.

**Table S2.** List of genes differentially expressed between purified wt and *PE<sup>tg</sup>* pre-leukemic pro-B cells based on three independent microarray experiments. Genes with an expression difference of > 1.5 fold and adjusted *P*-value of < 0.05 are listed in the table.

**Table S3.** List of genes differentially expressed between E17.5 fetal liver *Pax5<sup>-/-</sup>* pro-B cells transduced with either MIE-PAX5 or MIE retroviral vectors based on three independent microarray experiments. Genes with an expression difference of > 3 fold, and adjusted *P*-value of < 0.05 are listed in the table.

**Table S4.** List of genes overlapping between PAX5-activated and PAX5-ELN-repressed genes and between PAX5-repressed and PAX5-ELN-activated genes using our *ex vivo* PAX5 regulated genes (A) ; between PAX5-activated and PAX5-ELN-repressed genes and between PAX5-repressed and PAX5-ELN-activated genes using the well-established *in vivo* PAX5 gene signature (8) (B) ; between PAX5-activated and PAX5-ETV6-repressed genes and between PAX5-repressed and PAX5-ETV6-activated genes using our *ex vivo* PAX5 regulated genes (C); between PAX5-activated and PAX5-FOXP1-repressed genes and between PAX5-repressed and

PAX5-FOXP1-activated genes using our *ex vivo* PAX5 regulated genes (**D**) in pre-leukemic pro-B cells.

**Table S5.** Sequences of oligonucleotide primers used for qPCR to detect V(D)J gene rearrangement of the *IgH* locus ; for the resequencing of *Jak3*, *Kras*, *Pax5* and *Ptpn11* regions by next generation sequencing in murine and *PAX5*, *PTPN11*, *KRAS*, *NRAS*, *JAK3* and *JAK2* regions in human B-ALL samples ; for gene expression in RT-qPCR.

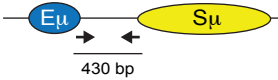
## References associated with supplementary information

1. Espeli M, Mancini SJ, Breton C, Poirier F, & Schiff C (2009) Impaired B-cell development at the pre-BII-cell stage in galectin-1-deficient mice due to inefficient pre-BII/stromal cell interactions. *Blood* 113(23):5878-5886.
2. Cobaleda C, Schebesta A, Delogu A, & Busslinger M (2007) Pax5: the guardian of B cell identity and function. *Nat Immunol* 8(5):463-470.
3. Nutt SL, Heavey B, Rolink AG, & Busslinger M (1999) Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* 401(6753):556-562.
4. Nutt SL, Urbanek P, Rolink A, & Busslinger M (1997) Essential functions of Pax5 (BSAP) in pro-B cell development: difference between fetal and adult B lymphopoiesis and reduced V-to-DJ recombination at the IgH locus. *Genes & development* 11(4):476-491.
5. Haddad D, *et al.* (2011) Sense transcription through the S region is essential for immunoglobulin class switch recombination. *The EMBO journal* 30(8):1608-1620.
6. Emig D, *et al.* (2010) AltAnalyze and DomainGraph: analyzing and visualizing exon expression data. *Nucleic Acids Res* 38(Web Server issue):W755-762.
7. Subramanian A, *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* 102(43):15545-15550.
8. Revilla IDR, *et al.* (2012) The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *The EMBO journal* 31(14):3130-3146.
9. Smeenk L, *et al.* (2017) Molecular role of the PAX5-ETV6 oncoprotein in promoting B-cell acute lymphoblastic leukemia. *The EMBO journal* 36(6):718-735.

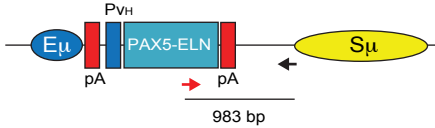
# Supplemental Figure 1

## A

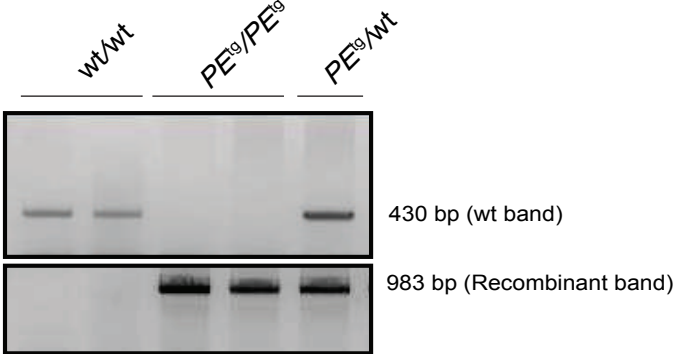
Wild type allele (wt)



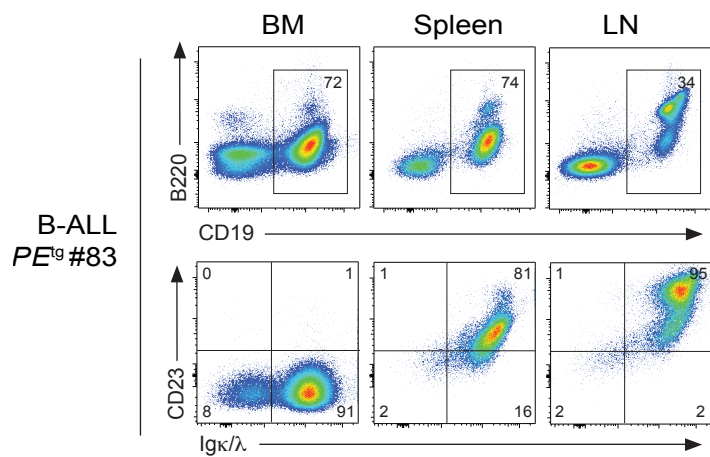
Mutant allele ( $PE^g$ )



## B



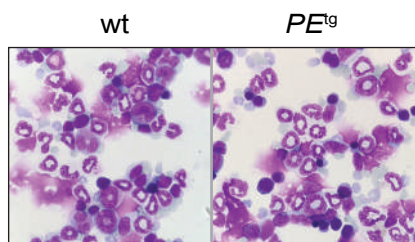
# Supplemental Figure 2



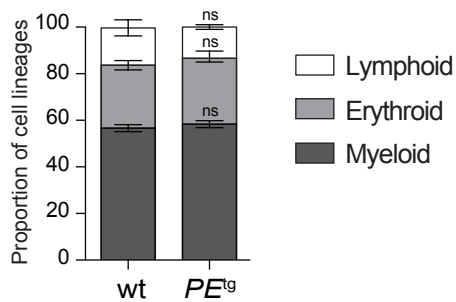


# Supplemental Figure 3

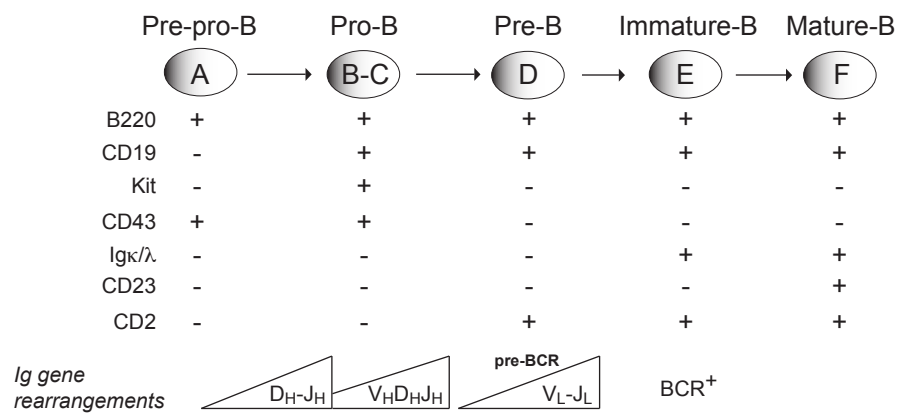
**A**



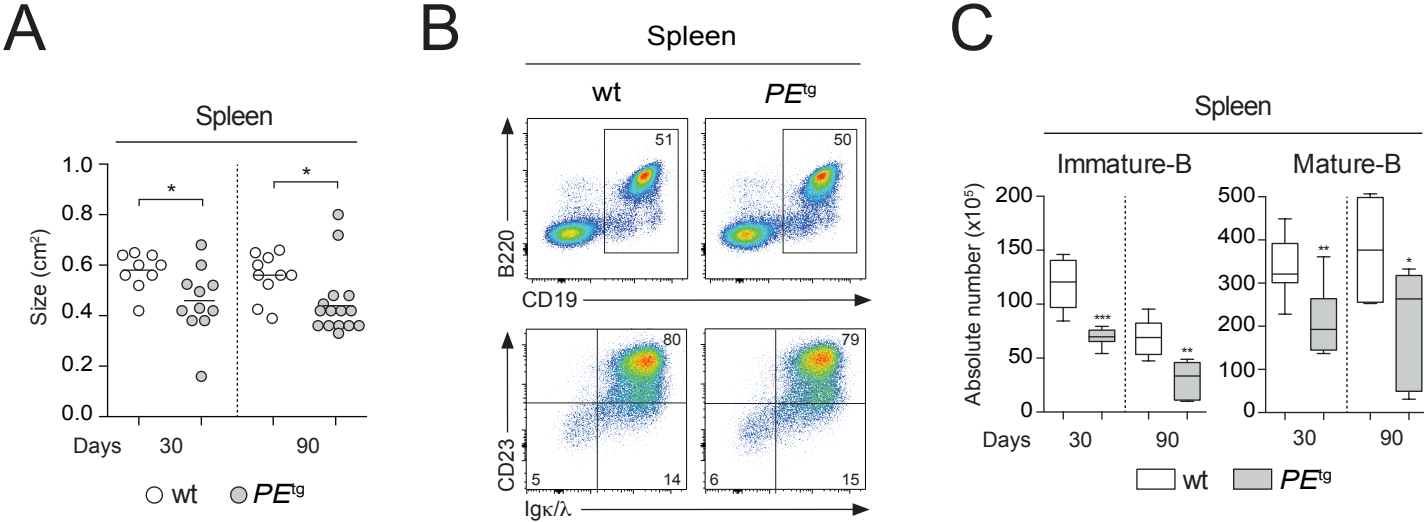
**B**



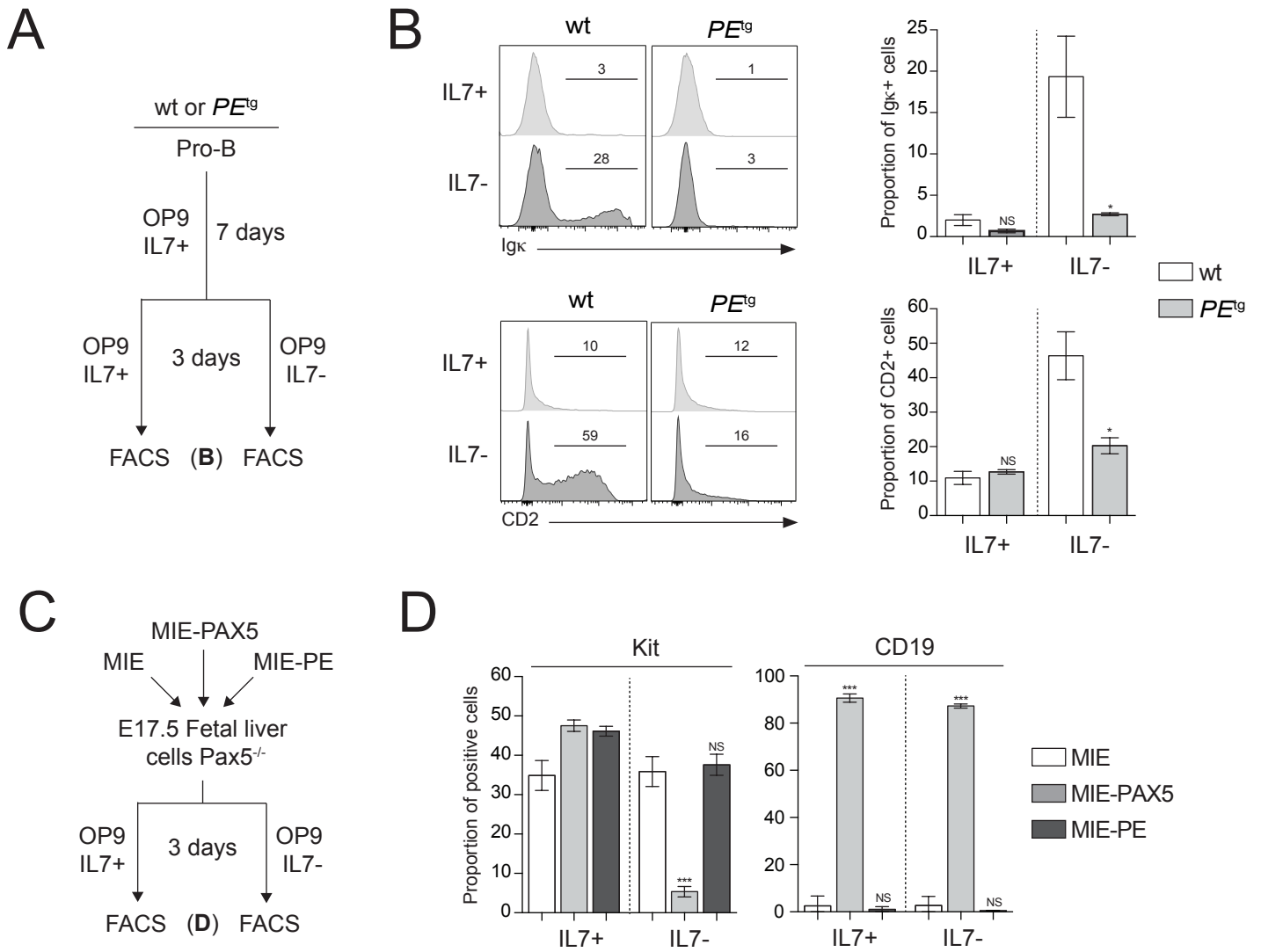
**C**



# Supplemental Figure 4

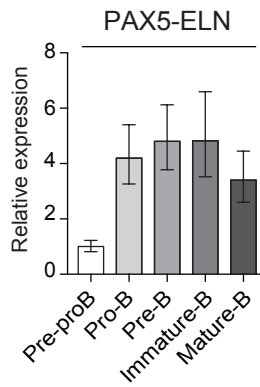


# Supplemental Figure 5

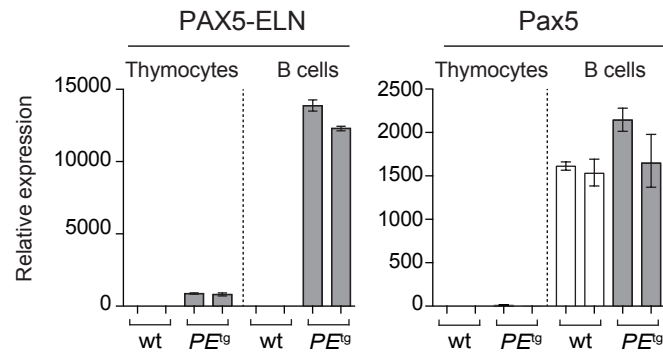


# Supplemental Figure 6

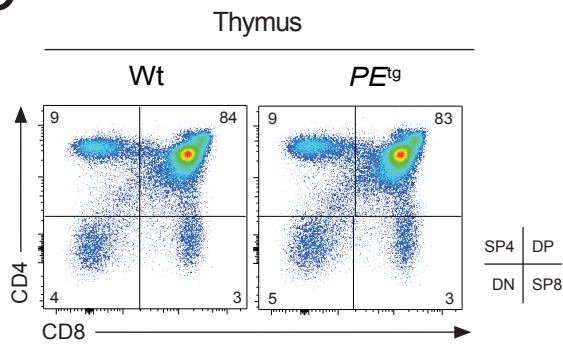
**A**



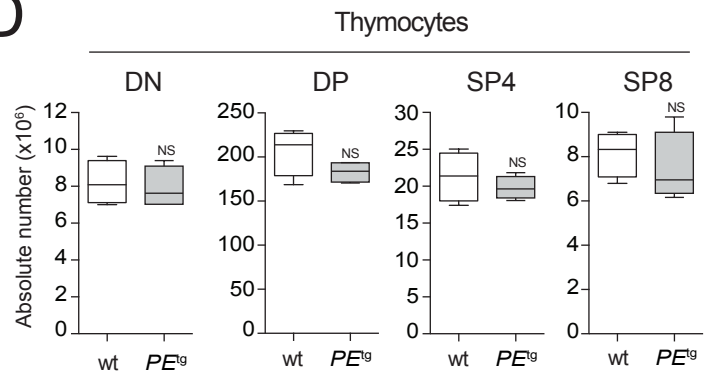
**B**



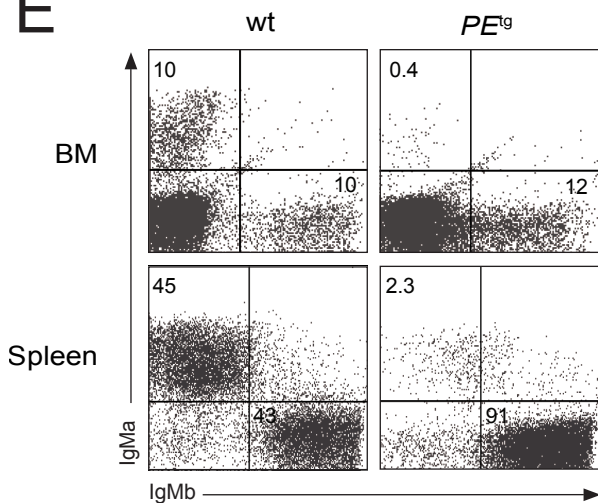
**C**



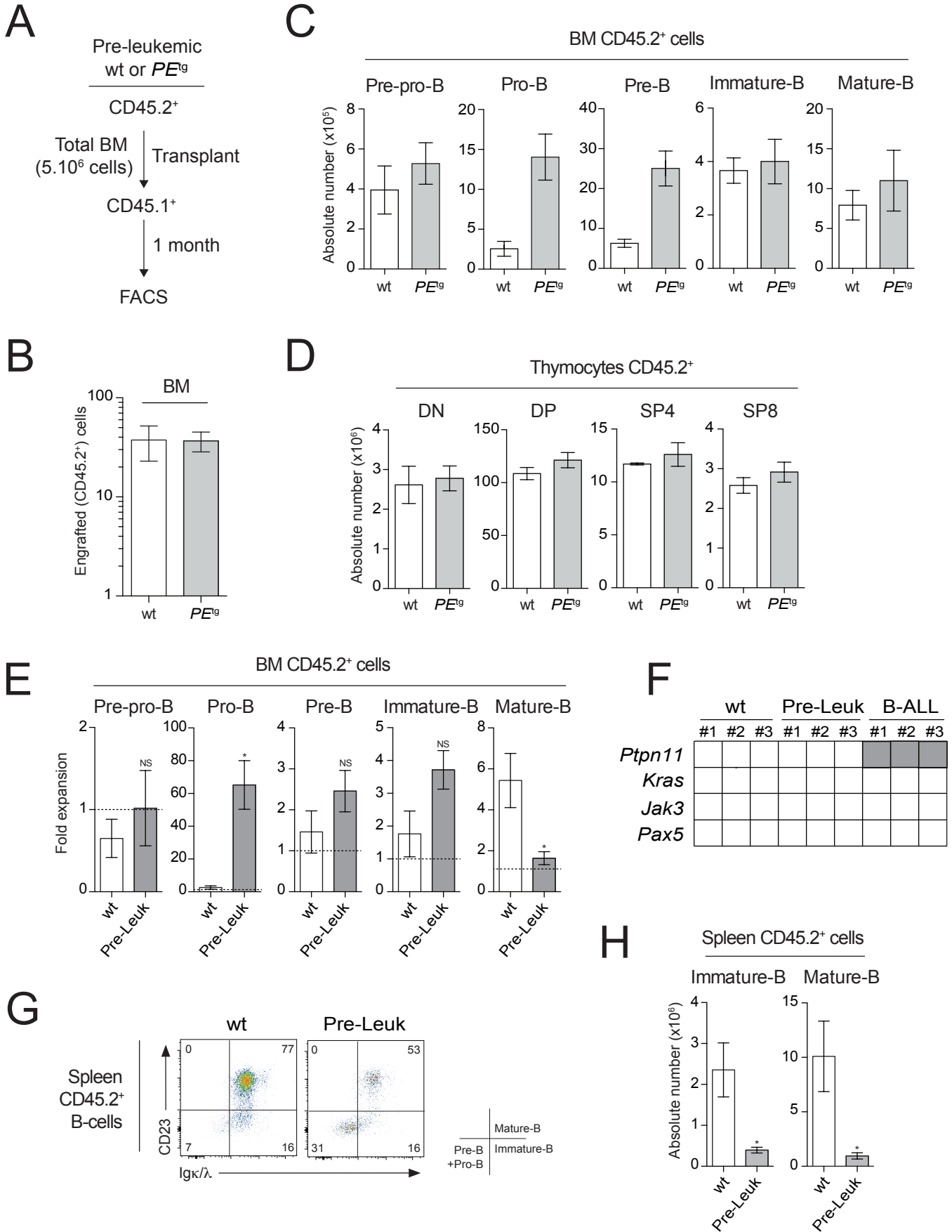
**D**



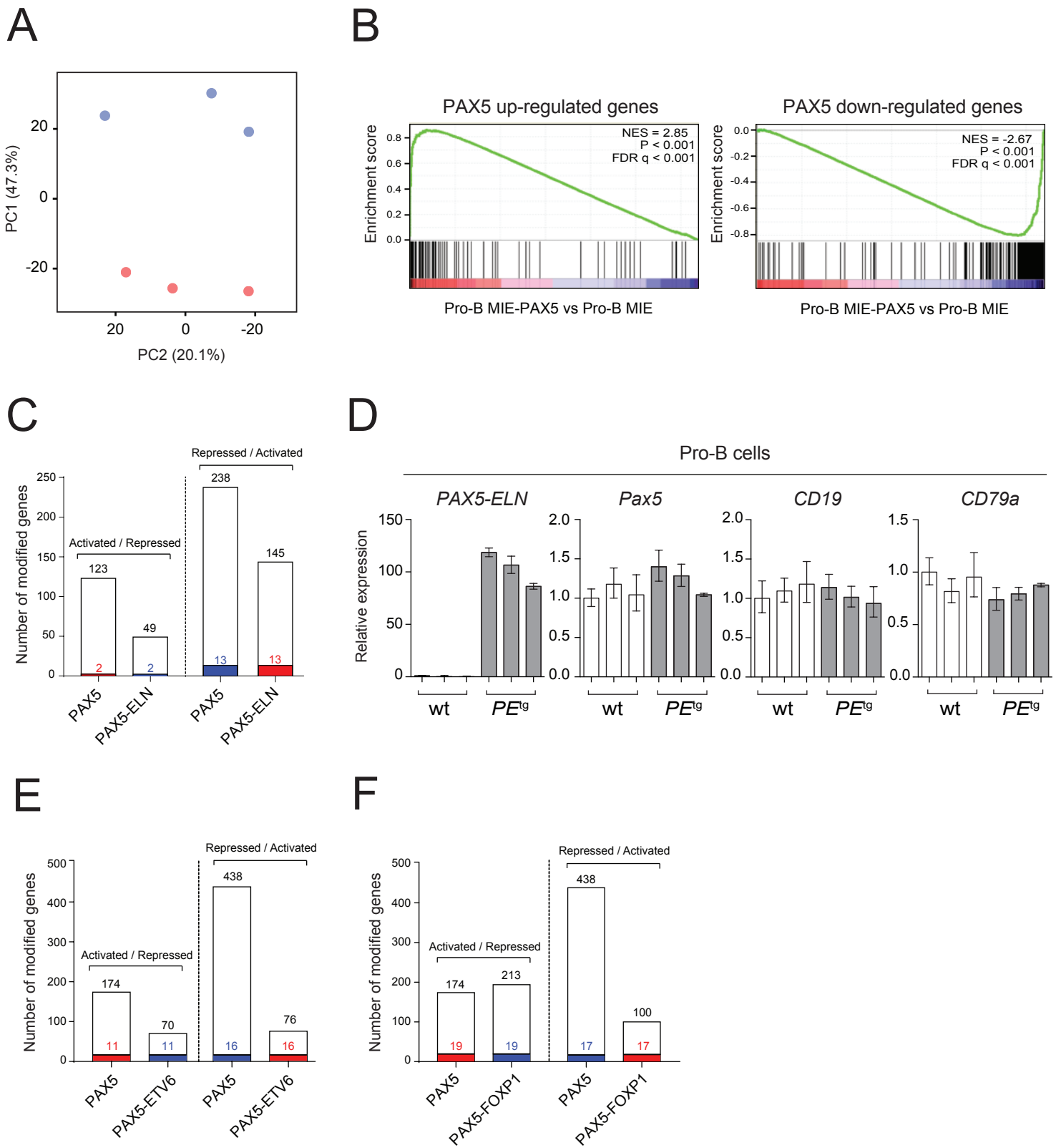
**E**



# Supplemental Figure 7



# Supplemental Figure 8

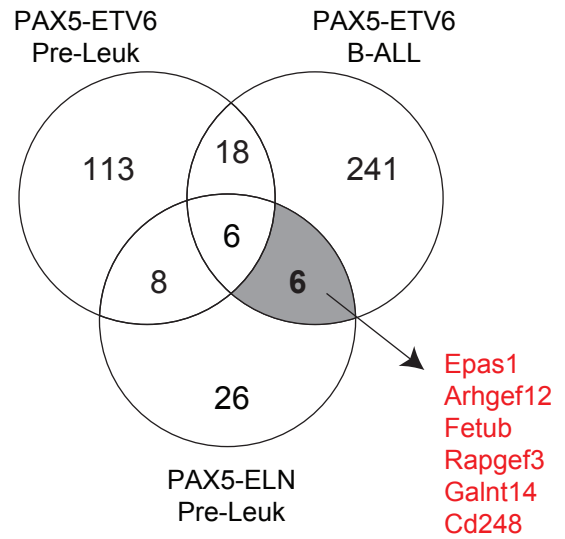


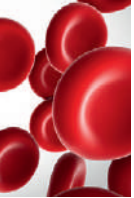
# Supplemental Figure 9

## A

Signal transducers	Secreted proteins	Surface receptors	Other functions
<i>Gimap3</i>	<i>Sema3g</i>	<i>Il2ra</i> <i>Trem1</i>	<i>Pcdh9</i> <i>Ly6d</i>

## B





### TO THE EDITOR:

# Germline *PAX5* mutation predisposes to familial B-cell precursor acute lymphoblastic leukemia

Nicolas Duployez,<sup>1,2,\*</sup> Laura A. Jamrog,<sup>3,\*</sup> Vincent Fregona,<sup>3</sup> Camille Hamelle,<sup>3</sup> Laurène Fenwarth,<sup>1,2</sup> Sophie Lejeune,<sup>4</sup> Nathalie Helevaut,<sup>1,2</sup> Sandrine Geffroy,<sup>1,2</sup> Aurélie Caillault,<sup>1,2</sup> Alice Marceau-Renaut,<sup>1,2</sup> Stéphanie Poulain,<sup>1,2</sup> Catherine Roche-Lestienne,<sup>2,5</sup> Laetitia Largeaud,<sup>3,6</sup> Naïs Prade,<sup>3,6</sup> Stéphanie Dufrechou,<sup>3,6</sup> Sylvie Hébrard,<sup>3</sup> Céline Berthon,<sup>2,7</sup> Brigitte Nelken,<sup>8</sup> José Fernandes,<sup>9</sup> Céline Villenet,<sup>10</sup> Martin Figeac,<sup>10</sup> Bastien Gerby,<sup>3</sup> Eric Delabesse,<sup>3,6</sup> Claude Preudhomme,<sup>1,2,†</sup> and Cyril Broccardo<sup>3,†</sup>

<sup>1</sup>Laboratory of Hematology, Biology and Pathology Center, Centre Hospitalier Universitaire (CHU) Lille, Lille, France; <sup>2</sup>INSERM, U1277, Lille, France; <sup>3</sup>Centre de Recherches en Cancérologie de Toulouse, INSERM UMR-1037, Université de Toulouse III Paul Sabatier, Toulouse, France; <sup>4</sup>Department of Clinical Genetics and <sup>5</sup>Institute of Medical Genetics, Jeanne de Flandre Hospital, CHU Lille, Lille, France; <sup>6</sup>Department of Hematology, CHU Toulouse, Toulouse, France; <sup>7</sup>Department of Clinical Hematology, Claude Huriez Hospital, and <sup>8</sup>Department of Pediatric Hematology, Jeanne de Flandre Hospital, CHU Lille, Lille, France; <sup>9</sup>Department of Clinical Hematology, Centre Hospitalier Valenciennes, Valenciennes, France; and <sup>10</sup>Université de Lille, Functional Genomics Platform, Lille, France

We describe here a family with a high incidence of B-cell precursor acute lymphoblastic leukemia (BCP-ALL) affecting 3 children (II.1, II.2, and II.3) at 11, 17, and 25 years old, respectively (Figure 1A,C; supplemental Figure 1, available on the *Blood* Web site). Both parents were asymptomatic. After a first relapse, the proband (II.2) underwent a familial allogeneic hematopoietic stem cell transplantation (HSCT) with his 11-year-old younger brother (II.3) leading to a complete remission (CR) for 20 years, but he relapsed a second time and died very soon after intracranial hemorrhage. His sister (II.1) developed BCP-ALL for which she was treated by chemotherapy leading to a CR even 30 years after her initial diagnosis. II.3 developed BCP-ALL 14 years after being an HSCT donor for II.2 and died of infectious complications after HSCT from an unrelated donor (supplemental Figure 1).

Because the family history was compatible with a germline transmission of the disease, whole-exome sequencing was performed on II.3, which identified a *PAX5* germline heterozygous c.113G>A mutation resulting in an R38H substitution. The R38H substitution is predicted to alter the protein affinity to DNA<sup>1</sup> and concerns a highly conserved residue located within the N-terminal DNA-binding paired domain (supplemental Figure 2A-C; supplemental Tables 1 and 2). This mutation is associated with leukemic progression in murine models with *Pax5* haploinsufficiency<sup>2</sup> or *PAX5-ELN* fusion.<sup>3</sup> This mutation was subsequently detected in all affected patients at BCP-ALL diagnosis and at remission, and in 1 asymptomatic parent without a history of malignancy (Figure 1B; supplemental Figure 1; supplemental Table 3). *PAX5* somatic alterations are present in one-third of the patients with sporadic BCP-ALL,<sup>4,5</sup> but *PAX5* germline mutations leading to BCP-ALL have been identified only recently.<sup>6-8</sup> *PAX5* encodes a critical transcription factor for B-cell differentiation<sup>9-11</sup> repressing B-lineage “inappropriate” gene expression and promoting the transcription of specific B-cell genes.<sup>10-14</sup>

As described in other germline *PAX5* pedigrees,<sup>6-8</sup> the existence of asymptomatic carriers and the absence of immunodeficiency before the onset of BCP-ALL suggest the requirement of additional genetic alterations that lead to the development of leukemia. Patients II.2 and II.3 had a normal karyotype at the BCP-ALL

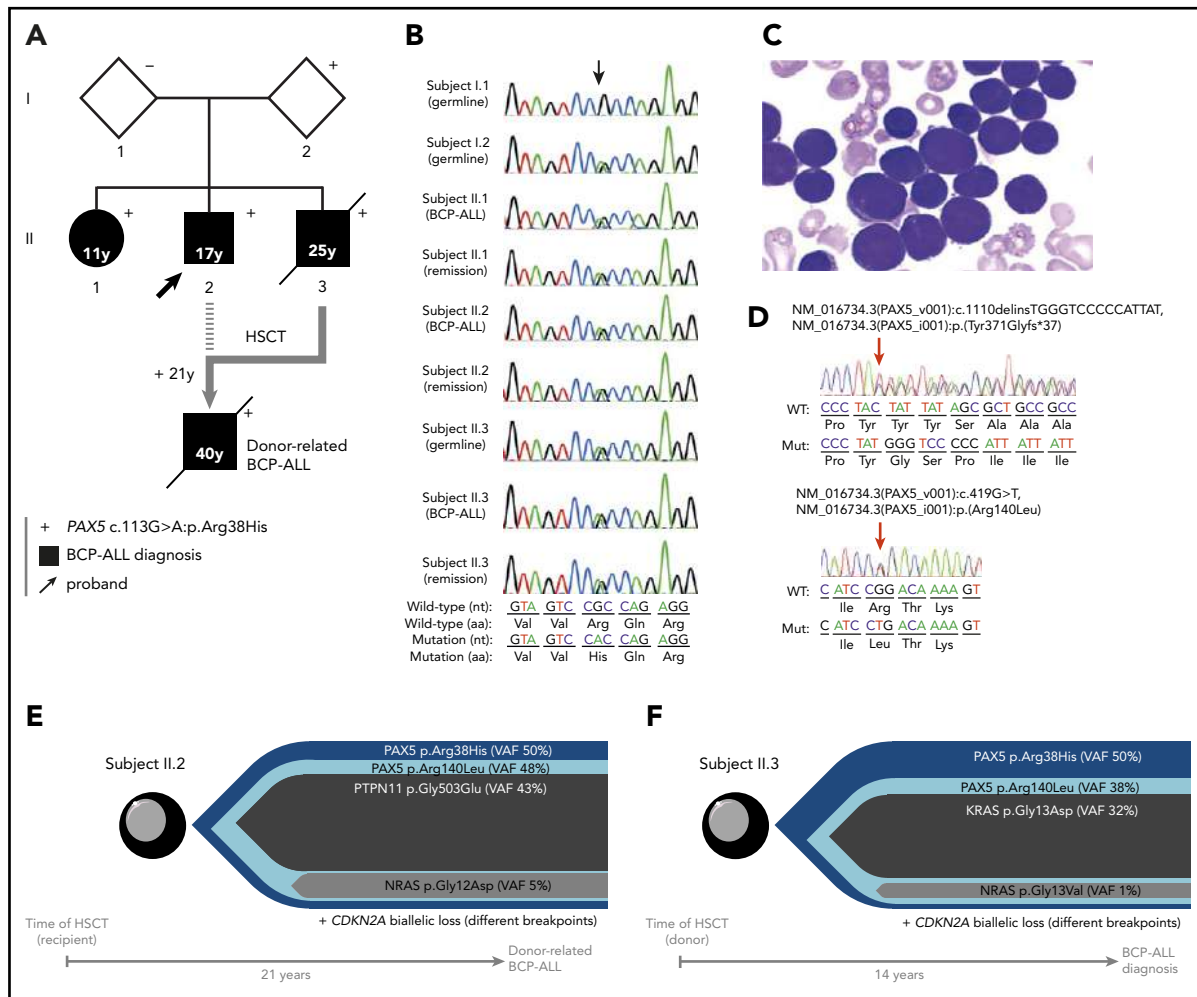
stage. Additional cytogenetic abnormalities were found in II.2 at his first (before transplantation) and second (after transplantation with his brother, consequently named II.3/2) relapses. Somatic *CDKN2A* homozygous loss and *RAS* pathway mutations, recurrent features of BCP-ALL, were detected in all analyzed BCP-ALL samples (II.2, II.3, and II.3/2) (supplemental Figure 3; supplemental Tables 3 and 4). A second *PAX5* mutation was detected at diagnosis: Y371fs in II.1 and R140L in II.3 and II.3/2 (Figure 1D-F; supplemental Table 3). Next-generation sequencing of II.3 complementary DNA (cDNA) showed that these 2 *PAX5* mutations were mainly detected on different alleles (data not shown).

More than 20 years separate HSCT allograft in II.2 from his donor-related (II.3/2) leukemia and 14 years from the allograft to diagnosis of BCP-ALL in II.3 (supplemental Figure 1). The leukemic samples from II.2 and II.3 harbored distinct *IGH* clonal rearrangements (supplemental Figure 4), suggesting that the R140L mutation occurred independently in both patients. Interestingly, the R140L mutation is recurrently associated with the R38H mutation in sporadic BCP-ALL because 10 of 11 patients reported to have an R140L mutation in the literature also had the R38H mutation, which suggests a cooperation between these 2 mutations<sup>15</sup> rather than just a biallelic inactivation of *PAX5*.<sup>16</sup>

The onset of BCP-ALL in patients with *PAX5* G183S germline mutations located in the octapeptide domain is very early with a median age of 2 years, and it is associated with a loss of chromosome 9p, leading to the simultaneous losses of the second *PAX5* allele and *CDKN2A*.<sup>6,7</sup> In contrast, the germline R38H variant (in this study and in Yazdanparast et al<sup>8</sup>) is associated with an older age of onset (11 to 25 years old) and a normal karyotype (supplemental Figure 5; supplemental Table 5).

To functionally address the impact of the R38H mutation on B-cell differentiation, we transduced *PAX5* wild-type (WT) and/or R38H murine fetal liver *Pax5*<sup>-/-</sup> B cells (Figure 2A) and the murine plasmacytoma 558L $\mu$ M cell line that does not express *Pax5* and *Cd79a* (supplemental Figure 6). As expected, *PAX5* WT restored the expression of *Cd19* in *Pax5*<sup>-/-</sup> cells and their

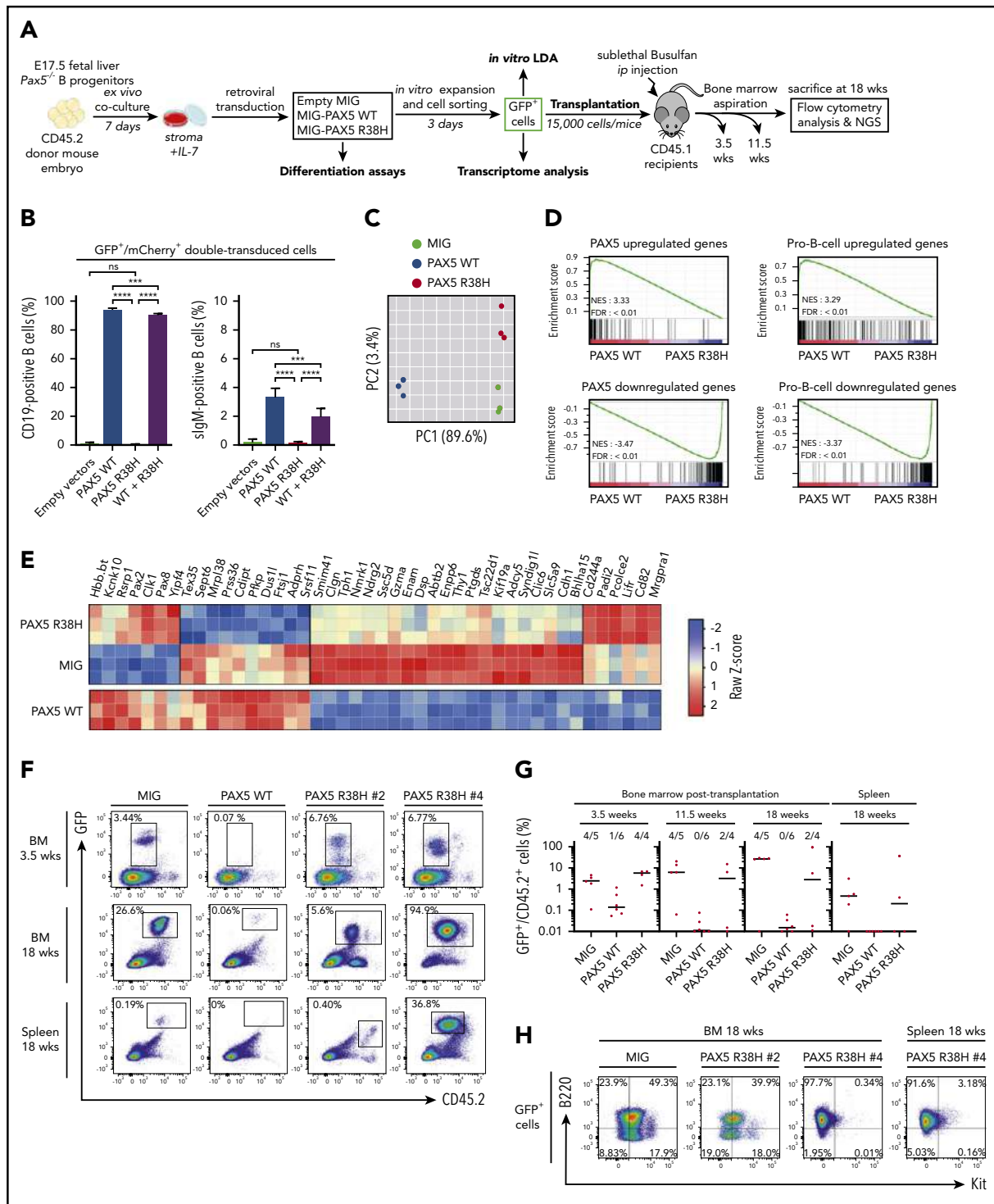




**Figure 1. Familial BCP-ALL with heterozygous germline PAX5 R38H mutation.** (A) The germline R38H variant was shown to be inherited from 1 parent (I.2) who is so far an asymptomatic carrier with no history of cancer at age 68 years. The proband (II.2, indicated by an arrow) developed BCP-ALL at age 17 years. He relapsed 2 years after his initial diagnosis and was allografted with his brother (II.3) as a donor 14 years before II.3 also developed BCP-ALL. II.2 relapsed at age 40 years and died soon thereafter as a result of an intracranial hemorrhage. His sister (II.1) developed BCP-ALL at age 11 years for which she received chemotherapy without HSCT. She is still in CR more than 30 years after her initial diagnosis. The proband's younger brother (II.3) also developed also BCP-ALL at age 25 years and died of infectious complications after HSCT from an unrelated donor. (+) Indicates the presence of the R38H germline mutation. (B) Sanger sequencing of PAX5 mutation in samples of I.1, I.2, II.1, II.2, and II.3; the location of the mutation is indicated by an arrow demonstrating a germline origin. The nucleotide and protein sequences are indicated at the bottom of the panel. (C) Representative image of May-Grünwald-Giemsa-stained marrow smear at BCP-ALL diagnosis in individual II.1. (D) Somatic PAX5 mutations in leukemic samples from individuals II.1 (top) and II.3 (bottom). Positions of mutations are indicated by red arrows. (E-F) Leukemic architecture at BCP-ALL diagnosis in individual II.3 (donor) (E) and at donor-related BCP-ALL in individual II.2 (recipient) (F). Variant allele frequency (VAF) for each mutation is indicated.

capacity to differentiate in contrast to R38H (Figure 2B). R38H was also unable to rescue surface immunoglobulin M expression in the 558L $\mu$ M cell line (supplemental Figure 6). In addition, R38H exhibited a weak competitive or additive effect on PAX5 WT when co-expressed in primary cells or in the 558L $\mu$ M cell line, respectively (Figure 2B; supplemental Figure 6). PAX5 WT significantly increased the in vitro clonal activity of transduced Pax5<sup>-/-</sup> B220<sup>+</sup> cells and led to the expression of Cd19 unlike R38H (supplemental Figure 7), confirming that R38H is not able to trigger B-cell differentiation. Furthermore, although no difference in frequencies of positive wells between PAX5 R38H- and MIG-infected cells was observed, the absolute number of B cells obtained per well was significantly higher (supplemental Figure 7). Altogether, our data demonstrate that R38H disables B-cell differentiation and partially maintains PAX5-dependent cell growth properties without overt dominant-negative effect on the normal PAX5 function.

The transcriptome of R38H-transduced Pax5<sup>-/-</sup> cells (Figure 2C-E) is similar, although not identical, to the negative control (MIG) and clearly distinct from the PAX5 WT condition (Figure 2C). Gene set enrichment analyses confirmed that R38H is unable to regulate PAX5 target genes and to lead to a pro-B-cell gene expression program in contrast to PAX5 WT<sup>17</sup> (Figure 2D; supplemental Figure 8). It is worth noting that the comparison of the residual transcriptional activity of PAX5 germline mutants showed that R38H has a stronger hypomorphic effect than G183S (supplemental Figure 9). Although the R38H transcriptome shares similarities with the one of PAX5 WT, supervised comparative analyses identified a molecular profile specific to R38H (Figure 2E). In particular, the expression of the leukemia inhibiting factor receptor (*Lif*) gene<sup>18</sup> implicated in stemness is downregulated in the PAX5 WT condition and upregulated in the R38H condition (Figure 2E). Our results suggest that PAX5 R38H confers some stemness features to the B cells. We thus



**Figure 2. PAX5 R38H behaves as a hypomorphic variant and can predispose to BCP-ALL in mice.** (A) Experimental scheme of retroviral complementation assays. (B) *Pax5*<sup>-/-</sup> cells were co-transduced with either empty MIG (MSCV-IRES-GFP) and MlmCherry (labeled as empty vectors), empty MlmCherry and MIG-PAX5 WT (labeled as PAX5 WT), empty MlmCherry and MIG-PAX5 R38H (labeled as PAX5 R38H), or MlmCherry-PAX5 WT and MIG-PAX5 R38H (labeled as WT + R38H) retroviral vectors. The proportions of cell markers were evaluated by fluorescence-activated cell sorting (FACS) for each condition (n = 8 per condition). Data are representative of 2 independent infections per experiment. Percentages indicate proportions of GFP/mCherry-double-positive that are CD19<sup>+</sup> or surface IgM (sIgM)-positive as indicated. (C) Principal component analysis (PCA) of the 2 best components for the 2000 most differentially expressed genes among the 3 conditions. Green dots represent empty MIG condition, blue dots MIG-PAX5 WT condition, and red dots MIG-PAX5 R38H condition. (D) Gene set enrichment analysis of PAX5 target gene and pro-B cell gene sets<sup>17</sup> in expression profiles of *Pax5*<sup>-/-</sup> cells transduced with PAX5 WT vs PAX5 R38H with false discovery rate (FDR) <0.01. (E) Comparative supervised heatmap using a Z-score and Spearman correlation clustering displaying the 44 most differentially expressed coding genes between PAX5 R38H-expressing *Pax5*<sup>-/-</sup> pro-B cells and MIG condition (fold >1.5; q < 0.05), with side comparison (lower panel) of corresponding gene expression in PAX5 WT-expressing *Pax5*<sup>-/-</sup> pro-B cells. (F) Representative FACS plots showing engraftment of the GFP<sup>+</sup> donor cells in bone marrow (BM) samples at 3.5 and 18 weeks and spleen samples at 18 weeks posttransplantation. Mice were numbered by #x. (G) Quantification of engraftment of CD45.2<sup>+</sup>/GFP<sup>+</sup> donor cells over time in BM at 3.5, 11.5, and 18 weeks posttransplantation and in spleen 18 weeks posttransplantation. Each dot represents individual mice (n = 4 to 6). Data show medians of engrafted mice (CD45.2<sup>+</sup>/GFP<sup>+</sup> cell proportion >1% of total cells). (H) Representative FACS plots showing B220 (CD45R) and Kit (CD117) expression in CD45.2<sup>+</sup>/GFP<sup>+</sup> cells of BM or spleen of PAX5 R38H- or MIG- transduced cells 18 weeks after transplantation. Mouse #4 shows BCP-ALL phenotype. ip, intraperitoneal; NES, normalized enrichment score; ns, not significant. Results are expressed as mean ± standard deviation. \*\*\*P < .001; \*\*\*\*P < .0001.

transplanted transduced CD45.2<sup>+</sup> Pax5<sup>-/-</sup> B cells (Figure 2A,F-H). As previously described,<sup>19,20</sup> MIG-transduced cells were able to home back to the bone marrow (BM) and to efficiently engraft at long term with an average of 30% CD45.2<sup>+</sup>/GFP<sup>+</sup> cells in BM 18 weeks after transplantation (Figure 2F-G). In contrast, PAX5 WT-transduced cells abrogated their long-term engraftment capacity in the BM and spleen of recipient mice (Figure 2G). Interestingly, the ectopic expression of R38H led to an intermediate phenotype 3.5 weeks after transplantation and, as expected, none of the engrafted GFP<sup>+</sup> cells acquired Cd19 (data not shown) because they do not express PAX5 WT. Interestingly, 2 of 4 mice maintained their engraftment 18 weeks posttransplantation and 1 (PAX5 R38H #4) developed a clonal B220<sup>+</sup>/IL7R<sup>+</sup>/Kit<sup>-</sup> leukemia (Figure 2H; supplemental Figure 10) with acquisition of a *Jak3*<sup>V670G</sup> mutation (supplemental Table 6).

In conclusion, we report a germline PAX5 p.R38H mutation in a family in which 3 children developed BCP-ALL. We demonstrated that R38H acts as a hypomorphic variant, does not abrogate the engraftment capacity of transduced Pax5<sup>-/-</sup> pro-B cells, and can predispose to BCP-ALL. The existence of asymptomatic patients without evidence of immunodeficiency before the onset of BCP-ALL shows that PAX5 germline cases might be underestimated and should be considered before any familial allograft.

## Acknowledgments

The authors thank all the patients and their families for consenting to participate in this study; Christophe Roumier (Tumor Bank, certification NF 96900-2014/65453-1, Centre Hospitalier Universitaire de Lille) for handling, conditioning, and storing patient samples; Meinrad Busslinger, MD, for the Pax5<sup>+/-</sup> mice and Michael Reth, for the 558L $\mu$ M cells; Manon Farcé from the cytometry and cell sorting facility of the Technology Cluster of the Cancer Research Center of Toulouse (INSERM U1037) for technical assistance; and the Anexplo/Genotoul platforms for technical assistance (UMS006).

This work was supported by grants from the Ligue Nationale Contre le Cancer (R19015BB), the Association "les 111 des Arts" (R18062BB), the association "Capucine" and the Société Française des Cancérogènes de l'Enfant (SFCE, R17094BB), the Région Occitanie (R16038BB), the Association "Cassandra" (R18041BB), and the Association "Constance La Petite Guerrière Astronaute" (R19043BB). L.A.J. was supported by grants from association "Capucine" and the SFCE (R17094BB) and Région Occitanie (RPH17006BBA).

## Authorship

Contribution: N.D., L.A.J., B.G., E.D., C.P., and C. Broccardo conceived and designed the experiments; N.D., L.F., N.H., S.G., A.C., A.M.-R., C.R.-L., C.V., and M.F. performed molecular and cytogenetic analyses; L.A.J., C.H., S.H., and B.G. performed and analyzed in vivo and in vitro studies; N.P. and S.D. performed the microarray experiments; L.L. analyzed the mouse next-generation sequencing data; V.F. analyzed the transcriptomic data; S.L. provided genetic counseling; S.P., C. Berthon, B.N., and J.F. provided samples and clinical data; and N.D., L.A.J., B.G., E.D., C.P., and C. Broccardo wrote the manuscript with feedback from all authors.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: N.D., 0000-0002-3927-1022; L.A.J., 0000-0003-2288-0806; L.F., 0000-0003-2808-4843; L.L., 0000-0001-5341-5427; C. Berthon, 0000-0002-3474-2577; M.F., 0000-0001-8523-3708; C.P., 0000-0002-1267-9546; C. Broccardo, 0000-0003-3016-6549.

Correspondence: Nicolas Duployez, Laboratory of Hematology, Biology and Pathology Center, Lille University Hospital, Boulevard Professeur

Leclercq, 59037 Lille, France; e-mail: nicolas.duployez@chru-lille.fr; and Cyril Broccardo, Centre de Recherche en Cancérologie de Toulouse, INSERM U1037-CRCT, 2 Avenue Hubert Curien, 31037 Toulouse Cedex 01, France; e-mail: cyril.broccardo@inserm.fr.

## Footnotes

Submitted 6 April 2020; accepted 23 September 2020; prepublished online on *Blood* First Edition 9 October 2020.

\*N.D. and L.A.J. contributed equally to this work.

†C.P. and C. Broccardo jointly directed this work.

Contact Cyril Broccardo (cyril.broccardo@inserm.fr) for original data.

The online version of this article contains a data supplement.

## REFERENCES

1. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics*. 2010; 11(1):548.
2. Dang J, Wei L, de Ridder J, et al. PAX5 is a tumor suppressor in mouse mutagenesis models of acute lymphoblastic leukemia. *Blood*. 2015; 125(23):3609-3617.
3. Jamrog L, Chemin G, Fregona V, et al. PAX5-ELN oncoprotein promotes multistep B-cell acute lymphoblastic leukemia in mice. *Proc Natl Acad Sci U S A*. 2018;115(41):10357-10362.
4. Coyaud E, Struski S, Prade N, et al. Wide diversity of PAX5 alterations in B-ALL: a Groupe Francophone de Cytogenétique Hematologique study. *Blood*. 2010;115(15):3089-3097.
5. Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*. 2007;446(7137): 758-764.
6. Shah S, Schrader KA, Waanders E, et al. A recurrent germline PAX5 mutation confers susceptibility to pre-B cell acute lymphoblastic leukemia. *Nat Genet*. 2013;45(10):1226-1231.
7. Auer F, Rüschemdorf F, Gombert M, et al. Inherited susceptibility to pre-B-ALL caused by germline transmission of PAX5 c.547G>A. *Leukemia*. 2014;28(5):1136-1138.
8. Yazdanparast S, Khatami SR, Galehdari H, Jaseb K. One missense mutation in exon 2 of the PAX5 gene in Iran. *Genet Mol Res*. 2015;14(4): 17768-17775.
9. Fuxa M, Busslinger M. Reporter gene insertions reveal a strictly B lymphoid-specific expression pattern of Pax5 in support of its B cell identity function. *J Immunol*. 2007;178(12):8222-8228.
10. Nutt SL, Heavey B, Rolink AG, Busslinger M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature*. 1999;401(6753):556-562.
11. Cresson C, Péron S, Jamrog L, et al. PAX5A and PAX5B isoforms are both efficient to drive B cell differentiation. *Oncotarget*. 2018;9(67): 32841-32854.
12. Cobaleda C, Schebesta A, Delogu A, Busslinger M. Pax5: the guardian of B cell identity and function. *Nat Immunol*. 2007;8(5):463-470.
13. 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): 269-281.
14. Nutt SL, Morrison AM, Dörfler P, Rolink A, Busslinger M. Identification of BSAP (Pax-5) target genes in early B-cell development by loss- and gain-of-function experiments. *EMBO J*. 1998;17(8):2319-2333.
15. Gu Z, Churchman ML, Roberts KG, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat Genet*. 2019;51(2): 296-307.

16. Bastian L, Schroeder MP, Eckert C, et al. PAX5 biallelic genomic alterations define a novel subgroup of B-cell precursor acute lymphoblastic leukemia. *Leukemia*. 2019;33(8):1895-1909.
17. Revilla-I-Domingo R, Bilic I, Vilagos B, et al. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J*. 2012;31(14):3130-3146.
18. Smith AG, Nichols J, Robertson M, Rathjen PD. Differentiation inhibiting activity (DIA/LIF) and mouse development. *Dev Biol*. 1992;151(2):339-351.

19. Rolink AG, Nutt SL, Melchers F, Busslinger M. Long-term in vivo reconstitution of T-cell development by Pax5-deficient B-cell progenitors. *Nature*. 1999;401(6753):603-606.
20. Schaniel C, Bruno L, Melchers F, Rolink AG. Multiple hematopoietic cell lineages develop in vivo from transplanted Pax5-deficient pre-B I-cell clones. *Blood*. 2002;99(2):472-478.

DOI 10.1182/blood.2020005756

© 2021 by The American Society of Hematology

## TO THE EDITOR:

# *RUNX1* germline variants in *RUNX1*-mutant AML: how frequent?

Martijn P. T. Ernst, François G. Kavelaars, Bob Löwenberg, Peter J. M. Valk,\* and Marc H. G. P. Raaijmakers\*

Department of Hematology, Erasmus Medical Center Cancer Institute, Rotterdam, The Netherlands

*RUNX1* mutations are recurrent aberrations in acute myeloid leukemia (AML) that are either somatically acquired or originate in the germline. Monoallelic pathogenic germline *RUNX1* variants cause familial platelet disorder with propensity to AML (FPD/AML), featuring predisposition to develop myeloid malignancies. Recently, Simon et al<sup>1</sup> found that 30% of *RUNX1*-mutated AML patients carried a (nonpolymorphic) *RUNX1* germline variant. This high frequency contrasted previously reported frequencies in AML cohort studies.<sup>2-5</sup>

To obtain further insight into the frequency of germline variants within *RUNX1*-mutated AML patients, we retrieved data from a large independent cohort of AML patients that was previously interrogated by paired genetic sequencing at diagnosis and in complete remission (CR).<sup>6</sup> Targeted next-generation sequencing (NGS) at diagnosis was performed in 763 AML patients, enrolled in the Dutch-Belgian Hemato-Oncology Cooperative Group (HOVON)102 clinical trial (2010-2013),<sup>7</sup> using the TruSight Myeloid Sequencing Panel (Illumina). The HOVON102 trial was designed to investigate the added value of clofarabine in combination with standard remission-induction chemotherapy in adult (18-65 years) AML or myelodysplastic syndrome patients. NGS data were analyzed as previously described,<sup>6</sup> and nonpolymorphic *RUNX1* variants were annotated to RefSeq ID NM\_001754.5.

At diagnosis, at least 1 *RUNX1* mutation was detected in 115 of 763 AML patients (15.1%) for a total of 142 *RUNX1* mutations (data not shown), corresponding to frequencies of 10% to 15% reported in literature.<sup>3,4,8,9</sup> We detected 2 or more mutations in 23 of 115 patients.

To distinguish germline variants from somatic mutations, we retrieved data from paired diagnostic and remission HOVON102 AML samples to identify persisting *RUNX1* mutations. Remission samples of 287 CR patients were available for NGS analysis.<sup>6</sup> Among these 287 AML patients, 48 *RUNX1* mutations were detected in 37 patients (12.9%), of whom 9 carried 2 or more mutations (supplemental Tables 1 and 2, available on the *Blood* Web site). The distribution of the variant allele frequencies (VAF) at diagnosis of the *RUNX1* mutations of the selected samples (n = 48 in 37/287 patients) was similar to the initial HOVON102

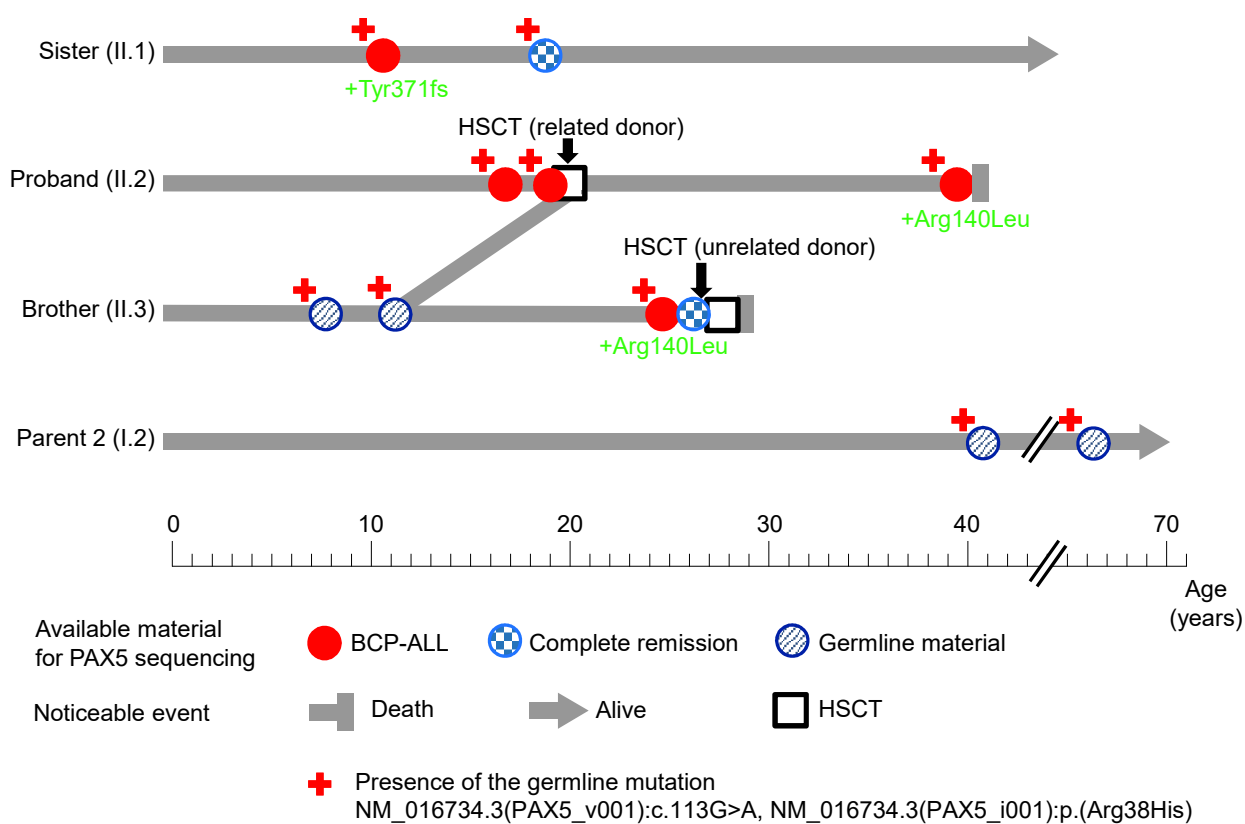
cohort (n = 142 in 115/763 patients) (Figure 1A). In the majority of cases (34/37), *RUNX1* mutations were acquired because the VAF of these mutations (n = 45) was  $\leq 10\%$  in CR (Figure 1A). However, in 3 of the 37 AML cases (8.1%), a *RUNX1* mutation was present at a VAF of 50% in both diagnostic and remission samples (Figure 1A), highly indicative for germline origin. Other cooccurring mutations present at diagnosis in these 3 AML patients were either cleared or persisted at much lower VAF in CR (Figure 1B-D), supporting the germline status of these *RUNX1* variants. Thus, in the current cohort of 287 AML patients that attained CR, 3 (1.0%) harbored a germline *RUNX1* variant.

We identified 2 different *RUNX1* germline variants in these 3 AML patients. A single patient (no. 748) carried *RUNX1* p.(Arg232Trp) (exon 7), a missense variant located between the runt homology domain and the transactivation domain (Figure 1E). The 2 other patients (no. 124 and no. 740) shared variant *RUNX1* c.97+1G>A, affecting the splice donor site of exon 3 (Figure 1E). A familial bond between these 2 patients could not be confirmed.

In accordance to recently published guidelines by the ClinGen Myeloid Malignancy Variant Curation Expert Panel,<sup>10</sup> we classified both identified *RUNX1* germline variants as variants of unknown significance (supplemental Table 3). Neither of these have been reported as *RUNX1* germline variants before in the context of FPD/AML traits,<sup>11</sup> although *RUNX1* p.(Arg232Trp) has been recurrently reported as somatically acquired mutation in AML.<sup>1,3,4</sup> Both variants occur outside of the runt homology domain, in which most missense *RUNX1* germline variants cluster.<sup>11</sup> *RUNX1* c.97+1G>A is presumed to specifically affect isoform *RUNX1C*, 1 of 3 major *RUNX1* isoforms, complicating its interpretation.<sup>10</sup> However, deletions of exon 2 and 3, also putatively affecting *RUNX1C* specifically, have been reported in the context of FPD/AML before.<sup>12</sup>

Additional somatic mutations were found in all 3 patients, most of which have been reported in the context of germline *RUNX1*-mutant AML before, including additional *RUNX1* mutations and concomitant mutations in *DNMT3A*, *FLT3*, and *GATA2*.<sup>1,11,13-15</sup> Secondary mutations in *RUNX1* are a frequent recurrent event in

## Supplemental Figure 1

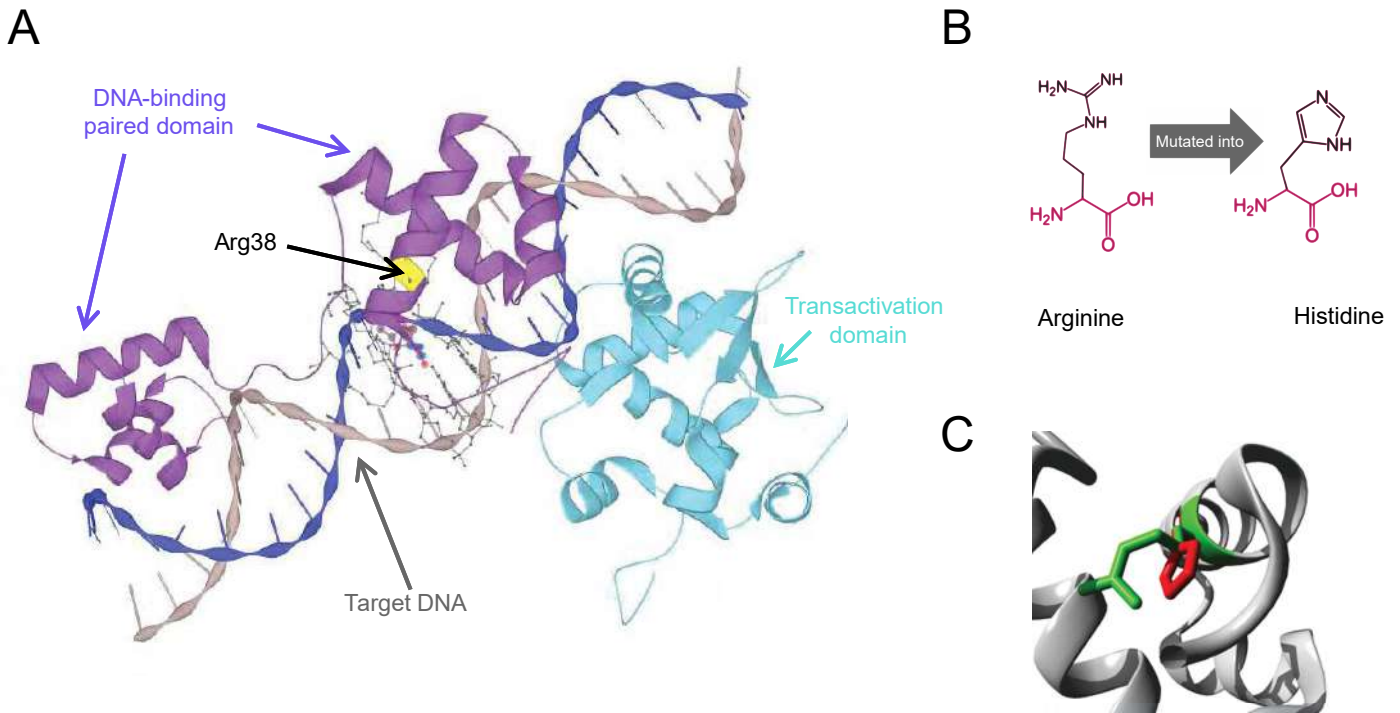


Supplemental Figure 1: Timeline of the clinical evolution of each member of the reported family with a germline PAX5 p.Arg38His mutation.

Each dot represents one sample. The ages of onset are depicted with a scale under the figure. Germline mutation presence is indicated with a + symbol. Acquired PAX5 mutations are detailed in green colour. Death of the patient is indicated with a stop line; fs: frameshift



## Supplemental Figure 2

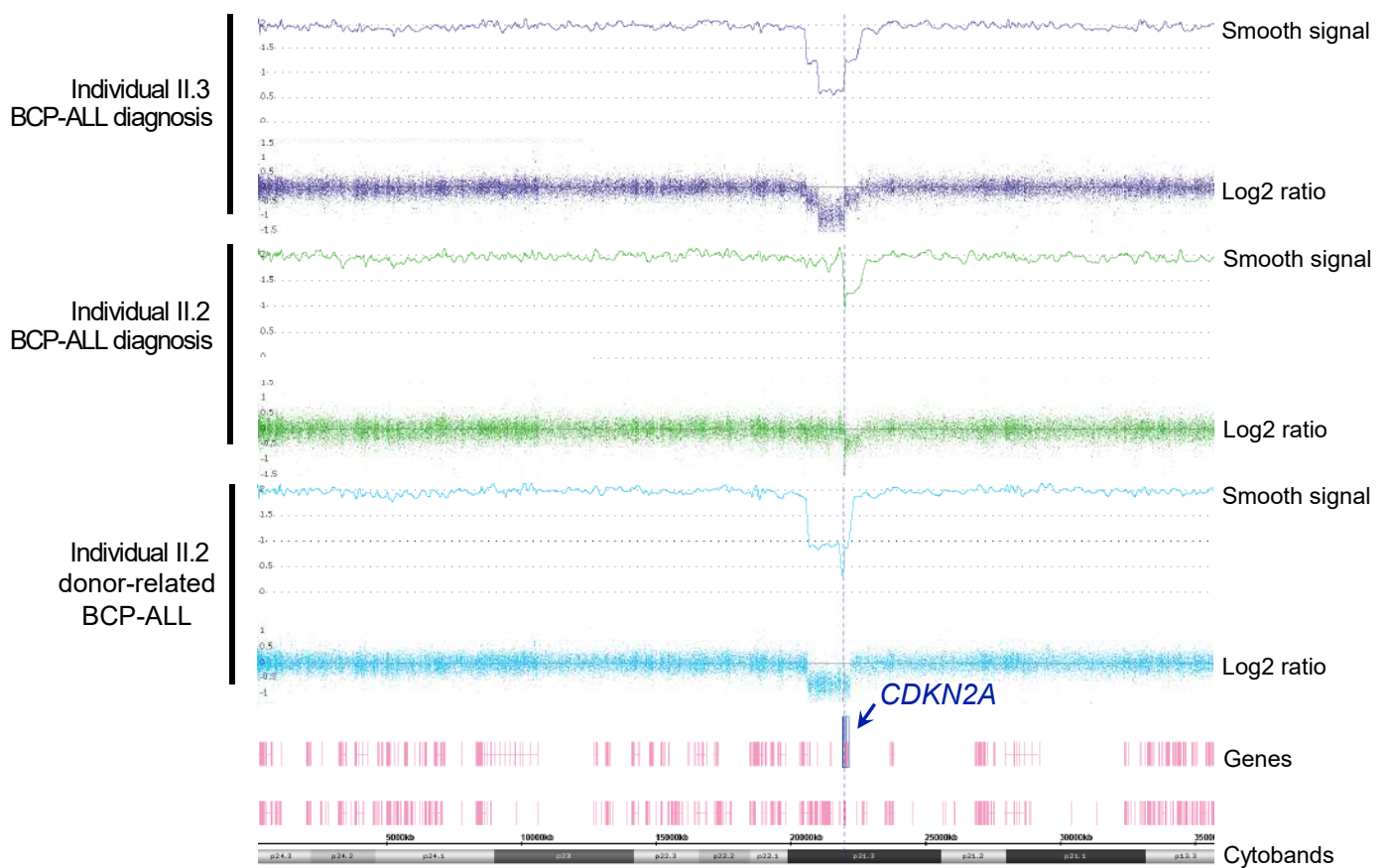


Supplemental Figure 2: Location of the p.Arg38His variant in the PAX5 protein.

(A) Tertiary structure of the DNA-binding domain of the PAX5 protein. Arginine 38 (Arg38) is indicated in yellow colour (from <https://www.uniprot.org>). (B) Schematic structure of the wild-type (Arginine) and the mutant (Histidine) amino acids (C) Close-up of the tertiary structure of the mutated site. The side chain of the wild type or the mutant amino acid is coloured in green or red respectively, the rest of the protein is coloured in grey (from HOPE web application: <http://www.cmbi.ru.nl/hope/>)<sup>1</sup>

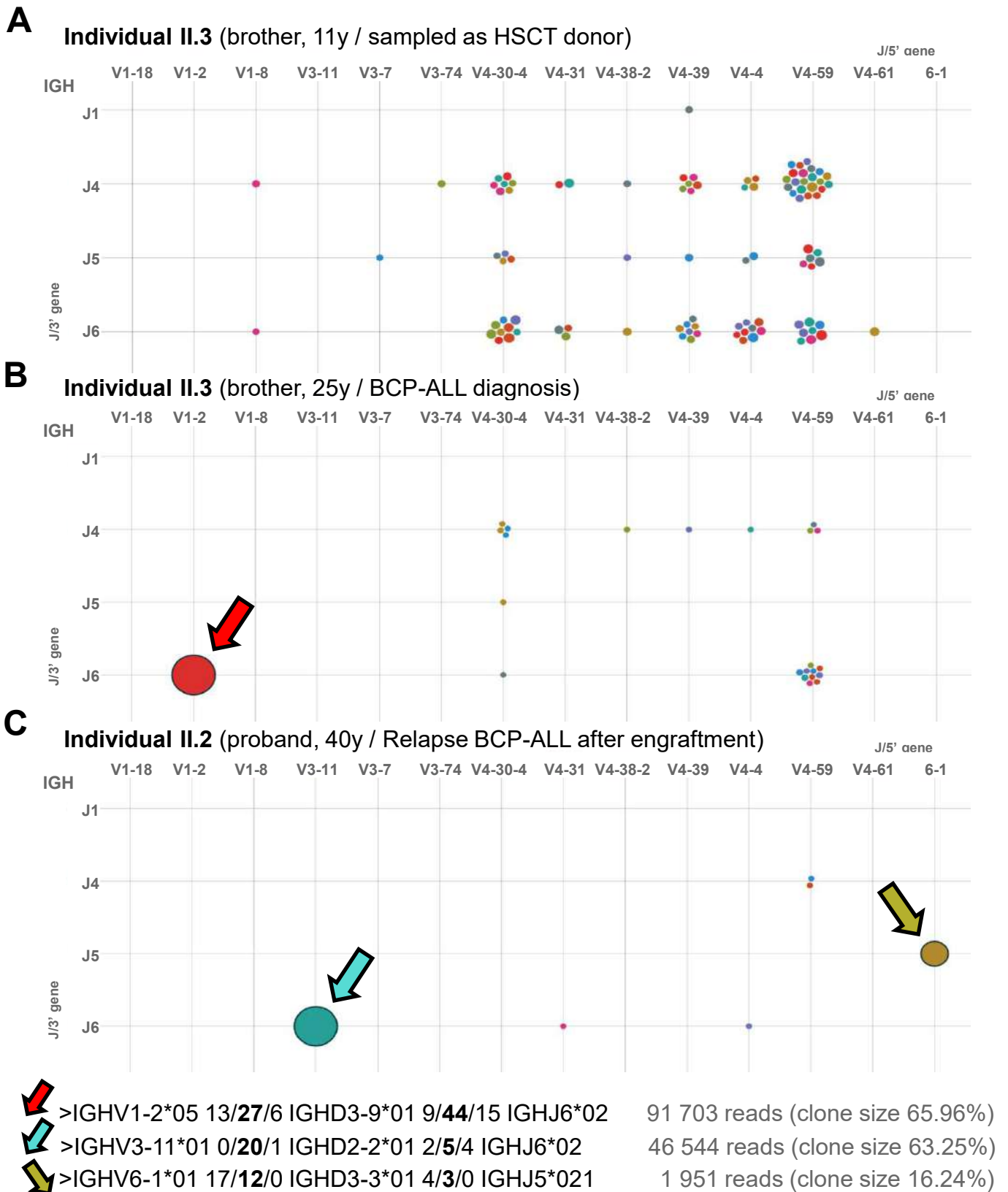
## Supplemental Figure 3

View: chr9: 0 - 36,135,435



Supplemental Figure 3: BCP-ALL samples show *CDKN2A* deletion SNP-array karyotyping showing cryptic biallelic deletions of *CDKN2A* at BCP-ALL stage in II.2 and II.3

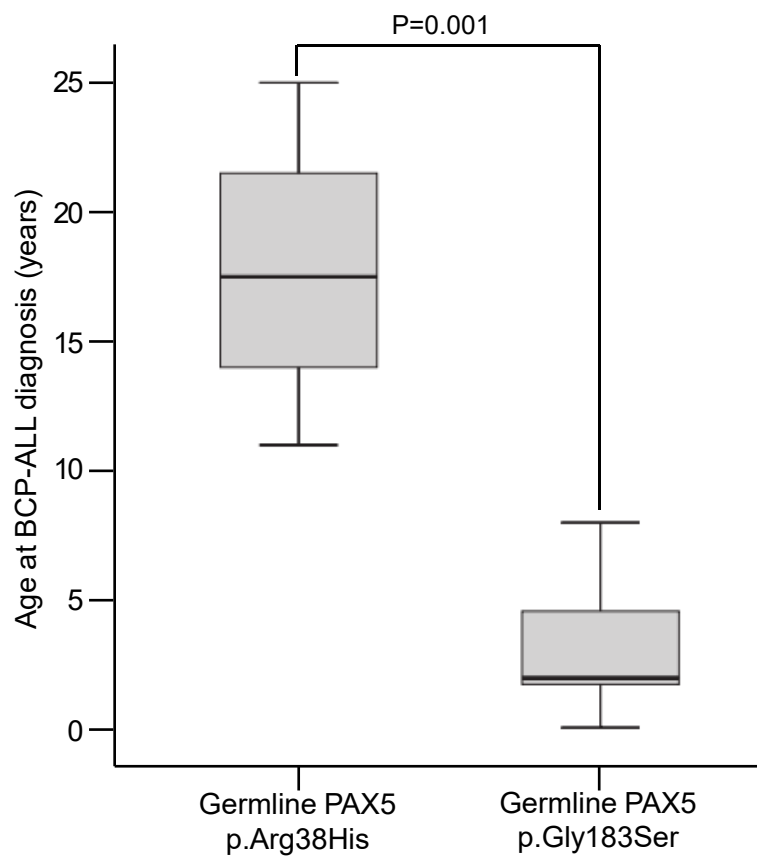
# Supplemental Figure 4



Supplemental Figure 4: *IGH* rearrangements of the two brothers. *IGH* rearrangements were visualized with the Vidjil web application. *IGH* clonotypes are displayed as circles whose size is proportional to the number of reads carrying a rearrangement (which is assumed to reflect the proportion of each rearrangement in the sample). The coordinates of the circles reflect the *IGH* genes involved in each rearrangement (V segments in the x axis, J segments in the y axis). The main clones found in leukemic samples from II.3 (panel B) and II.2 (panel C) are highlighted by arrows and their respective specific sequences are detailed below the figure. Both leukemic samples harbored leukemia-specific *IGH* rearrangements that were absent in sample from the individual II.3 (panel A) at time of stem cell donation.

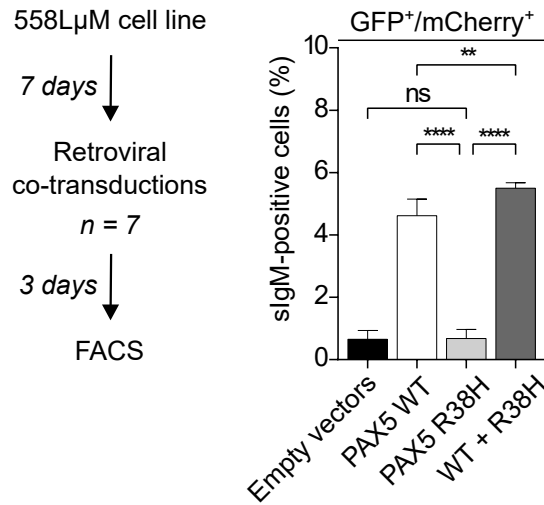


## Supplemental Figure 5



Supplemental Figure 5: Age of onset of BCP-ALL in patients with germline PAX5 mutations. Boxes show the interquartile range divided by the median value for age at BCP-ALL diagnosis in patients with germline PAX5 p.Arg38His mutation (n=4) versus PAX5 p.Gly183Ser mutation (n=13).

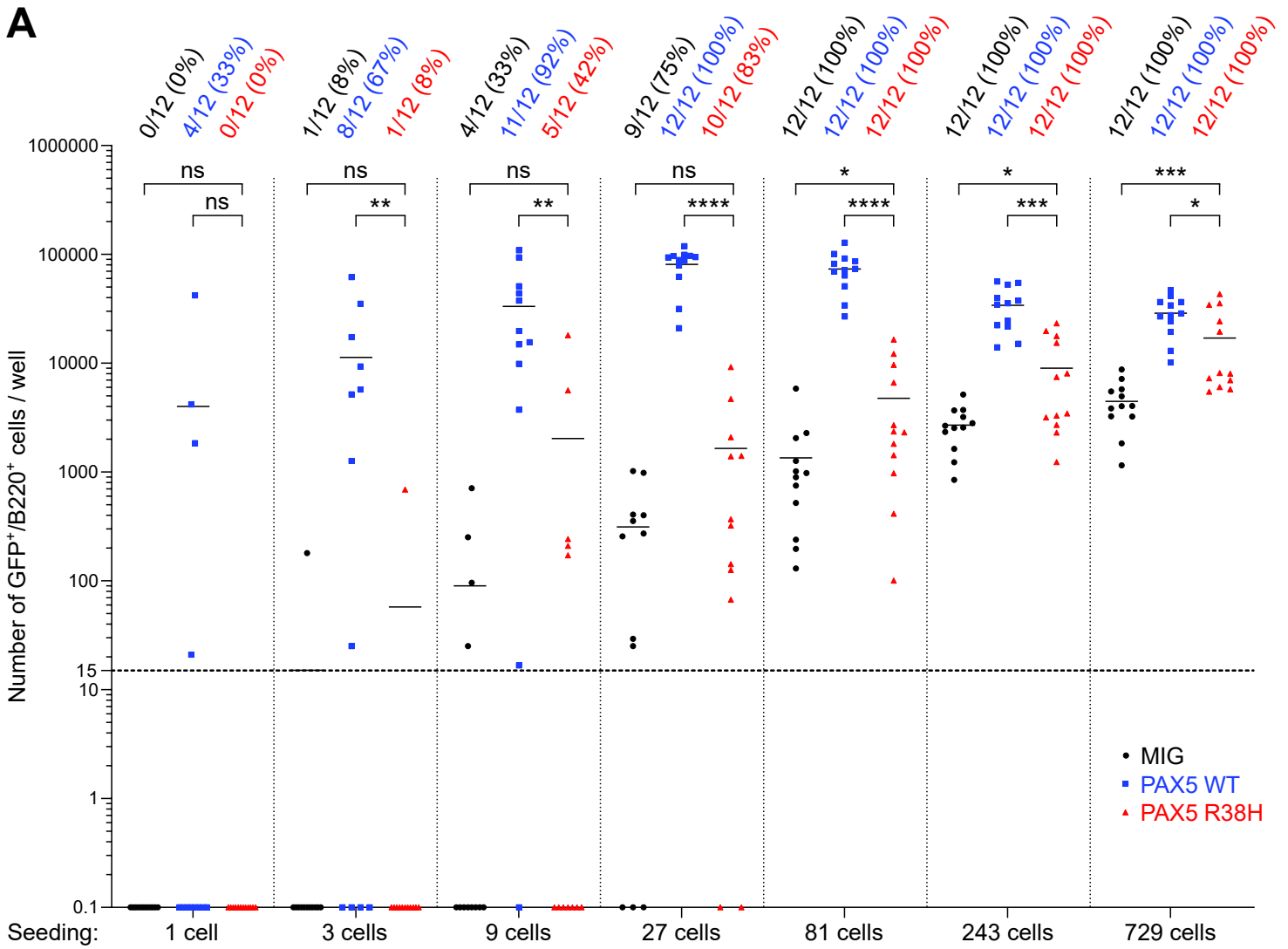
## Supplemental Figure 6



Supplemental Figure 6: PAX5 R38H does not rescue PAX5 WT activity and does not exert dominant-negative effect on PAX5 WT in 558LμM cell line.

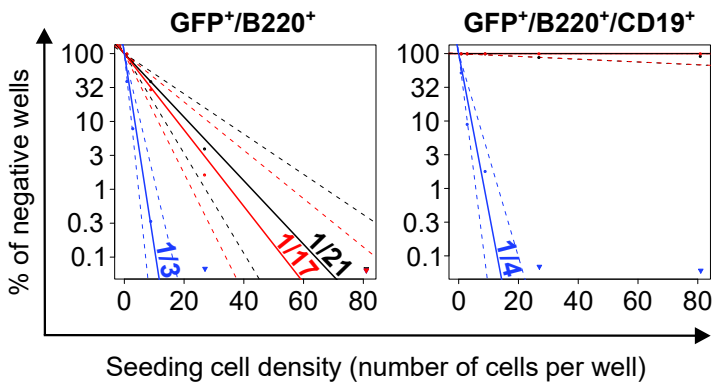
Left panel: Experimental scheme of 558LμM cell line retroviral complementation. Right panel: 558LμM cells were co-transduced with GFP (MIG)- and mCherry (MImCherry)-expressing vectors (labeled as empty vectors), with MIG and MImCherry-PAX5 WT (labeled as PAX5 WT), MImCherry and MIG-PAX5 R38H (labeled as PAX5 R38H) or MImCherry-PAX5 WT and MIG-PAX5 R38H (labeled as WT + R38H) retroviral vectors. Surface IgM (sIgM) expression was assessed by flow cytometry as a reflection of CD79 transactivation; cells were previously gated on GFP<sup>+</sup>/mCherry<sup>+</sup> criteria. Data are representative of two independent experiments (n=3 and n=4 independent infections per condition). Results are expressed as mean ± SD; \* p<0.05, \*\*p<0.01, \*\*\*p<0.001.

# Supplemental Figure 7



**B**

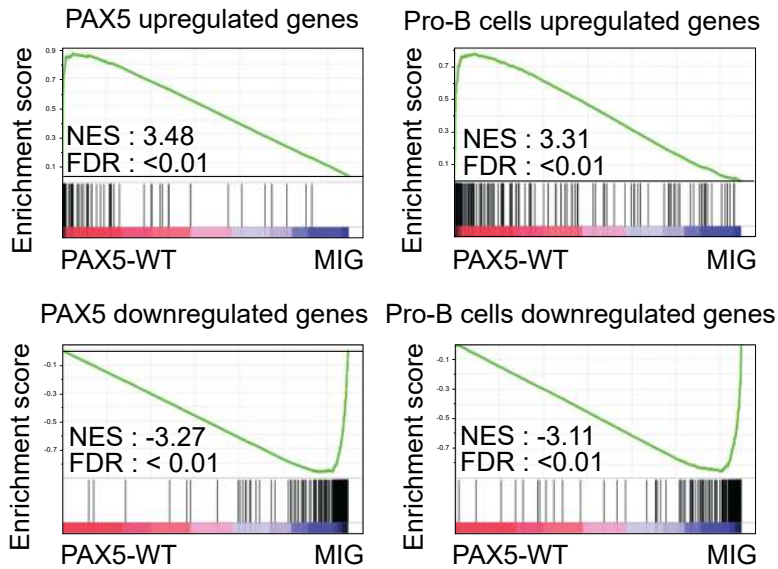
— MIG  
 — PAX5 WT  
 — PAX5 R38H



Fraction	Condition	Clonal frequency [95% CI]
GFP <sup>+</sup> /B220 <sup>+</sup> cells	MIG	1/21 [1/33 ; 1/13]
	PAX5 WT	1/3 [1/5 ; 1/2]
	PAX5 R38H	1/17 [1/28 ; 1/11]
GFP <sup>+</sup> /B220 <sup>+</sup> /CD19 <sup>+</sup> cells	MIG	
	PAX5 WT	1/4 [1/6 ; 1/2]
	PAX5 R38H	

Supplemental Figure 7: Clonogenic activity of transduced *Pax5*<sup>-/-</sup> fetal cells determined by in vitro limiting dilution assays. (A) Quantification by FACS of GFP<sup>+</sup>/B220<sup>+</sup> cells per well after one-week expansion. Bars represent means of 12 replicates per seeding density. Numbers of positive wells (more than 15 cells per well) and corresponding proportions in percentages are indicated. (B) Quantification of estimated frequency of transduced *Pax5*<sup>-/-</sup> fetal cells with ELDA (Extreme Limiting Dilution Analysis). ELDA (<http://bioinf.wehi.edu.au/software/elda/>)<sup>5</sup> was performed and the estimated frequency and range (95% Confidence Interval, 95% CI [lower value; upper value]) of transduced *Pax5*<sup>-/-</sup> cells are indicated in graph and table. The plain lines represent regression line fitting to the experimental data and the dotted lines represent the regression lines for the 95% CIs based on the likelihood ratio test of single-hit model. LDA analysis given by ELDA software also provides a Chi-square test resulting in non-significant difference between MIG and PAX5 R38H conditions.

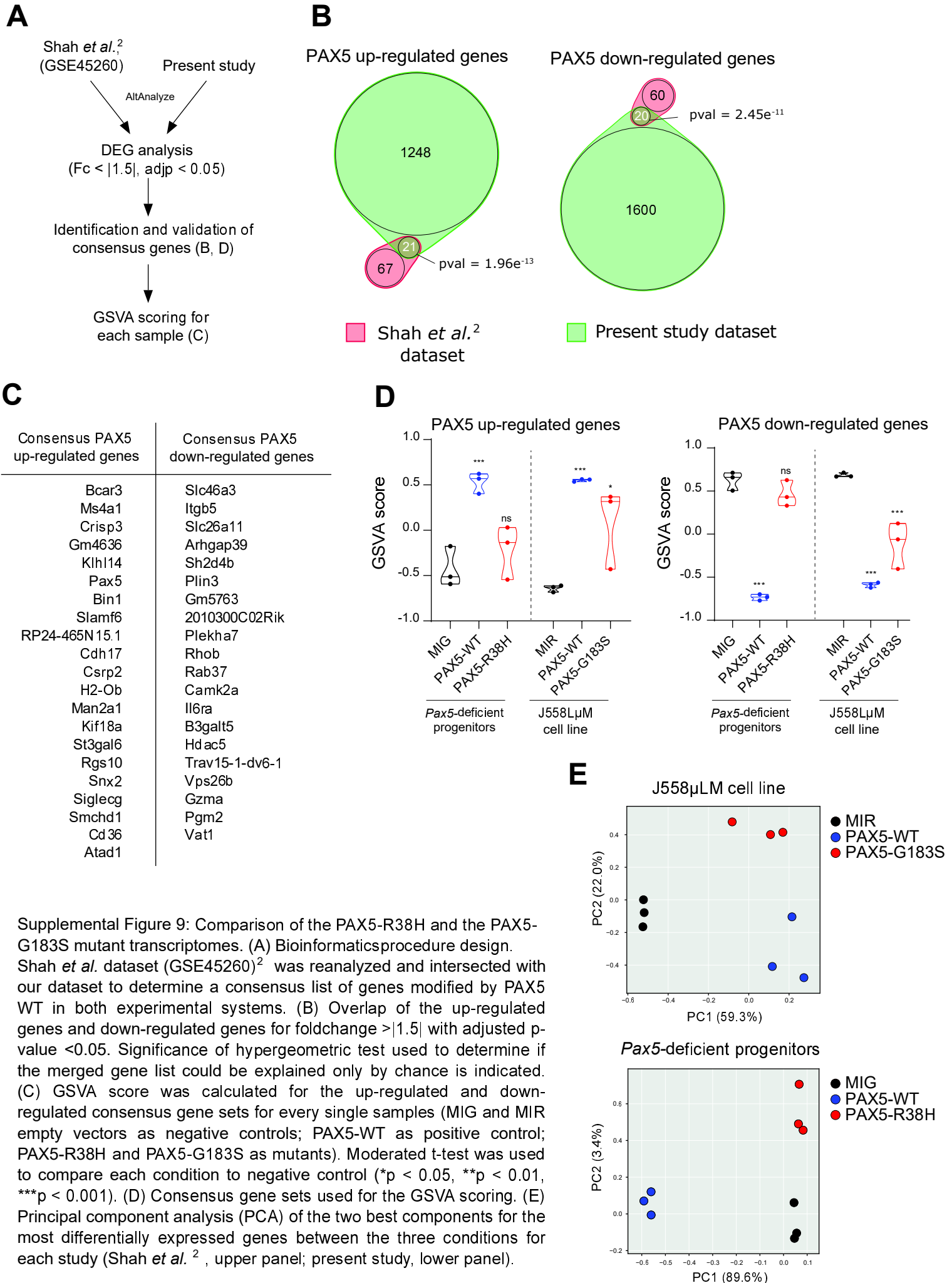
## Supplemental Figure 8



Supplemental Figure 8: GSEA of PAX5 target gene set and pro-B cell gene set in rescued *Pax5*<sup>-/-</sup> cells.

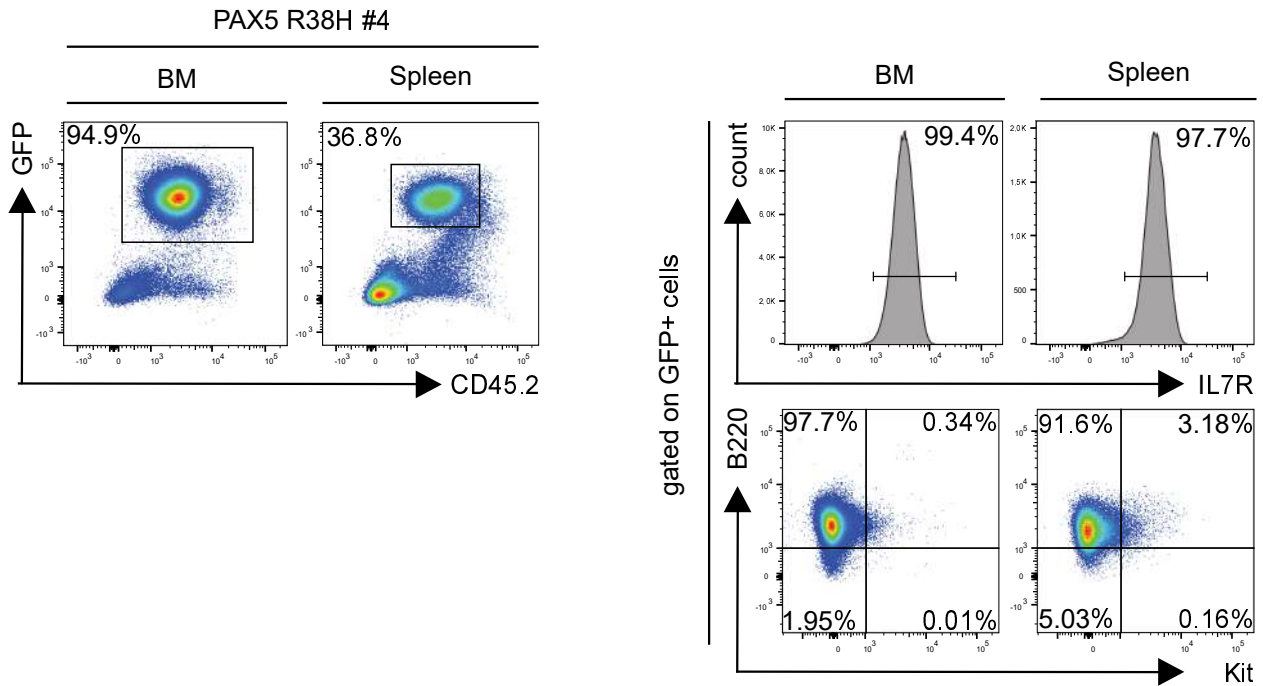
GSEA of PAX5 target gene set and pro-B cell gene set<sup>6</sup> in *Pax5*<sup>-/-</sup> cells transduced with PAX5 WT versus empty MIG. False Discovery Rate (FDR) < 0.01; NES: Normalized Enrichment Score.

# Supplemental Figure 9

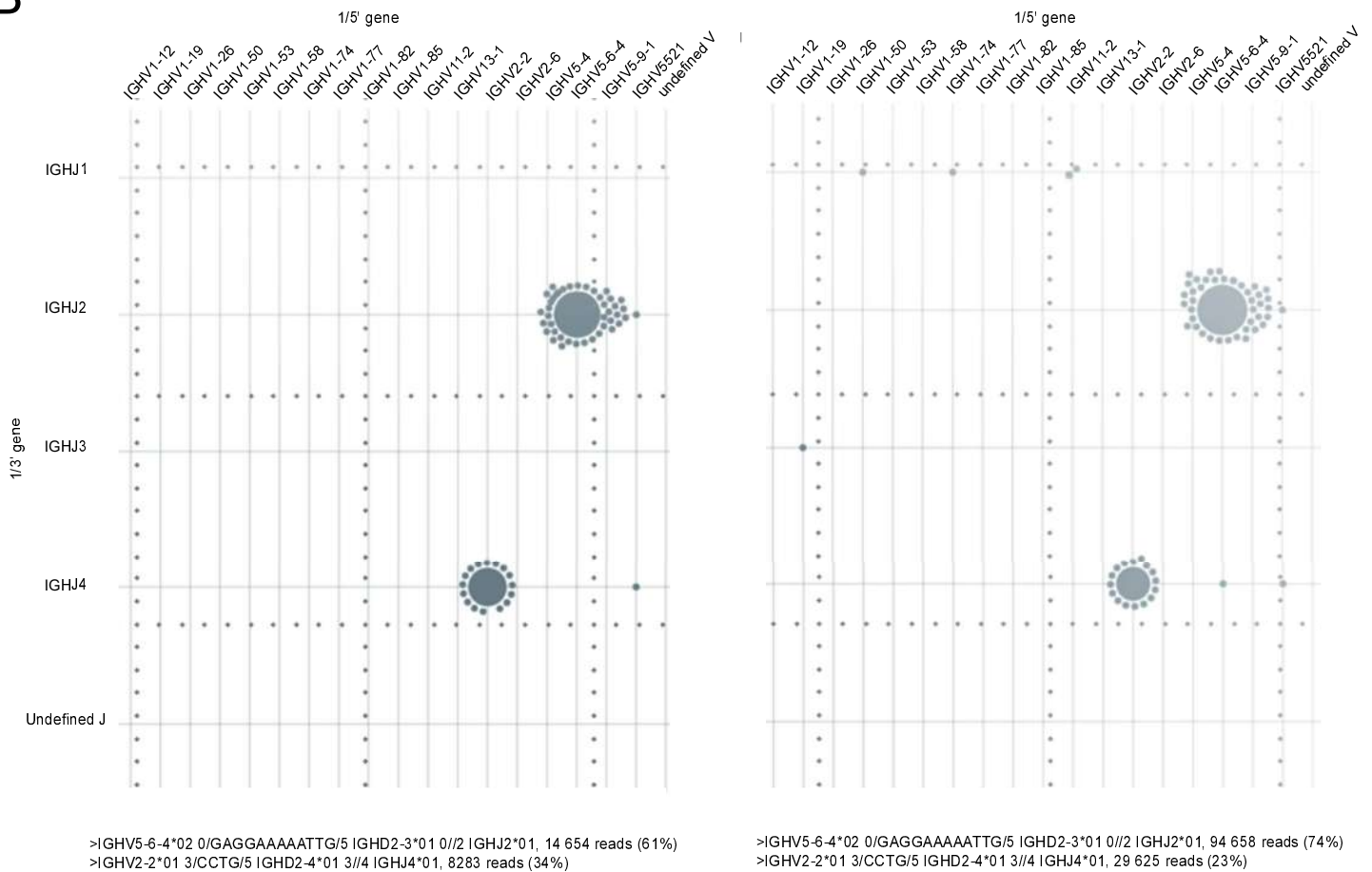


# Supplemental Figure 10

A



B



Supplemental figure 10 : Characterization of the leukemic R38H-expressing mouse

(A) Immunophenotyping of the B cells of the bone marrow (BM) and spleen of the leukemic mouse.

(B) Clonal *IGH* gene rearrangements of the BM cells from the leukemic mouse. *IGH* gene rearrangements were analyzed by PCR through IgH-primers binding to framework regions 1 and 3 (respectively FR1 and FR3) to determine clonality of B cells. Results were visualized with the Vidjil software. The rearrangement are displayed as circles whose size is proportional to the number of reads carrying rearrangements (which is assumed to reflect the proportion of each rearrangement in the sample). The coordinates of the positions of the circles correspond to the *IGH* gene segments involved in each rearrangement (V segments on the x-axis and J segments on the y-axis). The clonal rearrangements are detailed below the figure

## Supplemental Table 1

Species	Gene	Alignment				AA
Human	ENSG00000196092	N	FVNGRPLPDVV	R	QRIVELAHQGV	C 38
Human mutated		N	FVNGRPLPDVV	H	QRIVELAHQGV	C 38
<i>P. troglodytes</i>	ENSPTRG00000020936	N	FVNGRPLPDVV	R	QRIVELAHQGV	C 38
<i>M. mulatta</i>	ENSMMUG00000019639	N	FVNGRPLPDVV	R	QRIVELAHQGV	C 38
<i>F. catus</i>	ENSFCAG00000001628	N	FVNGRPLPDVV	R	QRIVELAHQGV	C 38
<i>M. musculus</i>	ENSMUSG00000014030	N	FVNGRPLPDVV	R	QRIVELAHQGV	C 38
<i>G. gallus</i>	ENSGALG00000016431	N	FVNGRPLPDVV	R	QRIVELAHQGV	C 38
<i>T. rubripes</i>	ENSTRUG00000003227	N	FVNGRPLPDVV	R	QRIVELAHQGV	C 38
<i>D. rerio</i>	ENSDARG000000037383	N	FVNGRPLPDVV	R	QRIVELAHQGV	C 38
<i>D. melanogaster</i>	FBgn0005561	N	FVNGRPLPDVV	R	QRIVELAHNGV	C 197
<i>C. elegans</i>	K06B9.5	N	FVNGRPLPDTI	R	AQIVEMSQHGT	C 113
<i>X. tropicalis</i>	ENSXETG00000006547	N	FVNGRPLPDVV	R	QRIVELAHQGV	C 38

Supplemental Table 1: Orthologue conservation of the Arginine 38 (table from <http://www.mutationtaster.org>). Position of mutated amino acid (in red) or its correspondence (AA) and *Ensembl* references are indicated.

## Supplemental Table 2



PAX proteins	Alignment																																				
Human PAX5	V	N	Q	L	G	G	V	F	V	N	G	R	P	L	P	D	V	V	R	Q	R	I	V	E	L	A	H	Q	G	V	R	P	C	D	I	S	R
Human PAX1	V	N	Q	L	G	G	V	F	V	N	G	R	P	L	P	N	A	I	R	L	R	I	V	E	L	A	Q	L	G	I	R	P	C	D	I	S	R
Human PAX2	V	N	Q	L	G	G	V	F	V	N	G	R	P	L	P	D	V	V	R	Q	R	I	V	E	L	A	H	Q	G	V	R	P	C	D	I	S	R
Human PAX3	V	N	Q	L	G	G	V	F	I	N	G	R	P	L	P	N	H	I	R	H	K	I	V	E	M	A	H	H	G	I	R	P	C	V	I	S	R
Human PAX4	M	N	Q	L	G	G	L	F	V	N	G	R	P	L	P	L	D	T	R	Q	Q	I	V	R	L	A	V	S	G	M	R	P	C	D	I	S	R
Human PAX6	V	N	Q	L	G	G	V	F	V	N	G	R	P	L	P	D	S	T	R	Q	K	I	V	E	L	A	H	S	G	A	R	P	C	D	I	S	R
Human PAX7	V	N	Q	L	G	G	V	F	I	N	G	R	P	L	P	N	H	I	R	H	K	I	V	E	M	A	H	H	G	I	R	P	C	V	I	S	R
Human PAX8	L	N	Q	L	G	G	A	F	V	N	G	R	P	L	P	E	V	V	R	Q	R	I	V	D	L	A	H	Q	G	V	R	P	C	D	I	S	R
Human PAX9	V	N	Q	L	G	G	V	F	V	N	G	R	P	L	P	N	A	I	R	L	R	I	V	E	L	A	Q	L	G	I	R	P	C	D	I	S	R

Supplemental Table 2: Parologue conservation of the Arginine 38 among the nine human PAX proteins. GenBank accession numbers for human PAX1 to 9 are NM\_006192, NM\_000278, NM\_181458, NM\_006193, NM\_016734, NM\_000280, NM\_001135254, NM\_003466 and NM\_006194 respectively. Letters in open boxes indicate amino acid discrepancy and in closed boxes identities compared to the PAX5 protein (in gray). The red arrow indicates the position of Arginine 38 in the PAX5 protein.



## Supplemental Table 3

Subject	Stage	Mutations
I.1	Asymptomatic	No mutation
I.2	Asymptomatic	NM_016734.3(PAX5_v001):c.113G>A, NM_016734.3(PAX5_i001):p.(Arg38His)
II.1§	BCP-ALL diagnosis	NM_016734.3(PAX5_v001):c.113G>A, NM_016734.3(PAX5_i001):p.(Arg38His) NM_016734.3(PAX5_v001):c.1110delinsTGGGTCCCCATTAT, NM_016734.3(PAX5_i001):p.(Tyr371Glyfs*37)
II.1§	Complete remission	NM_016734.3(PAX5_v001):c.113G>A, NM_016734.3(PAX5_i001):p.(Arg38His)
II.3	Asymptomatic	NM_016734.3(PAX5_v001):c.113G>A, NM_016734.3(PAX5_i001):p.(Arg38His) (VAF 50%)
II.3	Sampled as HSCT donor <sup>‡</sup>	NM_016734.3(PAX5_v001):c.113G>A, NM_016734.3(PAX5_i001):p.(Arg38His) (VAF 50%)
II.3	BCP-ALL diagnosis	NM_016734.3(PAX5_v001):c.113G>A, NM_016734.3(PAX5_i001):p.(Arg38His) (VAF 50%) NM_016734.3(PAX5_v001):c.419G>T, NM_016734.3(PAX5_i001):p.(Arg140Leu) (VAF 38%) NM_033360.4(KRAS_v001):c.38G>A, NM_033360.4(KRAS_i001):p.(Gly13Asp) (VAF 32%) NM_002524.5(NRAS_v001):c.38G>A, NM_002524.5(NRAS_i001):p.(Gly13Asp) (VAF 1%)
II.3	Complete remission	NM_016734.3(PAX5_v001):c.113G>A, NM_016734.3(PAX5_i001):p.(Arg38His) (VAF 50%)
II.2	BCP-ALL diagnosis	NM_016734.3(PAX5_v001):c.113G>A, NM_016734.3(PAX5_i001):p.(Arg38His) (VAF 50%) NM_033360.4(KRAS_v001):c.35G>A, NM_033360.4(KRAS_i001):p.(Gly12Asp) (VAF 13%) NM_002524.5(NRAS_v001):c.35G>A, NM_002524.5(NRAS_i001):p.(Gly12Asp) (VAF 2%) NM_002524.5(NRAS_v001):c.38G>A, NM_002524.5(NRAS_i001):p.(Gly13Asp) (VAF 5%)
II.2	BCP-ALL relapse	NM_016734.3(PAX5_v001):c.113G>A, NM_016734.3(PAX5_i001):p.(Arg38His) (VAF 50%) NM_002524.5(NRAS_v001):c.38G>A, NM_002524.5(NRAS_i001):p.(Gly13Asp) (VAF 5%) NM_001754.4(RUNX1_v001):c.364G>A, NM_001754.4(RUNX1_i001):p.(Gly122Arg) (VAF 2%) NM_001127208.2(TET2_v001):c.3602G>A, NM_001127208.2(TET2_i001):p.(Arg1201His) (VAF 4%)
II.3/2	BCP-ALL second relapse	NM_016734.3(PAX5_v001):c.113G>A, NM_016734.3(PAX5_i001):p.(Arg38His) (VAF 50%) NM_016734.3(PAX5_v001):c.419G>T, NM_016734.3(PAX5_i001):p.(Arg140Leu) (VAF 48%) NM_002524.5(NRAS_v001):c.35G>A, NM_002524.5(NRAS_i001):p.(Gly12Asp) (VAF 5%) NM_002834.4(PTPN11_v001):c.1508G>A, NM_002834.4(PTPN11_i001):p.(Gly503Glu) (VAF 43%)

Supplemental Table 3: Secondary mutations. Variant allele frequencies (VAF) for *PAX5* mutations were determined with droplet digital PCR. VAF for other mutations were determined by high-throughput sequencing. § High-throughput sequencing was not performed in this sample but only *PAX5* Sanger sequencing was performed. <sup>‡</sup> *PAX5* exon 4 c.419G>T:p.Arg140Leu <0.01%; fs\*: frameshift.

## Supplemental Table 4

Subject	Stage	Type	Chr.	Cytoband Start	Cytoband End	CN status	Start	End	Size (kb)	Genes	Interpretation
II.3	BCP-ALL diagnosis	Loss	2	q23.3	q23.3	1.0	154338514	154589071	250.557	no gene	germline*
II.3	BCP-ALL diagnosis	Loss	9	p21.3	p21.3	1.0	20560634	22514989	1954.355	30 genes including <i>CDKN2A</i> and <i>CDKN2B</i>	somatic
II.3	BCP-ALL diagnosis	Loss	9	p21.3	p21.3	0.0	21012118	21981583	969.465	24 genes including <i>CDKN2A</i>	somatic
II.3	BCP-ALL diagnosis	Loss	11	p15.1	p15.1	1.0	21370750	21497068	126.318	<i>NELL1</i>	germline*
II.3	BCP-ALL diagnosis	Loss	18	q21.2	q21.2	1.0	52987745	53755336	767.591	<i>TCF4</i> , <i>MIR4529</i> , <i>LOC100505474</i>	somatic
II.3	BCP-ALL diagnosis	Loss	22	q11.22	q11.22	1.0	22530228	22600828	70.600	<i>VPREB1</i>	somatic
II.3	complete remission	Loss	2	q23.3	q23.3	1.0	154338514	154589071	250.557	no gene	germline*
II.3	complete remission	Loss	11	p15.1	p15.1	1.0	21370750	21497068	126.318	<i>NELL1</i>	germline*
II.2	BCP-ALL diagnosis	Loss	9	p21.3	p21.3	1.0	21931526	22622698	691.172	<i>C9orf53</i> , <i>CDKN2A</i> , <i>CDKN2B-AS1</i> , <i>CDKN2B</i> , <i>DMRTA1</i>	somatic
II.2	BCP-ALL diagnosis	Loss	9	p21.3	p21.3	0.0	21976745	22015434	38.689	<i>CDKN2A</i> , <i>CDKN2B-AS1</i> , <i>CDKN2B</i>	somatic
II.2	BCP-ALL diagnosis	Loss	15	q14	q14	1.0	38427109	38548296	121.187	<i>SPRED1</i>	somatic
II.2	BCP-ALL relapse	Loss	7	p22.3	p11.1	1.0	43360	58025873	57982.513	439 genes	somatic
II.2	BCP-ALL relapse	Loss	9	p21.3	p21.3	1.0	21931526	22622698	691.172	<i>C9orf53</i> , <i>CDKN2A</i> , <i>CDKN2B-AS1</i> , <i>CDKN2B</i> , <i>DMRTA1</i>	somatic §
II. <sup>3</sup> / <sub>2</sub>	donor-related BCP-ALL	Loss	2	q23.3	q23.3	1.0	154338514	154589071	250.557	no gene	germline*
II. <sup>3</sup> / <sub>2</sub>	donor-related BCP-ALL	Loss	9	p21.3	p21.3	1.0	20659491	22231643	1572.152	28 genes including <i>CDKN2A</i> and <i>CDKN2B</i>	somatic
II. <sup>3</sup> / <sub>2</sub>	donor-related BCP-ALL	Loss	9	p21.3	p21.3	0.0	21881618	22011644	130.026	<i>C9orf53</i> , <i>CDKN2A</i> , <i>CDKN2B-AS1</i> , <i>CDKN2B</i>	somatic
II. <sup>3</sup> / <sub>2</sub>	donor-related BCP-ALL	Gain	10	p15.3	q26.3	3.0	100026	135427143	135327.117	whole chromosome	somatic
II. <sup>3</sup> / <sub>2</sub>	donor-related BCP-ALL	Loss	11	p15.1	p15.1	1.0	21370750	21497068	126.318	<i>NELL1</i>	germline*
II. <sup>3</sup> / <sub>2</sub>	donor-related BCP-ALL	Gain	14	q11.2	q32.33	3.0	20511672	107285437	86773.765	whole chromosome	somatic
II. <sup>3</sup> / <sub>2</sub>	donor-related BCP-ALL	Loss	20	q13.11	q13.32	1.0	42088007	56530487	14442.480	158 genes	somatic
II. <sup>3</sup> / <sub>2</sub>	donor-related BCP-ALL	Gain	X	p22.33	p22.31	2.0	168546	155233731	155065.185	whole chromosome	somatic

Supplemental Table 4: Copy-number abnormalities identified by single nucleotide polymorphism (SNP) array. For each aberration, chromosome location, copy number (CN) status, size and genes involved are indicated. \* These copy number variations were used as informative polymorphisms to confirm relationship between II.3 and donor-related BCP-ALL in subject II.2 (II<sup>3</sup>/<sub>2</sub>). § The low infiltration probably limited the resolution in this sample. Thus, homozygous *CDKN2A* deletion was likely to be missed.

## Supplemental Table 5

Pedigrees	Subjects	Germline mutation PAX5 (NM_016734)	Affected domain	Diagnosis	Age at diagnosis (years)	Cytogenetics	Acquired PAX5 abnormality	Acquired biallelic loss of CDKN2A	Ref.
1	IV.6 (proband)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	2	46,NN,t(1;9)(p13;p21),del(8)(p21),i(9)(q10)	Yes (LOH)	Yes	2
1	IV.5 (sibling)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	3	46,NN,i(9)(q10),del(17)(p11.2) [relapse]	Yes (LOH)	Yes	2
1	III.5 (parent's half-sibling)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	1	ND	ND	ND	2
1	IV.1 (cousin)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	<1	ND	ND	ND	2
1	IV.3 (cousin)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	8	ND	ND	ND	2
1	III.3 (parent)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier	NA	NA	NA	NA	2
1	IV.9 (sibling)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier	NA	NA	NA	NA	2
1	II.3 (grandparent)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier*	NA	NA	NA	NA	2
1	III.2 (parent's sibling)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier*	NA	NA	NA	na	2
2	III.2 (proband)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	5	47,NN,i(9)(q10),+mar	Yes (LOH)	Yes	2
2	IV.1 (child)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	2	47,NN,i(9)(q10),+11 [diagnosis] 44,NN,dic(4;9)(q10;p10),dic(9;12)(q10;q10),add(16)(q24) [relapse]	Yes (LOH)	Yes	2
2	IV.2 (child)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	2	46,NN,i(9)(q10),+11 [diagnosis] 46,NN,-9,i(9)(q10),add(12)(p12),add(14)(q32),+mar [relapse]	Yes (LOH)	Yes	2
2	III.4 (cousin)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	8	45,NN,dic(9;21)(p11;p11.1)	Yes (LOH)	ND	2
2	II.4 (parent's cousin)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	2	ND	ND	ND	2
2	II.2 (parent)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier	NA	NA	NA	NA	2
2	I.1 (grandparent)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier*	NA	NA	NA	NA	2
2	I.2 (grandparent's sibling)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier*	NA	NA	NA	NA	2
2	II.3 (parent's sibling)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier*	NA	NA	NA	NA	2
3	C6 (proband)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	1	del(9)(p13p22)	Yes (LOH)	Yes	3
3	C7 (sibling)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	2	i(9)(q10)	Yes (LOH)	Yes	3
3	C3 (cousin)	c.547G>A:p.Gly183Ser	Octapeptide	BCP-ALL	5	dic(9;12)(p11-13;p11)	Yes (LOH)	Yes	3
3	A1 (grandparent)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier	NA	NA	NA	NA	3
3	B2 (parent's sibling)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier	NA	NA	NA	NA	3
3	B3 (parent's sibling)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier	NA	NA	NA	NA	3
3	B4 (parent)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier	NA	NA	NA	NA	3
3	C2 (cousin)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier	NA	NA	NA	NA	3
3	C4 (cousin)	c.547G>A:p.Gly183Ser	Octapeptide	unaffected carrier	NA	NA	NA	NA	3
4	II.1 (proband)	c.113G>A:p.Arg38His	Paired domain	BCP-ALL	18	ND	ND	ND	4
4	I.1 (parent)	c.113G>A:p.Arg38His	Paired domain	unaffected carrier	NA	NA	NA	NA	4
5	II.2 (proband)	c.113G>A:p.Arg38His	Paired domain	BCP-ALL	17	46,XY [diagnosis] 46,XY,del(7)(p10) [relapse]	No	Yes	This study
5	II.1 (sister)	c.113G>A:p.Arg38His	Paired domain	BCP-ALL	11	ND	Yes	ND	This study
5	II.3 (brother)	c.113G>A:p.Arg38His	Paired domain	BCP-ALL	25	46,XY	Yes	Yes	This study
5	I.1 (mother)	c.113G>A:p.Arg38His	Paired domain	unaffected carrier	NA	NA	NA	NA	This study
5	I.I3/2 (after proband-graft)	c.113G>A:p.Arg38His	Paired domain	BCP-ALL relapse	40	49,XY,+X,+10,+14,del(20)(q13)	Yes	Yes	This study

Supplemental Table 5: Review of described pedigrees with germline PAX5 mutation-related BCP-ALL (B-cell precursor acute lymphoblastic leukemia); LOH: loss of heterozygosity; NA: not applicable; ND: not defined; \*untested

## Supplemental Table 6

Mouse	Engraftment at 18 weeks	<i>Ptpn11</i>	<i>Kras</i>	<i>Jak3</i>		VAF %
		Mutation	Mutation	Mutation		
MIG #2	Yes	NM	NM	NM		
MIG #3	Yes	NM	NM	NM		
MIG #4	Yes	NM	NM	NM		
MIG #5	Yes	NM	NM	NM		
MIG #6	Yes	NM	NM	NM		
PAX5 WT #3	No	NM	NM	NM		
PAX5 R38H #1	No	NM	NM	NM		
PAX5 R38H #2	Yes	NM	NM	NM		
PAX5 R38H #3	No	NM	NM	NM		
PAX5 R38H #4	Yes	NM	NM	V670G	NM_001190830.1:c.2009T>G, p.Val670Gly	92

Supplemental Table 6: Murine targeted NGS on bone marrow of recipient mice. The engraftment status and the mutational status of *Ptpn11*, *Kras* and *Jak3* genes and the variant allele frequency (VAF) of the mutation if applicable are indicated. NM for not mutated.

## Supplemental Table 7

Name	Forward primer	Reverse primer
<i>PAX5</i> exon 1	5'_TTATTCCGACCTGTGAGCGG_3'	5'_CAAACACGCCGCAGTTAGAC_3'
<i>PAX5</i> exon 2	5'_CACAGCGGTGCTTCTCCTAT_3'	5'_GGCTACTTGGAATTGGGGCT_3'
<i>PAX5</i> exon 3	5'_GGGTAACAGAGACCGAGCAG_3'	5'_GGCCAGAGTAGCCCGTTATT_3'
<i>PAX5</i> exon 4	5'_CACAGGGAAACAGAAGCAAAGT_3'	5'_CGCATTAGTACGTGTGCTGAAG_3'
<i>PAX5</i> exon 5	5'_CTCGTGGCTCAGGTGTGG_3'	5'_CTCGCTCCTCTGCAGGTAAG_3'
<i>PAX5</i> exon 6	5'_AGTTGGGGTCAGGTCCTCTT_3'	5'_TG TAGCGACTGCCCAAGTTT_3'
<i>PAX5</i> exon 7	5'_GGCACAGGCATGGACTAACT_3'	5'_TCTCATCACAGCTGCAACCC_3'
<i>PAX5</i> exon 8	5'_GAAGCGTAGAGGTCACCCAG_3'	5'_GCTGGAGATGTGTGTGCATG_3'
<i>PAX5</i> exon 9	5'_GACAGCTGCCCACTCCATAA_3'	5'_ATGTGTCTGGGCAAGCTACC_3'
<i>PAX5</i> exon 10	5'_AGAGGAAGAGGGGAGCTTCA_3'	5'_TGATGGGACTTGCCATCTGG_3'

Supplemental Table 7: Primers used for *PAX5* Sanger sequencing.

## Supplemental Table 8

Primers for targeted NGS

Name	Forward primer	Reverse primer
Jak3_L507	5'_CTGCTTTCTTGCCATCCAGA_3'	5'_ACAGTTGAAGGTGGCTTCCA_3'
Jak3_A568	5'_GAAGTAGCTTCTAGTGGGACTTCTGA_3'	5'_AGAGGCATGCAGGTAGAGCA_3'
Jak3_R653	5'_GCAGGAGAGTTTGTCTACAA_3'	5'_TAGTGGGTGAGGCTAGCCTGTA_3'
Kras_G12	5'_AGACTGTAGAGCAGCGTTACCTCTAT_3'	5'_TTGATAATCTTGTGTGAGACATGTTCT_3'
Ptpn11_E76	5'_TGGTCACGAAGATGTCACTGA_3'	5'_ATACGCCTTCTCTCAATGGACTA_3'

Primers for *IGH* FR1 and FR3 rearrangements

Region	Name	Forward Primer	Name	Reverse primer
FR1	MH1	5'_CAGGTGCAGCTKMAGGAGTCWGG_3'	JH1_4	5'_CTTACCTGAGGAGACGGTGAC_3'
	MH2	5'_CAGGTCCARCTGCAGCAGYCTGG_3'	JH2	5'_AGGACTCACCTGAGGAGAC_3'
	MH3	5'_GAGGTYCAGCTGCARCAGTCTGG_3'	JH3	5'_AGGACTCACCTGCAGAGAC_3'
	VH1A	5'_AGGTGCAGCTTCAGGAGTCAGG_3'		
	VH3A	5'_AGGTGAAGCTGGTGGAGTCTG_3'		
	VH3B	5'_AGGTGAAGCTTCTCGAGTCTGG_3'		
	VH3C	5'_AAGTGAAGCTTGAGGAGTCTGG_3'		
	FR3	VH1-1	5'_CCAGCACAGCCTACATGSAGCT_3'	JH1_4
Vh1-2		5'_CAAGGCCACAYTGA CTGYAGA_3'	JH2	5'_AGGACTCACCTGAGGAGAC_3'
VH2_8		5'_CTSASCATCASCAAGGACA ACTC_3'	JH3	5'_AGGACTCACCTGCAGAGAC_3'
VH3_12		5'_ACTMGWGACACWTCYAAGA ACCA_3'		
VH4_11		5'_ACCCTGTACCTGCARATGAGCAA_3'		
VH5		5'_TACCTGCAAATGASCAGTCTGA_3'		
VH6_7_10_13		5'_TACCTGCAAATGASCAGTCTGA_3'		
VH9_14_15		5'_ATGAAYRSCYTRAGAGCTGARGA_3'		

Supplemental Table 8: Primers used for targeted NGS of murine *Jak3*, *Kras*, and *Ptpn11* genes and for murine *IGH* rearrangement study.

### Legend Supplemental Figures:

Supplemental Figure 1: **Timeline of the clinical evolution of each member of the reported family with a germline *PAX5* p.Arg38His mutation.**

Each dot represents one sample. The ages of onset are depicted with a timeline under the figure. Germline mutation presence is indicated with a + symbol. Acquired *PAX5* mutations are detailed in green colour. Death of the patient is indicated with a stop line; fs: frameshift.

Supplemental Figure 2: **Location of the p.Arg38His variant in the *PAX5* protein.**

(A) Tertiary structure of the DNA-binding domain of the *PAX5* protein. Arginine 38 (Arg38) is indicated in yellow colour (from <https://www.uniprot.org>). (B) Comparative structure of the wild-type (Arginine) and mutant (Histidine) amino acids (C) Close-up of the tertiary structure of the mutated site. The side chain of the wild-type or the mutant amino acid is coloured in green or red respectively, the rest of the protein is coloured in grey (from HOPE web application: <http://www.cmbi.ru.nl/hope/>)<sup>1</sup>.

Supplemental Figure 3: **BCP-ALL samples show *CDKN2A* deletion.**

SNP-array karyotyping showing cryptic biallelic deletions of *CDKN2A* at BCP-ALL stage in II.2 and II.3.

Supplemental Figure 4: ***IGH* rearrangements of the two brothers.**

*IGH* rearrangements were visualized with the Vidjil software. *IGH* clonotypes are displayed as circles whose size is proportional to the number of reads carrying a rearrangement (which is assumed to reflect the proportion of each rearrangement in the sample). The coordinates of the circles reflect the *IGH* genes involved in each rearrangement (V segments in the x axis, J segments in the y axis). The main clones found in leukemic samples from II.3 (panel B) and II.2 (panel C) are highlighted by arrows and their respective specific sequences are detailed below the figure. Both leukemic samples harbored leukemia-specific *IGH* rearrangements that were absent in sample from the individual II.3 (panel A) at time of stem cell donation.

Supplemental Figure 5: **Age of onset of BCP-ALL in patients with germline PAX5 mutations.**

Boxes show the interquartile range divided by the median value for age at BCP-ALL diagnosis in patients with germline PAX5 p.Arg38His mutation (n=4) versus PAX5 p.Gly183Ser mutation (n=13).

Supplemental Figure 6: **PAX5 R38H does not rescue PAX5 WT activity and does not exert dominant-negative effect on PAX5 WT in 558L $\mu$ M cell line.**

Left panel: Experimental scheme of 558L $\mu$ M cell line retroviral complementation. Right panel: 558L $\mu$ M cells were co-transduced with GFP (MIG)- and mCherry (MImCherry)-expressing vectors (labeled as empty vectors), with MIG and MImCherry-PAX5 WT (labeled as PAX5 WT), MImCherry and MIG-PAX5 R38H (labeled as PAX5 R38H) or MImCherry-PAX5 WT and MIG-PAX5 R38H (labeled as WT + R38H) retroviral vectors. Surface IgM (sIgM) expression was assessed by flow cytometry as a reflection of CD79 transactivation; cells were previously gated on GFP<sup>+</sup>/mCherry<sup>+</sup> criteria. Data are representative of two independent experiments (n=3 and n=4 independent infections per condition). Results are expressed as mean  $\pm$  SD; \* p<0.05, \*\*p<0.01, \*\*\*p<0.001.

Supplemental Figure 7: **Clonogenic activity of transduced Pax5<sup>-/-</sup> fetal cells determined by *in vitro* limiting dilution assays.**

(A) Quantification by FACS of GFP<sup>+</sup>/B220<sup>+</sup> cells per well after one-week expansion. Bars represent means of 12 replicates per seeding density. Numbers of positive wells (more than 15 cells per well) and corresponding proportions in percentages are indicated. (B) Quantification of estimated frequency of transduced Pax5<sup>-/-</sup> fetal cells with ELDA (Extreme Limiting Dilution Analysis). ELDA (<http://bioinf.wehi.edu.au/software/elda/>)<sup>5</sup> was performed and the estimated frequency and range (95% Confidence Interval, 95% CI [lower value; upper value]) of transduced Pax5<sup>-/-</sup> cells are indicated in graph and table. The plain lines represent



regression line fitting to the experimental data and the dotted lines represent the regression lines for the 95% CIs based on the likelihood ratio test of single-hit model. LDA analysis given by ELDA software also provides a Chi-square test resulting in non-significant difference between MIG and PAX5 R38H conditions.

Supplemental Figure 8: **GSEA of PAX5 target gene set and pro-B cell gene set in rescued *Pax5*<sup>-/-</sup> cells.**

GSEA of PAX5 target gene set and pro-B cell gene set<sup>6</sup> in *Pax5*<sup>-/-</sup> cells transduced with PAX5 WT versus empty MIG. False Discovery Rate (FDR) < 0.01; NES: Normalized Enrichment Score.

Supplemental Figure 9 **Comparison of the PAX5-R38H and the PAX5-G183S mutant transcriptomes.**

(A) Bioinformatics procedure design. Shah *et al.* dataset (GSE45260)<sup>2</sup> was reanalyzed and intersected with our dataset to determine a consensus list of genes modified by PAX5 WT in both experimental systems. (B) Overlap of the up-regulated genes and down-regulated genes for foldchange >|1.5| with adjusted p-value <0.05. Significance of hypergeometric test used to determine if the merged gene list could be explained only by chance is indicated. (C) GSVA score was calculated for the up-regulated and down-regulated consensus gene sets for every single samples (MIG and MIR empty vectors as negative controls; PAX5-WT as positive control; PAX5-R38H and PAX5-G183S as mutants). Moderated t-test was used to compare each condition to negative control (\*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001). (D) Consensus gene sets used for the GSVA scoring. (E) Principal component analysis (PCA) of the two best components for the most differentially expressed genes between the three conditions for each study (Shah *et al.* <sup>2</sup>, upper panel; present study, lower panel).

Supplemental Figure 10: **Characterization of the leukemic R38H-expressing mouse.**

(A) Immunophenotyping of the B cells of the bone marrow (BM) and spleen of the leukemic mouse. (B) Clonal *IGH* gene rearrangements of the BM cells from the leukemic mouse.

*IGH* gene rearrangements were analyzed by PCR through IgH-primers binding to framework regions 1 and 3 (respectively FR1 and FR3) to determine clonality of B cells. Results were visualized with the Vidjil software. The rearrangement are displayed as circles whose size is proportional to the number of reads carrying rearrangements (which is assumed to reflect the proportion of each rearrangement in the sample). The coordinates of the positions of the circles correspond to the *IGH* gene segments involved in each rearrangement (V segments on the x-axis and J segments on the y-axis). The clonal rearrangements are detailed below the figure

## Legend Supplemental Tables:

### Supplemental Table 1: **Orthologue conservation of the Arginine 38.**

(Table from <http://www.mutationtaster.org>). Position of mutated amino acid (in red) or its correspondence (AA) and *Ensembl* references are indicated.

### Supplemental Table 2: **Parologue conservation of the Arginine 38 among the nine human PAX proteins.**

GenBank accession numbers for human PAX1 to 9 are NM\_006192, NM\_000278, NM\_181458, NM\_006193, NM\_016734, NM\_000280, NM\_001135254, NM\_003466 and NM\_006194 respectively. Letters in open boxes indicate amino acid discrepancy and in closed boxes identities compared to the PAX5 protein (in gray). The red arrow indicates the position of Arginine 38 in the PAX5 protein.

### Supplemental Table 3: **Secondary mutations.**

Variant allele frequencies (VAF) for *PAX5* mutations were determined with droplet digital PCR. VAF for other mutations were determined by high-throughput sequencing. § High-throughput sequencing was not performed in this sample but only *PAX5* Sanger sequencing was performed. † *PAX5* exon 4 c.419G>T;p.Arg140Leu <0.01%.; fs\*: frameshift.

### Supplemental Table 4: **Copy-number abnormalities identified by single nucleotide polymorphism (SNP) array.**

For each aberration, chromosome location, copy number (CN) status, size and genes involved are indicated. \* These copy number variations were used as informative polymorphisms to confirm relationship between II.3 and donor-related BCP-ALL in subject II.2 (II<sup>3/2</sup>). § The low infiltration probably limited the resolution in this sample. Thus, homozygous *CDKN2A* deletion was likely to be missed.

Supplemental Table 5: **Review of described pedigrees with germline *PAX5* mutation-related BCP-ALL.**

LOH: loss of heterozygosity; NA: not applicable; ND: not defined; \*untested.

Supplemental Table 6: **Murine targeted NGS on bone marrow of recipient mice.**

The engraftment status and the mutational status of *Ptpn11*, *Kras* and *Jak3* genes and the variant allele frequency (VAF) of the mutation if applicable are indicated. NM for not mutated.

Supplemental Table 7: **Primers used for *PAX5* Sanger sequencing.**

Supplemental Table 8: **Primers used for targeted NGS** of murine *Jak3*, *Kras*, and *Ptpn11* genes and for murine *IGH* rearrangement study.

**References mentioned in Supplemental Figures and Tables:**

1. Venselaar, H., Te Beek, T. A., Kuipers, R. K., Hekkelman, M. L. & Vriend, G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics* **11**, 548 (2010).
2. Shah S, Schrader KA, Waanders E, et al. A recurrent germline PAX5 mutation confers susceptibility to pre-B cell acute lymphoblastic leukemia. *Nat Genet.* 2013;45(10):1226-1231.
3. Auer F, Ruschendorf F, Gombert M, et al. Inherited susceptibility to pre B-ALL caused by germline transmission of PAX5 c.547G>A. *Leukemia.* 2014;28(5):1136-1138.
4. Yazdanparast S, Khatami SR, Galehdari H, Jaseb K. One missense mutation in exon 2 of the PAX5 gene in Iran. *Genet Mol Res.* 2015;14(4):17768-17775.
5. Hu Y, Smyth GK. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods.* 2009;347(1-2):70-78.
6. Revilla IDR, Bilic I, Vilagos B, et al. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J.* 2012;31(14):3130-3146.

## **Supplemental material**

### **Methods**

#### ***Consent and compliance with ethical guidelines***

Informed consent was obtained from all individuals at diagnosis. The study was conducted according to the declaration of Helsinki and was approved by the Human Research Committee of Lille and the internal review board of the Tumor Bank from the University Hospital Center of Lille (certification NF 96900-2014/65453-1).

Mice were housed under specific pathogen-free conditions (Anexplo UMS006 CREFRE, Toulouse, France) in accordance with the European Directive 2010/63/EU and the French Institutional Guidelines for animal handling. Mice were handled according to protocols approved by the Regional Ethical Committee (#A31555010).

#### ***DNA samples***

Human DNA was extracted from bone marrow (BM) at BCP-ALL diagnosis or relapse and from peripheral blood in complete remission using the QIAamp Tissue Kit (Qiagen) following the manufacturer's instructions. For individual II.1, DNA was obtained from BM smears at BCP-ALL diagnosis. Murine DNA extraction of BM samples was performed using a QIAamp DNA Mini Kit (#51304) following manufacturer's instruction (Qiagen).

#### ***Human whole Exome Sequencing (WES)***

Exome sequencing was performed for individual II.3 at diagnosis and complete remission and II.2 at relapse. The genomic library was prepared using the Nextera Rapid Capture Exome Kit (Illumina, San Diego, CA) on a HiSeq2500 (Illumina, San Diego, CA) with 2x100 HiSeq SBS Kit v4 chemistry. The Fastq files were aligned to the human genome assembly GRCh37 (hg19) with bwa (v0.7.12)<sup>1</sup>, followed by IndelRealigner from the GATK toolkit<sup>2</sup>. Variants were called with Freebayes<sup>3</sup> and filtered using an in-house database selecting only high quality (>40) variants. Synonymous, deep intronic, intergenic variants and variants already documented in 1000g/ESP with a MAF  $\geq$  1% were discarded.

### **Human SNP-array karyotyping**

Genomic DNA from II.2 (at BCP-ALL diagnosis, relapse and donor-related BCP-ALL) and II.3 (at BCP-ALL diagnosis and complete remission) were hybridized to Cytoscan HD array (Affymetrix) following the manufacturer's protocol. Data were analyzed using the Chromosome Analysis Suite (ChAS) software (Affymetrix) as previously described<sup>4</sup>. Alterations were validated by visual inspection and annotated using GRCh37.

### **Mutations analysis**

The coding sequence of *PAX5* was investigated using Sanger sequencing and the relevant primers described in Supplemental Table 7.

The frequency of somatic *PAX5* c.419G>T:p.Arg140Leu in II.2 (at time of donor-related BCP-ALL) and II.3 (at BCP-ALL diagnosis and at time of sample collection for HSCT) was quantified using Droplet Digital PCR assays by Bio-Rad reagents, primers and probes (HEX-labeled wild-type allele; FAM-labeled mutant alleles) in duplicate using 90 ng of DNA. PCR products were determined using the QX100 droplet reader (Bio-Rad) and QuantaSoft software (version 1.7.4.0917, Bio-Rad) with a detection limit of 0.01%.

For R38H and R140L resequencing, a PCR amplicon based library was generated with two specific primers (Fwd : 5'-TTTGTGAATGGACGGCCACT-3' and Rv : 5'-TGGTTGGGTGGCTGCTGTA-3') to obtain an amplicon including the two mutations. The library was sequenced with a MiSeq sequencer (Illumina, San Diego, USA) and Miseq Reagent kit V2 (paired-end sequencing 2 x 150 cycles). Alignment was performed using BWA aligner and visualized with IGV.

Targeted resequencing of 36 genes recurrently mutated in hematological malignancies<sup>5</sup> was performed in II.3 (at BCP-ALL diagnosis) and II.2 (at BCP-ALL diagnosis, relapse and donor-related BCP-ALL). Libraries were prepared using the Ampliseq System following the manufacturer's instructions and run on Ion Proton (ThermoFisher, Waltham, MA, USA).

Murine libraries were generated using a multiplex PCR covering 7 recurrent regions in a murine *PAX5*-ELN model (Supplemental Table 8 and <sup>6</sup>) and run on a MiSeq sequencer (Illumina, San

Diego, USA) with a Miseq Reagent Kit V2 (paired-end sequencing 2 x 150 cycles). Alignment was performed using bwa and variant calling was performed using FreeBayes and Mutect2 variant callers.

### ***IGH rearrangements sequencing***

Immunoglobulin heavy chain (*IGH*) rearrangements were investigated by high-throughput sequencing in individuals II.3 (at time of stem cell donation and BCP-ALL diagnosis) and II.2 (at time of donor-related BCP-ALL). For every sample, *IGH* rearrangements were amplified on genomic DNA with a high fidelity polymerase (FastStart™ High Fidelity PCR System dNTPack, Roche Dignostics, Meylan, France) with primers described by the Biomed-2 group as described previously<sup>7</sup>. For human, the library was prepared using the Ion Plus Fragment Library Kit according to the manufacturer's instructions and run on Ion S5 (Thermo Fisher, Waltham, MA, USA). The raw Ion Torrent flow was transformed to demultiplexed FastQ sequences using the Torrent Server (Thermo Fisher) and the reads were processed using the Vidjil software ([www.vidjil.org](http://www.vidjil.org)). For mouse *IGH* rearrangements, the library was generated using FR1 and FR3 specific primers (Supplemental Table 2).

### ***Cell lines, primary cells and culture media***

Phoenix-eco retroviral packaging cell line ( $\phi$ NXe)<sup>8</sup> was cultured in Dulbecco's Modified Eagle Medium (DMEM Gibco #31966) containing 10% fetal bovine serum (FBS, Gibco #10270). Murine plasmacytoma cell line 558L $\mu$ M was kindly provided by Pr. Michael Reth<sup>9</sup> and grown in RPMI-1640 media (Sigma #R0883) as previously described<sup>10</sup>. The murine stromal cell lines MS5 and OP9 (ATCC CRL-2749)<sup>11</sup> were grown in Alpha Minimum Essential Medium (Gibco MEM  $\alpha$ , no nucleosides #22561) supplemented with 20% FBS. *Pax5*<sup>-/-</sup> fetal liver cells were harvested at E17.5, dissociated on a 70  $\mu$ m cell strainer and co-cultured on murine stromal cells in IMDM with 5% FBS for Mouse Pre-B Lymphoid Cells (Stemcell Technologies #06350) and 5 ng/mL recombinant murine IL-7 (Peprotech #217-17). All transduction experiments were performed after an initial *Pax5*<sup>-/-</sup> pro-B-cell expansion on stromal cells with 2 ng/mL of IL-7.

### **DNA constructs and retroviral production**

PAX5 WT and PAX5 R38H complementary DNAs (cDNAs) were synthesized, verified by Sanger sequencing and sub-cloned into pMSCV-IRES-GFP II (MIG) and pMSCV-IRES-mCherryFP II plasmids (MImCherry, gift from Dr. Dario Vignali<sup>12</sup>) to generate bicistronic IRES-driven retroviral vectors co-expressing PAX5 WT with mCherryFP (MImCherry-PAX5 WT) or GFP (MIG-PAX5 WT) or PAX5 R38H with GFP (MIG-PAX5 R38H). The Phoenix-eco retroviral packaging system/cell line ( $\phi$ NXe) (<http://www.stanford.edu/group/nolan>) was used to generate ecotropic retrovirus. Retroviral supernatants were produced by transient transfection of Phoenix cells pre-incubated in DMEM (Gibco #61965), one hour before to transfection and kept along transfection. After overnight incubation, cells were treated for 6 h with 10 mM sodium butyrate (Sigma #B5887) diluted in medium. Viral supernatants were harvested after at least 24 h of incubation at 32°C and filtered (0.45  $\mu$ m filter, Merck #SE2M230104).

### **Mice**

C57BL/6J mice were purchased from Charles River Laboratories (France) and CD45.1 B6.SJL congenic mice from the Jackson Laboratory (Bar Harbor, ME, USA). *Pax5*<sup>-/-</sup> mice were obtained by breeding *Pax5*<sup>+/-</sup> mice kindly provided by Pr. Meinrad Busslinger<sup>13</sup> (Vienna, Austria) and backcrossed on C57BL/6J background.

### **Retroviral transductions and differentiation assays**

Retroviral transductions were performed at 1,000 x g at 32°C for 90 min. Half a million 558L $\mu$ M cells were transduced in six-well dishes with 3 mL of retroviral supernatants diluted at 1:3 in presence of 8  $\mu$ g/mL polybrene (Sigma #H9268). Flow analysis was performed 3 days post-transduction.

The *ex vivo* pro-B-cell expansion and differentiation protocol used was adapted from Rolink *et al.*<sup>14</sup>. Fifty thousand *Pax5*<sup>-/-</sup> pro-B cells were co-cultivated on OP9-derived stromal cells in 100 mm dishes for one week in presence of 2 ng/mL of IL-7 and transduced in 24-well plates during exponential growth. Retroviral supernatants were diluted at 1:3, supplemented with 5.4  $\mu$ g/mL polybrene and incubated overnight at 37°C. After one week expansion, pro-B cells were washed out of IL-7 and plated on stromal cells in pro-B medium with or without IL-7 for the



differentiation assay. After 3 days, single-cell suspensions were stained using standard protocols and subjected to flow cytometry. For transcriptome and transplantation assays, retrovirally transduced pro-B cells were then sorted on the GFP expression criteria.

### ***Flow cytometry and cell sorting***

BM cell suspensions were initially enriched in B cells using CD45R(B220) immunomagnetic microbeads (#130-049-501 Miltenyi Biotec) and separated with QuadroMACS™ Starting LS Kit (#130-091-051 (Miltenyi Biotec). This B-cell fraction was subsequently stained with V500-labelled anti-CD45R/B220 (clone RA3-6B2; BD Biosciences # 561227), BV711-labelled anti-CD117/Kit (clone 2B8; BD Biosciences #563160), FITC-labelled anti-Igk (clone 187.1; BD Biosciences #550003) antibodies. B-cell differentiation was analyzed using BV421-labelled anti-CD45R/B220 (clone RA3-6B2; BD Biosciences #562922), APC-conjugated anti-CD19 (clone 1D3; eBioscience #17-0193-82), BB700-labelled anti-IgM (clone II/41; BD Biosciences #746099) and PE-Cy7 anti-Igk (clone 187.1; BD Biosciences #560667) antibodies. The 558L $\mu$ M cells were stained with BB700-conjugated anti-mouse IgM (clone II/41; BD Biosciences #746099). Marker expressions were analyzed in the GFP and mCherryFP double-positive populations.

### ***Transplantation assay***

A 30 mg·kg<sup>-1</sup> sublethal dose of busulfan (Sigma #B2635) was injected<sup>15</sup> into the peritoneal cavity of 6-week-old recipient mice 24 h before intravenously transplantation of 15,000 transduced GFP-positive sorted cells. Chimerism was quantified by flow cytometry (FACS) as the percentage of donor-derived cells (% CD45.2<sup>+</sup> cells) found in the CD45.1 recipient BM and spleen.

### ***Limiting Dilution Assays (LDA)***

*Pax5*<sup>-/-</sup> fetal cells were transduced as described above and sorted four days later as GFP<sup>+</sup>/B220<sup>+</sup>/Kit<sup>+</sup> at 1, 3, 9, 27, 81, 243 or 729 cells per well (12 replicates per seeding density) in 96-well plates containing pro-B medium supplemented with 5 ng/mL IL-7 and pre-plated with 2,000 OP9 stromal cells per well.

### ***Murine transcriptome***

Transduced cells were purified according to mCherryFP or GFP expression. Total RNA was purified using an RNeasy Plus Micro Kit (#74034) following manufacturer's instruction (Qiagen). RNA quality was assessed using Agilent High Sensitivity RNA ScreenTape System (5067-5579, 5067-5580, 5067-5581) on Agilent 2200 TapeStation instrument (Agilent Technologies). RNA samples were prepared according to the manufacturer's protocol with the GeneChip Whole Transcript (WT) Pico Reagent Kit (ThermoFischer) and hybridized on Clariom D GeneChip Mouse Transcriptome 1.0 arrays.

Raw CEL files pre-processing and differential expression analysis were carried out using AltAnalyze 2.1.0<sup>16</sup>. The probesets with a detection above background (DABG) p-value above 0.05 or expression below 1 (non-log) were discarded from the downstream analysis. Gene expression levels were determined using the constitutive probeset values following by quantile normalization. Differentially expressed genes (DEGs) analysis was performed by the limma package<sup>17</sup> of the Bioconductor suite and R studio environment (version 1.2.1335). Principal component analysis (PCA) was performed with AltAnalyze using the 2,000 most differentially expressed genes, clustered by sample log fold method. Unsupervised hierarchical clustering was performed using one-minus Pearson's correlation distance metrics and the average linkage method. Genes were ranked using the t value calculated by limma. Pre-ranked gene set enrichment analysis (GSEA) from the Broad institute<sup>18</sup> was used to statistically test pathway enrichment between the samples. Statistical analysis was performed by determining false discovery rate (FDR q-value) based on 10,000 random permutations.

Transcriptome micro-array data from Shah *et al.* (GSE45260)<sup>19</sup> were reanalyzed with the same pipeline of our dataset to prevent disparity. First, PAX5 WT condition of each array study was compared to the corresponding empty vector condition to determine the differentially expressed genes ( $F_c < |1.5|$ , adjusted p-value  $< 0.05$ ) induced by PAX5 WT in each system. Next, the two PAX5 up-regulated gene lists (as well as down-regulated gene lists) from both array studies were intersected to identify the two consensus gene sets and results were represented by nVenn diagram<sup>20</sup>. Hypergeometric tests were performed to validate the

consistency of the consensus gene sets. Single sample enrichment was performed through gene set variation analysis (GSVA)<sup>21</sup> for the two consensus gene sets and moderated t-test was performed as described<sup>21</sup>.

### **Statistics**

The number of biological replicates is indicated in the legend of the relevant figures. Error bars for pooled replicates represent standard deviation (SD). Statistical differences were determined using a two-tailed unpaired Student's t test for comparison of quantitative variable, assuming normality and equal distribution of variance between the different groups analyzed. For LDA, estimated frequencies and 95% confidence intervals were determined using L-calc software (version 1.1 StemCell Tech.) and ELDA software: (<http://bioinf.wehi.edu.au/software/elda/>)<sup>22</sup>. Statistical analyses of cell growth in LDA were performed using Mann-Whitney test. All statistical analyses were performed using GraphPad Prism software, version 7 (GraphPad). A *p* value lower than 0.05 was considered as statistically significant (\**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001; \*\*\*\**p* < 0.0001).

### **References of the methods section**

1. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.
2. DePristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011;43(5):491-498.
3. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. *arXiv*. 2012;1207.3907.
4. Duployez N, Boudry-Labis E, Roumier C, et al. SNP-array lesions in core binding factor acute myeloid leukemia. *Oncotarget*. 2018;9(5):6478-6489.
5. Marceau-Renaut A, Duployez N, Ducourneau B, et al. Molecular Profiling Defines Distinct Prognostic Subgroups in Childhood AML: A Report From the French ELAM02 Study Group. *Hemasphere*. 2018;2(1):e31.
6. Jamrog L, Chemin G, Fregona V, et al. PAX5-ELN oncoprotein promotes multistep B-cell acute lymphoblastic leukemia in mice. *Proc Natl Acad Sci U S A*. 2018;115(41):10357-10362.
7. Ferret Y, Caillault A, Sebda S, et al. Multi-loci diagnosis of acute lymphoblastic leukaemia with high-throughput sequencing and bioinformatics analysis. *Br J Haematol*. 2016;173(3):413-420.
8. Pear WS, Nolan GP, Scott ML, Baltimore D. Production of high-titer helper-free retroviruses by transient transfection. *Proc Natl Acad Sci U S A*. 1993;90(18):8392-8396.

9. Hombach J, Leclercq L, Radbruch A, Rajewsky K, Reth M. A novel 34-kd protein co-isolated with the IgM molecule in surface IgM-expressing cells. *EMBO J.* 1988;7(11):3451-3456.
10. Maier H, Colbert J, Fitzsimmons D, Clark DR, Hagman J. Activation of the early B-cell-specific mb-1 (Ig-alpha) gene by Pax-5 is dependent on an unmethylated Ets binding site. *Mol Cell Biol.* 2003;23(6):1946-1960.
11. Kodama H, Nose M, Niida S, Nishikawa S, Nishikawa S. Involvement of the c-kit receptor in the adhesion of hematopoietic stem cells to stromal cells. *Exp Hematol.* 1994;22(10):979-984.
12. Holst J, Szymczak-Workman AL, Vignali KM, Burton AR, Workman CJ, Vignali DA. Generation of T-cell receptor retrogenic mice. *Nat Protoc.* 2006;1(1):406-417.
13. Urbanek P, Wang ZQ, Fetka I, Wagner EF, Busslinger M. Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell.* 1994;79(5):901-912.
14. Rolink A, Kudo A, Karasuyama H, Kikuchi Y, Melchers F. Long-term proliferating early pre B cell lines and clones with the potential to develop to surface Ig-positive, mitogen reactive B cells in vitro and in vivo. *EMBO J.* 1991;10(2):327-336.
15. Andersson G, Illigens BM, Johnson KW, et al. Nonmyeloablative conditioning is sufficient to allow engraftment of EGFP-expressing bone marrow and subsequent acceptance of EGFP-transgenic skin grafts in mice. *Blood.* 2003;101(11):4305-4312.
16. Emig D, Salomonis N, Baumbach J, Lengauer T, Conklin BR, Albrecht M. AltAnalyze and DomainGraph: analyzing and visualizing exon expression data. *Nucleic Acids Res.* 2010;38(Web Server issue):W755-762.
17. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.
18. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545-15550.
19. Shah S, Schrader KA, Waanders E, et al. A recurrent germline PAX5 mutation confers susceptibility to pre-B cell acute lymphoblastic leukemia. *Nat Genet.* 2013;45(10):1226-1231.
20. Perez-Silva JG, Araujo-Voces M, Quesada V. nVenn: generalized, quasi-proportional Venn and Euler diagrams. *Bioinformatics.* 2018;34(13):2322-2324.
21. Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics.* 2013;14:7.
22. Hu Y, Smyth GK. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods.* 2009;347(1-2):70-78.



## Partie III : Conclusions & Discussion

### Modèle murin de leucémogénèse B induite par l'oncoprotéine de fusion PAX5-ELN

Étant donné la fréquence élevée des altérations de *PAX5* dans les LAL-B (Coyaud, Struski, Prade, et al., 2010; Familiades et al., 2009; Mullighan et al., 2007), *PAX5* semble une cible thérapeutique privilégiée. Cependant, les facteurs de transcription sont des protéines difficiles à cibler, ce dû à leur localisation nucléaire, et les translocations chromosomiques ne sont pas réparables. L'approche innovante des *CAR-T cells* exploite l'expression quasi-systématique du CD19 par les lymphoblastes leucémiques comme cible pour amener les lymphocytes T modifiés *ex vivo* à éliminer les blastes. Or la transcription du gène *CD19* est exclusivement activée par *PAX5*. Certes, l'expression de CD19 est quasi normale en dépit d'une inactivation mono-allélique de *PAX5* (Mullighan et al., 2007). Mais par déduction, le succès thérapeutique de cette immunothérapie peut être limité chez les patients atteints de LAL-B de sous-type *PAX5alt* puisqu'une altération du second allèle de *PAX5* survient fréquemment (Bastian et al., 2019; Gu et al., 2019). Afin de contourner ces difficultés, il est cependant envisageable de trouver des cibles thérapeutiques alternatives parmi les voies de signalisation dérégulées par l'oncoprotéine de fusion *PAX5-X*. Il est nécessaire de travailler à l'aide d'outils *in vivo* pour révéler et comprendre dans leur ensemble les mécanismes sous-jacents à l'initiation et à la progression tumorale induits par les lésions initiatrices, et potentiellement trouver de nouvelles cibles thérapeutiques ou identifier une vulnérabilité thérapeutique à exploiter – intégrant les concepts de létalité synthétique (Lord et al., 2015), de thérapie différenciatrice ou de réactivation de voie divergente comme barrière contre la transformation maligne (Chan et al., 2020). Or ces modèles *in vivo* sont rares.

Le recombinoïde de *PAX5* est imposant et le gène de l'élastine a été initialement identifié comme partenaire de fusion par notre laboratoire (Bousquet et al., 2007). Dans ce contexte, notre équipe a décidé de modéliser l'expression du transcrite de fusion *PAX5-ELN* chez la souris, issu de la translocation chromosomique t(7;9)(q11.2;p13) chez les patients. La création de modèle murin répond à un besoin d'amélioration de l'approche thérapeutique dans le cadre du traitement des LAL-B notamment à pronostic défavorable ou sujets à la rechute. Bien que le sous-type *PAX5alt* soit actuellement associé à un pronostic intermédiaire et à un pic d'incidence prépondérant chez les patients de moins de 39 ans, avec un âge médian de 10 ans au diagnostic (Tableau 6), trop peu de cas *PAX5-ELN* ont été décrits pour établir une corrélation significative entre les réarrangements t(7;9)(q11.2;p13) et une appréciation particulière du pronostic.

Case	Age	Sex	Phenotype	Karyotype	PAX5 fusion gene	Reference
1	37	M	B-III	46,XY,t(7;9)(q11;p13)[17]	PAX5 ex7/ELN ex2	[2]
2	16	M	B-II	46,XY,t(7;9)(q11;p13),del(9)(p21)[14]/ 46,XY[1]	PAX5 ex7/ELN ex2	[2]
3	4	M	B-I	46,XY,t(7;9)(q11;p13)[2]/ 46,XY[18]	PAX5 ex7/ELN ex2	[4]
4	1.4	M	B-ALL	NA (9906_037)	PAX5 ex5/ELN*	[8]
5	19.4	M	B-III	46,XY,?add(4)(q2?),?add(5)(p14),del(8)(p21),der(9)t(7;9)(q11.2;p13), del(16)(p13),inc[cp16]/46,idem,?dup(1)(q21)[3]/46,XY[4]	PAX5 ex7/ELN ex5	This work
6	3.1	M	B-III	45,XY,-7,der(9)t(7;9)(q11;p13)[15]/ 46,XY[1]	PAX5 ex6/AUTS2 ex4	[5]
7	2.8	F	B-III	45,XX,-7,der(9)t(7;9)(q11;p13),dup(16)(p11p13)[14]/ 46,XX[2]	PAX5 ex6/AUTS2 ex6	[9]
8	0.6	M	B-II/III	46,XY,t(7;9)(q11;p13)[8]	PAX5 ex6/AUTS2 ex5	[10]
9	2.1	M	B-III	46,XY,del(7)(q22q33)?,del(9)(q22?),del(12)(p11)[8]	PAX5 ex5/POM121 ex4	[1]
10	1	M	B-III	46,XY,t(7;9)(q11;p13)[21]/ 46,XY[3]	PAX5 ex5/POM121 ex4	[4]
11	1.7	F	B-ALL	45,XX,-7,der(9)t(7;9)(q11;p13)[17] 46,XX[2]	NA	[12]
12	1.3	M	B-ALL	45,XY,inv(1)(p35q32),-7,der(9)t(7;9)(q11;p13)[31] 46,XY[17]	NA	[13]

All cases, except cases 4 and 9, are listed in the Mitelman database from which also the respective karyotypes have been retrieved [11].  
\*PAX5 exon 5 is fused to an unknown ELN exon.  
NA, data not available.

Tableau 11 : Récapitulatif des caractéristiques des patients présentant une translocation chromosomique t(7; 9)(q11.2 ;p13). Tableau issu de (Denk et al., 2014).

Néanmoins, les réarrangements chromosomiques de *PAX5* avec le gène de l'élastine ont tendance à être de pronostic défavorable. En effet, parmi les patients exprimant le transcrite de fusion *PAX5-ELN* résultant de la translocation t(7;9)(q11;p13), un patient (un garçon de moins de 2 ans, pris en charge au St. Jude Children's Research Hospital, USA) s'est présenté avec une LAL-B à risque élevé, comportant une mutation de *JAK1* et un profil d'expression génique *Ph-like* (Mullighan, Zhang, et al., 2009), et deux patients (deux jeunes hommes de 16 et 37 ans) ont subi une rechute post-allogreffe et sont décédés 16 mois après leur diagnostic initial (Bousquet et al., 2007). En revanche, un patient (un jeune homme de 19 ans) a atteint la rémission complète huit mois après son diagnostic (Denk et al., 2014).

Par ailleurs, avec du recul sur le recombinome de *PAX5*, il s'avère que le locus 7q11.2 est réarrangé de manière récurrente avec *PAX5*, engendrant les fusions *PAX5-ELN*, *PAX5-AUTS2* et *PAX5-POM121* (Tableau 11 ; Denk et al., 2014). Ce dernier gène-partenaire de fusion code une nucléoporine transmembranaire. Deux des cas de *PAX5-POM121* notamment décrits sont des petits garçons de 1 an et 2 ans (Coyaud, Struski, Prade, et al., 2010; Nebral et al., 2009). L'un présentait une LAL-B à risque élevé tandis que l'autre était en rémission complète (RC) 20 mois après le diagnostic. Comptant deux rechutes précoces fatales et un patient en RC plus de 2 ans après le diagnostic, les patients (3 cas pédiatriques) exprimant *PAX5-AUTS2* (un gène-partenaire de fusion également situé au locus 7q11 mais de fonction inconnue) ont plutôt un pronostic défavorable (Coyaud, Struski, Dastugue, et al., 2010; Denk et al., 2012; Kawamata et al., 2008) étayant l'hypothèse de **translocations t(7;9)(q11.2;p13) davantage associées à un pronostic défavorable**. L'étude de Denk et al. (2014) souligne un biais de genre extrêmement rare (étant donné qu'une LAL est d'ores déjà une maladie sporadique) : 83% des patients (10/12) avec une translocation de *PAX5* avec le locus 7q11.2 (tous partenaires de fusion confondus) sont du genre masculin, donnant un sexe ratio homme/femme de 5. En particulier, tous les patients *PAX5-ELN* sont du genre masculin. Tous les cas *PAX5-AUTS2* et *PAX5-POM121* se sont présentés très jeunes, tous avant 3 ans et avec un immunophénotype principalement B-III (LAL pré-B), alors que les cas *PAX5-ELN* se distinguent par leur âge médian au diagnostic, de 16 ans, convergeant vers l'âge médian du sous-type moléculaire *PAX5Alt* (10 ans ; Kimura & Mullighan, 2020) et leurs immunophénotypes diversifiés.

Les souris transgéniques exprimant *PAX5-ELN* développent efficacement un phénotype de LAL-B de patient, comprenant un envahissement lymphoblastique de la moelle osseuse, une splénomégalie et une lymphadénopathie. Ce modèle murin *in vivo* récapitule les étapes majeures de la leucémogénèse B en référence au modèle proposé (section 3.2.5 ; Inaba et al., 2013), comprenant une période de latence minimale de 3 mois durant laquelle les souris ne manifestent pas de signes externes de pathologie.



La transformation clonale maligne est associée à l'acquisition de mutations somatiques secondaires affectant soit la voie Ras/MAPK soit la voie Jak/STAT dans 86% des cas murins, conformément aux cohortes de patients de patients pédiatriques atteints de LAL-B ainsi qu'à d'autres modèles murins de LAL-B (Dang et al., 2015; Duque-Afonso et al., 2015; Martín-Lorenzo et al., 2015, 2018; van der Weyden et al., 2015).

Il n'existe pas de consensus concernant la *cellule initiateur de leucémie (LIC, pour leukemia-initiating cell, également appelée LSC pour leukemic stem cell)* à l'origine des LAL, et souvent des rechutes. Par exemple, concernant la translocation de fusion *ETV6-RUNX1* qui peut avoir lieu *in utero*, il s'agirait de cellules  $CD34^+CD38^{low}CD19^+$  composant un mélange de cellules souches et de pro-B (D. Hong et al., 2008), alors que les gènes de fusion *PAX5-X* sont exprimés à partir du locus de *PAX5* endogène chez les patients, ce qui exclut probablement une cellule souche en tant que cellule préleucémique initiateur. Contrairement aux deux seuls modèles murins *knock-in* de transcrit de fusion *PAX5-X* publiés à ce jour (février 2021), *id est Pax5<sup>ETV6/+</sup> et Pax5<sup>Foxp1/+</sup>* exprimés de façon constitutive sous le contrôle du locus endogène *Pax5* murin (Smeenk et al., 2017), le transgène *PAX5-ELN* est exprimé sous le contrôle du locus *IgH* de la souris et son expression est régulée par le promoteur pVH et l'*enhancer* E $\mu$ . Cette construction génique permet une expression précoce du transgène *PAX5-ELN* et restreinte au compartiment lymphoïde B. Ce timing d'expression est comparable à celui des gènes de fusion *PAX5-X*, à l'instar de *TCF3-PBX1* (voir texte ci-dessous, Duque-Afonso et al., 2015) mais à la différence de la translocation de fusion *BCR-ABL1* qui doit avoir lieu dans une cellule souche hématopoïétique (CSH ; Castellanos et al., 1997; Claudia S. Huettner et al., 2003; Martín-Lorenzo et al., 2018), et du transcrit de fusion *ETV6-RUNX1* qui doit être exprimé dans les CSH et pas uniquement à partir de progéniteurs lymphoïdes B précoces pour induire un développement leucémique chez la souris (Schindler et al., 2009). Bien que le locus *IgH* possède des promoteurs et activateurs transcriptionnels puissants, notre modèle murin transgénique (*Pax5<sup>+/+</sup>* au locus *Pax*; *PAX5-ELN<sup>tg/WT</sup>* au locus *IgH*) n'induit pas de transformation leucémique fulgurante, comme c'est le cas d'au moins une des tentatives de modélisation de leucémie *Ph+*, où les souris double-transgéniques *MMTVtTA/BCR-ABL1* développent une LAL-B entre 3 semaines et 11 semaines (en fonction de la lignée fondatrice) suite à l'induction de l'expression de BCR-ABL1 (p210) grâce au retrait de la tétracycline dans les biberons d'eau (Huettner et al., 2000). De plus, nous avons vérifié que le niveau d'expression protéique était similaire entre la protéine endogène Pax5 et l'oncoprotéine PAX5-ELN, et qu'il n'y avait pas de fuite d'expression inappropriée (fraction cellulaire B220-négative de la moelle osseuse et cellules thymiques).

Notre lignée murine suffit à elle-même et parvient à un juste milieu en termes de contexte de recherche en cancérologie et d'échelle de temps : la transformation maligne aiguë survient après un délai acceptable chez une majorité de souris exprimant *PAX5-ELN*, avec une forte pénétrance de 80% à 10 mois et une médiane de survie des animaux à 6 mois. À l'inverse de notre modèle et des travaux de Castellanos et al., dont la cohorte de souris chimériques *bcr-ABLp190/bcr+* présentait une grave détresse à partir de 4 mois d'âge (Castellanos et al., 1997), les souris transgéniques exprimant BCR-ABL1 (p190) sous le contrôle du promoteur de Ly6E (*Sca1*, *stem cell antigen 1* ; Martín-Lorenzo et al., 2018) limitant son expression au compartiment des cellules souches et progéniteurs (Miles et al., 1997) développent une LAL-B avec une très faible pénétrance (13%). De plus, ce modèle tg(*TCF3-PBX1*) impose une latence de 13 mois. De même, les souris transgéniques exprimant conditionnellement *TCF3-PBX1* (via le système inductible *Cre-Lox*) développent des LAL-B avec une incidence à 12 mois variant de 7% à 53% et 59% selon le promoteur contrôlant la Cre recombinase (*Cd19*, *Mb1* ou *Mx1* respectivement ; Duque-Afonso et al., 2015), suggérant que plus l'expression est précoce durant le développement lymphoïde B, plus la pénétrance des leucémies TCF3-PBX1 est élevée. Par conséquent, ces deux groupes de chercheurs ont choisi de croiser leurs lignées murines avec la lignée hétérozygote *Pax5+/-* (Urbánek et al., 1994). L'haploinsuffisance de *Pax5* sert d'événement coopératif pour réduire la longue latence et augmenter la pénétrance des leucémies B dans leurs modèles murins.

Ainsi, notre modèle murin de LAL-B combine une latence convenable et pratique pour étudier les étapes précoces de la leucémogénèse, et une forte pénétrance rendant le modèle rentable en termes de coût de zootechnie et de bioéthique, et sans toutefois requérir un *challenge* oncogénique supplémentaire (sans intervention humaine). En effet, à la différence de *ETV6-RUNX1* nécessitant une mutagenèse chimique (Schindler et al., 2009), et de *Pax5<sup>Etv6/+</sup>* nécessitant un croisement avec une lignée murine délétée des gènes suppresseurs de tumeur *Cdkn2a/b* (Smeenk et al., 2017) pour induire la transformation maligne *in vivo*, notre modèle présente l'avantage de ne solliciter aucune manipulation expérimentale supplémentaire pour induire la leucémogénèse.

Par ailleurs, l'équipe du Professeur Busslinger a modélisé la translocation *PAX5-JAK2* en créant une lignée murine exprimant l'oncoprotéine de fusion Pax5-JAK2 à partir du locus endogène de *Pax5* murin (Jurado et al., 2015). Il s'agit d'une translocation de nature particulière car elle joue un double rôle (*dual-hit*), induisant à la fois un blocage de la différenciation B due à l'altération de *PAX5* et une suractivation de de la voie JAK/STAT, correspondant au second événement oncogénique nécessaire à transformation leucémique. En effet, *PAX5-JAK2* possède un domaine kinase actif mais perd son domaine d'auto-inhibition lors de la fusion avec *PAX5*. Par conséquent, *PAX5-JAK2* est capable de

s'autophosphoryler dans le noyau, de phosphoryler ses cibles STATs en retour et ainsi d'activer constitutivement la voie de signalisation JAK-STAT (Schinnerl et al., 2015). Jusqu'à présent, *PAX5-JAK2* est le seul réarrangement *PAX5-X* à appartenir au sous-type de LAL-B à risque élevé *Ph-like* (Roberts, 2018). Conformément à ces caractéristiques, les souris hétérozygotes *Pax5*<sup>*PAX5-JAK2/+*</sup> développent rapidement une LAL-B agressive (en 12 semaines) sans nécessiter de mutation coopérative additionnelle, corroborant l'hypothèse du double hit (Jurado et al., 2015).

Au stade préleucémique, *PAX5-ELN* octroie un potentiel d'expansion aberrant au compartiment pro-B, associé à l'expression aberrante du marqueur Kit, mais au détriment du développement des sous-compartiments de cellules B immatures et matures de la moelle osseuse dont le nombre a significativement diminué. Un tel déséquilibre est également promu par *ETV6-RUNX1* (Tsuzuki et al., 2004) et le développement des cellules B est également altéré chez les souris *PAX5-ETV6* et *PAX5-FOXP1* au stade préleucémique (Smeenk et al., 2017). Il s'avère que *PAX5-ELN* permet aux cellules pro-B normales de réacquérir des capacités et un potentiel de CSH, conformément au modèle de leucémogénèse (section 3.2.5; Inaba et al., 2013) qui postule que la lésion initiatrice confère des propriétés d'auto-renouvellement aux cellules initiatrices, CSH ou progéniteurs lymphoïdes. Ainsi, *PAX5-ELN* comme *TCF3-PBX1* confèrent des propriétés d'auto-renouvellement spécifiquement aux cellules pro-B (Duque-Afonso et al., 2015) à l'image des clones préleucémiques générées par l'expression d'*ETV6-RUNX1* (D. Hong et al., 2008). Malgré cette accumulation aberrante de clones pro-B, les souris transgéniques *PAX5-ELN* exhibent une distribution des compartiments hématopoïétiques matures (lymphoïde B/lymphoïde T/myéloïde) comparables aux souris sauvages contrôles. Ce n'est pas un trait de caractère isolé. Par exemple, la protéine de fusion *ETV6-RUNX1* stimule également l'expansion des clones préleucémiques sans impacter sur la différenciation myéloïde malgré sa présence dans des CSH (Tsuzuki et al., 2004).

Intrinsèquement, *PAX5-ELN* n'est pas capable de restaurer *in vitro* la fonction différenciatrice de *PAX5* sauvage dans des cellules déficientes en *Pax5* et les cellules pro-B *Pax5*<sup>*+/+*</sup> *tg(PAX5-ELN)* ne sont pas non plus capables de se différencier *ex vivo*, suggérant que *PAX5-ELN* induit un arrêt de la différenciation en antagonisant la fonction de *Pax5* sauvage. Pourtant, allant à l'encontre de l'effet d'inhibition compétitive de *PAX5-ELN* démontré *in vitro* (Bousquet et al., 2007), *PAX5-ELN* n'exerce pas de d'effet trans-dominant-négatif *in vivo*. En effet, les cibles transcriptionnelles activées par *Pax5* dans les cellules pro-B ne sont pas inhibées par *PAX5-ELN* et les cibles transcriptionnelles réprimées par *Pax5* ne sont pas réexprimées dans les pro-B exprimant *PAX5-ELN* (cf GSEA réalisées par Vincent Fregona ; Jamrog, Chemin, Fregona et al., 2018). Comme démontré ainsi par ce modèle murin, *PAX5-ELN* est permissif

vis-à-vis du déroulement de la différenciation B en présence de l'expression endogène de Pax5 *in vivo*. Similairement, ni PAX5-ETV6 ni PAX5-FOXP1 n'interfère avec l'expression de la plupart des gènes régulés par Pax5 au stade préleucémique *in vivo* (Smeenk et al., 2017) alors qu'un rôle dominant-négatif leur avait été attribué *in vitro* (Fazio et al., 2008; Mullighan et al., 2007). Outre la réattribution d'un potentiel d'auto-renouvellement, PAX5-ELN réintroduit un état de quiescence dans un sous-ensemble de cellules pro-B, résultant de la reprogrammation oncogénique induite par PAX5-ELN (Fregona et al., 2021), à l'image du maintien des CSH en phase quiescente (G0) induit par ETV6-RUNX1 (Schindler et al., 2009). Collectivement, ces caractéristiques permettent d'expliquer comment l'oncoprotéine PAX5-ELN induit (ainsi indirectement) un blocage partiel de la différenciation lymphoïde B *in vivo* au fur et à mesure du développement murin, sans pour autant appliquer un effet dominant-négatif sur la forme normale de Pax5. Ainsi, le rôle des protéines de fusion PAX5-X ne se limite ni à une perte de fonction due à la perte des domaines de régulation transcriptionnelle de PAX5 (TAD/ID), ni à un simple effet de compétiteur inhibitif. Certaines protéines de fusion (telles que PAX5-ETV6, PAX5-DACH1, PAX5-POM121) pourraient également activer des gènes cibles (Fazio et al., 2013; Fortschegger et al., 2014). Chez la souris, PAX5-ELN induit une signature moléculaire spécifique et active notamment des programmes transcriptionnels d'adhésion et de migration cellulaire, à l'instar de *Pax5*<sup>ETV6/+</sup> *in vivo* (Smeenk et al., 2017). De plus, dans les blastes leucémiques, PAX5-ETV6 induit une suractivation transcriptionnelle de *LCK* qui en retour induira l'hyper-phosphorylation de STAT5 (V. Cazzaniga et al., 2015). En comparaison, le mode d'action d'ETV6-RUNX1 se traduit vraisemblablement par une répression constitutive des gènes cibles de RUNX1 s'accompagnant d'une perte de sa fonction transactivatrice par perte d'association avec p300 (Zelent et al., 2004).

En conclusion, les souris exprimant le transgène PAX5-ELN fournissent un robuste modèle *in vivo* de leucémogénèse progressive multi-étapes. Ce modèle murin est donc un atout pour investiguer le stade préleucémique et il peut éventuellement être exploité comme outil de criblage ou d'évaluation préclinique. Ce modèle murin en particulier ne se réduit pas à seulement représenter les patients exprimant PAX5-ELN mais plus globalement des altérations de *PAX5* initiatrices de leucémogénèse (sous-type de LAL-B *PAX5alt*, soit 10% des cas de LAL-B), et pourrait s'étendre à modéliser les mécanismes en commun avec les LAL-B ETV6-RUNX1 qui correspondent à 25% des LAL-B pédiatriques.

Il existe des porteurs sains exprimant les transcrits de fusion *BCR-ABL1*, *ETV6-RUNX1*, et *TCF3-PBX1*, mais pas de *PAX5-X* à ma connaissance. La présence de tels oncogènes considérés comme des oncogènes « faibles » ne conduit pas nécessairement au développement de la leucémie (section 3.5.6) mais génère principalement une susceptibilité. À l'inverse, les oncogènes de fusion *PAX5-X* (notamment *PAX5-ETV6*, *PAX5-JAK2* et *PAX5-*

*ELN*) sont davantage considérés comme des oncogènes « puissants », produisant des oncoprotéines de fusion fortes de leur fonction de facteurs de transcription, capables de promouvoir la leucémogénèse en induisant une reprogrammation oncogénique incluant une empreinte transcriptionnelle et un potentiel de réservoir de clones préleucémiques ayant la capacité de s'auto-renouveler. *PAX5* est avant tout un gène suppresseur de tumeur haploinsuffisant, en lien avec son rôle central dans le maintien de l'identité B (surnommé « gardien de l'identité B », Cobaleda, Schebesta, et al., 2007) qui est elle-même considérée comme une barrière métabolique depuis récemment (Chan & Müschen, 2017). En effet, *PAX5*, en collaboration avec *IKZF1*, fonctionne comme un *gatekeeper* métabolique (« gardien métabolique ») puisqu'il permet notamment de restreindre l'apport en glucose dans les cellules B qui pourrait alimenter les clones leucémiques en devenir (Martín-Lorenzo et al., 2018; Sadras et al., 2020), élargissant davantage son rôle pivot dans le développement normal et pathologique des cellules lymphoïdes B.

Cette forte potentialité soulève la question de l'addiction oncogénique à l'événement initiateur. Par exemple, la dépendance à *BCR-ABL1*, qui est considéré comme un oncogène faible (Bose et al., 1998), dépend notamment du point de cassure (*BCR-ABL1* p210 ou p190) et du temps d'exposition à la lésion initiatrice engendrant une susceptibilité à accumuler d'autres lésions. En utilisant un système de répression inductible dépendant de l'administration de tétracycline, l'expression transitoire de *BCR-ABL1*p190 est suffisante pour initier le développement leucémique tandis la suppression de l'expression de *BCR-ABL*p190 dans des souris leucémiques ne parvient pas à inverser le phénotype leucémique, indiquant que cette lésion initiatrice n'est pas requise pour maintenir la maladie chez la souris. Au contraire, le phénotype s'aggrave avec la diminution de l'expression de *BCR-ABL1* (Pérez-Caro et al., 2007). À l'inverse, la suppression de l'expression de *BCR-ABL*p210 reverse le phénotype leucémique, indiquant que l'expression continue de *BCR-ABL*p210 est requise pour le maintien de la leucémie (Huettner et al., 2000). Dans le modèle murin exprimant *BCR-ABL*p190 sous le contrôle du promoteur de *Sca1*, des blastes leucémiques ont été détectés sans l'anomalie (Martín-Lorenzo et al., 2018). Les leucémies induites par *BCR-ABL*p190 seraient davantage dépendantes du ou des driver-s additionnel-s. une fois la transformation maligne ayant eu lieu. À l'inverse, dans le modèle murin exprimant *BCR-ABL*p210 sous le contrôle du promoteur *Sca1*, le maintien de la leucémie est au moins en partie dépendant de l'expression de *BCR-ABL1* puisque l'administration d'un traitement éradiquant les cellules exprimant *BCR-ABL1* a entraîné une amélioration de la survie des souris (Pérez-Caro et al., 2009).

*ETV6-RUNX1* est également considéré comme un oncogène faible (Schindler et al., 2009), et il semble que l'oncoprotéine de fusion soit constamment nécessaire pour maintenir un phénotype leucémique (Montaño et al., 2020).

Il serait donc intéressant d'inactiver l'expression de PAX5-ELN à plusieurs stades de la leucémogénèse : dans les cellules préleucémiques et dans les lignées primaires établies à partir des blastes de moelles osseuses des souris transgéniques leucémiques. Afin d'adresser l'addiction versus l'indépendance à l'oncogène, il serait possible de réprimer l'expression des transcrits *PAX5-ELN* à l'aide de shRNA ou d'inactiver le transgène à l'aide de sgRNA (guides CRISPR) – les souris transgéniques PAX5-ELN ayant été croisées avec les souris transgéniques exprimant constitutivement l'endonucléase Cas9 (non-GFP) – et d'effectuer une série de tests fonctionnels *in vitro* (e.g. test de différenciation et LDA) et *in vivo* (e.g. tests de transplantation).

## Émergence de la composante héréditaire dans la prédisposition aux LAL-B

Il s'avère que certains des régulateurs transcriptionnels clés de l'hématopoïèse fréquemment ciblés par des mutations **somatiques** sont également affectés par des mutations **germinales** (se référer à la section 3.2.6.1. *Susceptibilité génétique*) et constituent de la sorte des facteurs de prédisposition génétique.

### **Mutations germinales délétères à forte pénétrance :**

Les découvertes simultanées du variant germlinal *PAX5 G183S* dans des familles non-apparentées ont étendu la contribution des altérations de *PAX5* dans la leucémogénèse des LAL-B (Auer et al., 2014; Shah et al., 2013). L'établissement de la récurrence de cette mutation germinale *PAX5 R38H* (Duployez et al., 2021; Yazdanparast et al., 2015) et, plus largement, la diversification des mutations germinales de *PAX5* (Auer et al., 2014; Shah et al., 2013) et l'identification d'autres variants germinaux de gènes codant des facteurs de transcription comme *ETV6* et *IKZF1* (Churchman et al., 2018; Moriyama, Metzger, et al., 2015; se référer à la section 3.2.6.1. *Susceptibilité génétique*) permettent d'asseoir une composante héréditaire de la prédisposition au développement (leucémique) d'une LAL-B (composante actuellement émergente et probablement sous-estimée jusqu'à présent). La prise en compte d'un tel facteur de risque est primordiale lorsqu'il s'agit de considérer une allogreffe de CSH provenant d'un membre de la même famille (qui représente la première option de donneur envisagé dans le cas d'un patient réfractaire aux polychimiothérapies ou récidivant), puisque la conversion leucémique des clones préleucémiques latents pourraient induire une nouvelle LAL chez le patient-hôte.

### **Plus qu'une haploinsuffisance :**

Notre étude fonctionnelle de *PAX5 R38H* met en évidence le **rôle causal** des mutations ponctuelles de *PAX5* ayant de prime abord un caractère perte-de-fonction.

Des travaux d'autres équipes de recherche fournissaient d'ores et déjà des pistes quant à la perte de fonctionnalité de *PAX5* muté à son arginine (R) en position 38. Les premiers travaux *in vitro* sur *PAX5* avaient montré que la substitution R38A réduisait l'habilité à lier la séquence consensus de gènes cibles de plus d'un facteur 10 (Czerny et al., 1993). En 2015, Dang et ses collègues avaient identifié des mutations secondaires induites par leur mutagenèse qu'ils jugeaient perte-de-fonction de *Pax5*, situées majoritairement dans le domaine PAIRED, dont R38H (Dang et al., 2015).

Quant à Martín-Lorenzo et ses collègues, ils avaient démontré la perte partielle d'activité transcriptionnelle du variant *PAX5 R38C* (Martín-Lorenzo et al., 2018; Figure ci-dessous).

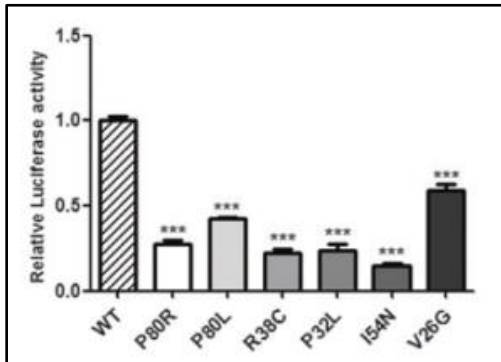


Figure issue de Martín-Lorenzo et al., 2018

Grâce à des analyses fonctionnelles et moléculaires, j'ai pu démontrer que *PAX5 R38H* agissait comme un variant fortement hypomorphique de *PAX5* (similairement à *PAX5 P80R* qui affecte également le domaine de liaison à l'ADN ; Mullighan et al., 2007), dérégulant les niveaux d'expression des gènes cibles de *PAX5* et inhibant les propriétés fonctionnelles de *PAX5* dans la différenciation lymphoïde B.

De plus, j'ai validé *in vivo* la susceptibilité au développement leucémique par des tests de transplantation syngénique. En effet, l'expression ectopique de *PAX5 R38H* dans des cellules progénitrices B déficientes en *Pax5* (*Pax5* KO) peut intrinsèquement induire le développement d'un phénotype de LAL-B chez la souris et ce dans le même laps de temps que le modèle murin *PAX5-ELN*. Leurs blastes possèdent un immunophénotype très immature, à l'instar de la moitié des tumeurs de souris soumises à une mutagenèse induite par ENU (N-ethyl-N-nitrosourea / N-nitroso-N-éthylurée) ou MMLV (Moloney Murine Leukemia Virus) qui expriment B220 mais pas CD19 ou BP1 (correspondant à la fraction de Hardy A donc au stade de développement pré-pro-B, Dang et al., 2015; Tableau ci-dessous).

Stage	CLP2	Early pro-B	Late pro-B	Pre-B	Immature B
Hardy	A	B	C	D	E
Key marker	B220	CD19	BP1	CD25	IgM
<b>N of <i>Pax5</i><sup>+/-</sup> tumors</b>					
ENU	9	7	2	3	0
MMLV	10	9	3	2	3

CD25 immunophenotypic data were available only for a subset of tumors, see supplemental Table 1. CLP, common lymphoid progenitor.

Tableau issu de Dang et al., 2015



Ainsi, nos résultats corroborent le postulat selon lequel les mutations N-terminales de PAX5 correspondent à des mutations gain-de-fonction. Nos conclusions rejoignent également les études concluant qu'une mutation génétique perte-de-fonction pouvait correspondre à un phénotype cellulaire gain-de-fonction (Ortmann et al., 2015).

Du fait de l'identification de porteurs asymptomatiques dans les familles non-apparentées, R38H semble suivre un mode de transmission récessif (Duployez et al., 2021; Yazdanparast et al., 2015) : la mutation germinale (sans coopération de mutation somatique) n'est pas suffisante pour développer une leucémie et n'induirait qu'un développement lymphocytaire B « subnormal » à l'instar de *PAX5 G183S* (Auer et al., 2014; Shah et al., 2013).

La progression leucémique chez les patients porteurs de mutations germinales de *PAX5* (9p13) est associée à un profil mutationnel spécifique : incluant une perte homozygote (bi-allélique) du suppresseur de tumeur *CDKN2A/B* (9p21.3), l'acquisition de mutations récurrentes somatiques de la voie RAS et l'acquisition d'une altération perte-de-fonction du second allèle de *PAX5* (Shah et al., 2013, Auer et al., 2014, Duployez, Jamrog et al., 2021).

### **Association récurrente entre *PAX5 R38H* et *PAX5 R140L* :**

D'un point de vue général, la substitution d'une arginine par une histidine est la substitution d'acide aminé la plus fréquente (Tsuber et al., 2017). *PAX5 R38H* correspond déjà à une mutation récurrente somatique rapportée dans au moins 11 cas de LAL-B non-familiales (Bastian et al., 2019; Gu et al., 2019). Les algorithmes de profilage transcriptomique classent la majorité des patients exprimant somatiquement *PAX5 R38H* dans le sous-type moléculaire hétérogène des altérations de *PAX5* (*PAX5alt*, Gu et al., 2019, Figure ci-dessous).

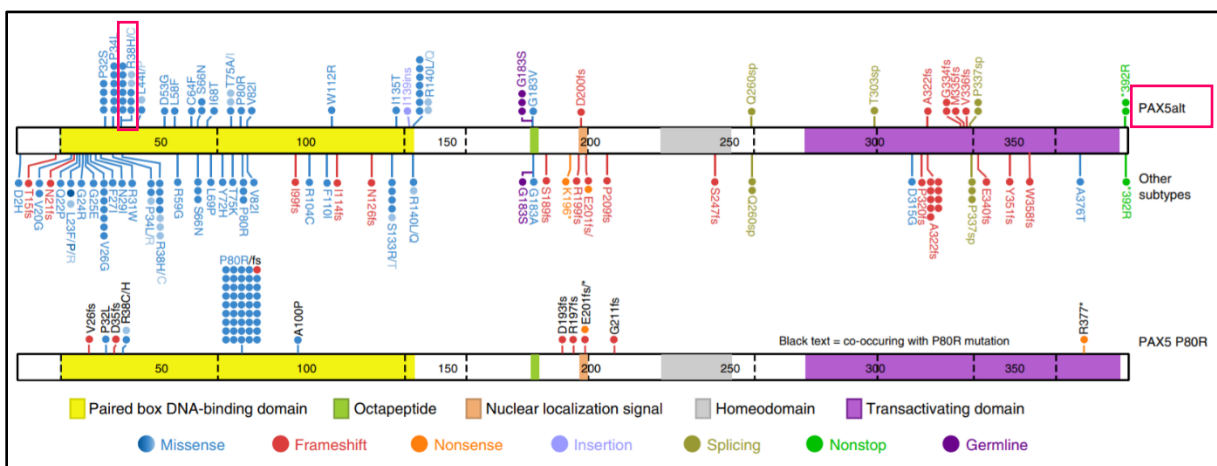


Figure issue de Gu et al., 2019.

Plus précisément, *PAX5 R38H* (affectant le sous-domaine PAI du domaine de liaison à l'ADN de *PAX5*) survient fréquemment de pair avec la mutation secondaire *PAX5 R140L* (affectant le sous-domaine RED du domaine de liaison à l'ADN de *PAX5*), conduisant à l'inactivation bi-allélique de *PAX5* et permettant de classer ces patients dans la catégorie *PAX5-plus* (Bastian et al., 2019, Figure ci-dessous).

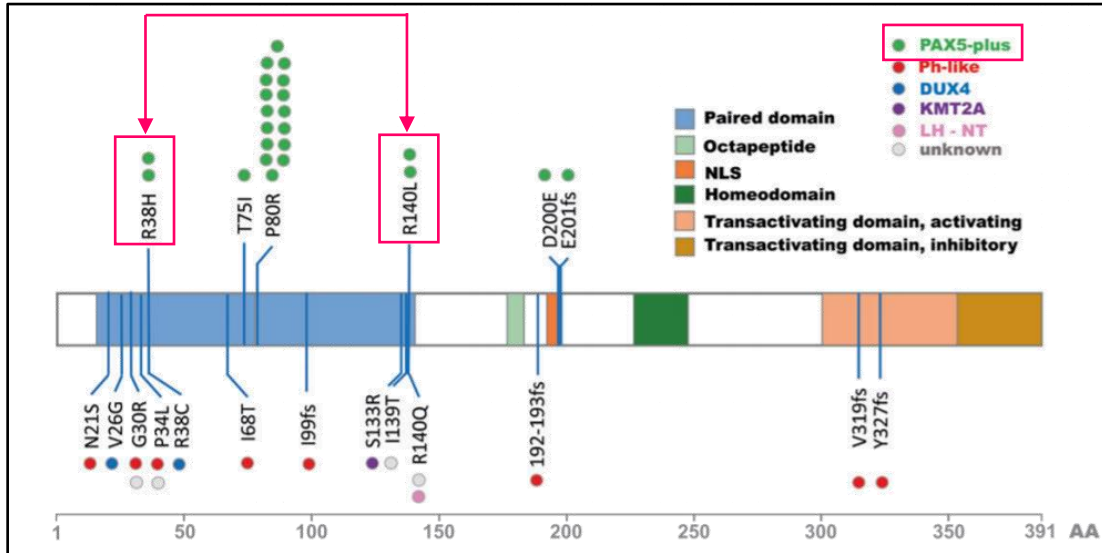


Figure issue de Bastian et al., 2019.

En somme, près de 90% (10/11 cas) des patients exprimant *PAX5 R140L* expriment aussi *PAX5 R38H*, (cohortes compilées de Gu et al., 2019 & Bastian et al., 2019). Similairement, deux des patients atteints de LAL-B familiales (II.2 et II.3, de respectivement 17 ans et 25 ans au diagnostic) pris en charge au CHU de Lille et à l'origine de cette étude fonctionnelle ont également acquis la mutation *PAX5 R140L* détectée au diagnostic (VAF à 48% et VAF à 38% respectivement). Il semblerait que les mutations *R38H* et *R140L* affectent deux allèles différents de *PAX5* (Duployez et al., 2021). Dans l'une des cohortes comparées (Bastian et al., 2019), l'âge au diagnostic des patients exprimant la mutation somatique *PAX5 R38H* (21 ans, 1 homme et 1 femme), ainsi que leurs profils mutationnels (délétion de *CDKN2A*, mutation de *NRAS*) concordent avec ceux des patients de la famille française (Tableau ci-dessous). **Cette association récurrente entre l'expression de *PAX5 R38H* et *PAX5 R140L* suggère une coopération spécifique de ces mutations en faveur de la transformation leucémique.**

Table 1 Overview of subgroup-specific genomic alterations in the *PAX5-plus* cohort ( $n = 19$ )

ID	Age (years)	Sex	Immuno-phenotype	<i>PAX5</i> sequence variants (VAF)	Chromosome 9 aberrations	<i>PAX5</i> Copy Number Variation (log2 N)	<i>CDKN2A</i> Copy Number Variation (log2 N)	RAS-activating mutations (VAF)	JAK/STAT-activating mutations (VAF)	further chromosome aberrations
BerInML3915	21	m	common	p.R38H (0.45); p.R140L (0.45)	no structural aberration	0	-2	NRAS p.G12D (0.44)		
BerIn_561921	21	f	common	p.R38H (0.53); p.R140L (0.40)	t(9;22)(p13;q13)	0	-2	NRAS p.G12D (0.15), NRAS p.G12S (0.15)		

Tableau tronqué issu de Bastian et al., 2019.

**Récurrence de mutations secondaires de *JAK3* :**

L'acquisition d'une mutation de *JAK3* (*JANus Kinase 3*, ou *Just Another Kinase 3*) constitue un point commun entre nos modèles d'étude de mutants de *PAX5*. Nous avons détecté les mutations ponctuelles *Jak3 R653H* et *Jak3 V670A* chez les souris transgéniques *PAX5-ELN* leucémiques (voir Figure 2B de Jamrog et al., 2018). De plus, un des patients avec *PAX5-ELN* comporte une délétion d'un exon de *JAK3* (c.1702-22\_1786+28del, VAF 44% ; voir Figure 2C de Jamrog et al, 2018). Quant au modèle murin d'expression ectopique de *PAX5 R38H*, une souris transplantée a acquis une mutation *Jak3 V670G* avec une VAF de 92% indiquant une expansion clonale, une caractéristique que j'ai confirmée par des résultats expérimentaux non publiés :

Souris transplantées leucémiques	Mutations activatrices de Ras- ou Jak/Stat identifiées par séquençage ciblé (NGS)	VAF (fréquence allélique du variant)
PAX5 <sup>R38H</sup> mice #4	<i>Jak3 V670G</i>	92%
PAX5 <sup>R38H</sup> mice #5	<i>Ptpn11 E69K</i> ; <i>Jak3 R653H</i>	24% ; 37%
PAX5 <sup>R38H</sup> mice #6	<i>Jak3 R653H</i> ; <i>Jak3 V674M</i>	63% ; 3,1%
PAX5 <sup>R38H</sup> mice #7	<i>Jak3 R653H</i>	94%
PAX5 <sup>R38H</sup> mice #9	<i>Ptpn11 A72V</i> ; <i>Jak3 R653H</i>	90% ; 7,8%

Par ailleurs, les souris hétérozygotes pour l'inactivation génique de *Pax5* exposées à des agents infectieux courants (zootéchnie conventionnelle) acquièrent secondairement des mutations somatiques de *Jak3* identiques (*R653H*, *V670A* ; Tableau ci-dessous de Martín-Lorenzo et al., 2015) qui engendrent un compartiment progéniteur aberrant {hyper}sensible à l'IL-7 (Martín-Lorenzo et al., 2015).

**Table 1. Genotype and phenotype of pB-ALL-bearing *Pax5*<sup>+/-</sup> mice**

Mouse number	Mouse age at disease (months)	CD19 phenotype	<i>Jak3</i> mutation in mouse <sup>a</sup>	<i>Jak3</i> mutation (allele frequency) <sup>b</sup>	<i>JAK3</i> human homolog	<i>Pax5</i> mutation in mouse <sup>a</sup>	<i>PAX5</i> human homolog	<i>Jak1</i> mutation in mouse <sup>a</sup>	
1	0361	8.8	CD19 <sup>+</sup>	R653H	R653H (50%)	R657Q		S645P	
2	0388	9.5	CD19 <sup>-</sup>	R653H	R653H (11%)	R657Q		—	
3	5665	6.5	CD19 <sup>-</sup>	—	R563H (2%)	R657Q	P80R	P80R	V657F
4	0332	15.3	CD19 <sup>-</sup>	—	—	—			—
5	W495	10.2	CD19 <sup>-</sup>	—	—	—			—
6	S748	14.2	CD19 <sup>+</sup>	V670A	V670A (27%)	V674A	P80L	Not described	—
7	W893	9.8	CD19 <sup>-</sup>	R653H	R653H (49%)	R657Q			—
8	S767	15.3	CD19 <sup>+</sup>	R653C	R653C (63%)	R657Q			—
9	W634	12	CD19 <sup>+</sup>	R653H	R653H (35%)	R657Q			—

<sup>a</sup>Mutations confirmed by Sanger sequencing.<sup>b</sup>Mutations confirmed by deep sequencing.

Tableau issu de Martín-Lorenzo et al., 2015.

La plupart des mutations ponctuelles de *JAK3* identifiées affectent ses 7 domaines de structure JH1 à JH7 avec une prédominance de hotspots mutationnels dans le JH2 correspondant au domaine pseudokinase. Les mutations dites gain-de-fonction de *JAK3* mènent à leur activation constitutive de leur activité kinase et conduisent subséquemment à l'activation aberrante de la voie de signalisation JAK/STAT, aboutissant à l'hyperactivation de la voie STAT5. Elles représentent ainsi un moyen de promouvoir la survie cellulaire. La récurrence de ces mutations de *JAK3* s'aligne avec la description d'autres mécanismes oncogéniques convergeant vers l'hyperactivation de la voie STAT5, induite par des fusions de *PAX5* comme *PAX5-ETV6* (V. Cazzaniga et al., 2015), mais également induite directement par la protéine de fusion BCR-ABL1 (Heltemes-Harris et al., 2011). Cela constitue donc un mécanisme leucémogène commun, de plus associé à un mauvais pronostic (Heltemes-Harris et al., 2011). Cela représente un signal de survie alternatif à la signalisation du pré-BCR régulée négativement dans les cellules exprimant une translocation de *PAX5* (V. Cazzaniga et al., 2015) et non fonctionnelle dans les cellules transgéniques *PAX5-ELN* (données non publiées).

#### **Des anomalies fonctionnelles deux-en-une :**

En somme, les altérations génomiques de *PAX5* conduisent inexorablement à une haploinsuffisance impactant essentiellement sa fonction de garant de l'identité B, et favorisant donc un arrêt du développement lymphoïde B à un stade précoce (lymphoblastique) propice à la prolifération et la survie cellulaires (Martín-Lorenzo et ses collègues démontrent une augmentation du nombre de cellules pro-B et pré-B et immatures B dans la MO et diminution des cellules B dans le sang périphérique). Outre cette perte de fonction, *PAX5-ELN* et *PAX5 R38H* induisent l'un comme l'autre des signatures moléculaires (transcriptomiques) spécifiques, pouvant indiquer un potentiel de reprogrammation oncogénique. La synergie de ces deux propriétés oncogéniques principales aboutit à la dérégulation du programme transcriptionnel lymphoïde B.

Dans son ensemble, ces travaux mettent en évidence l'importance de la dérégulation du réseau transcriptionnel B sans l'inhiber complètement (activité transcriptionnelle réduite, Shah et al., 2013), compte tenu que les souris hétérozygotes pour *Pax5* ne montrent aucun phénotype délétère (Urbánek et al., 1994) et présentent un développement normal de la lignée lymphoïde B (Martín-Lorenzo et al., 2015; Smeenk et al., 2017). Les souris *Pax5*<sup>+/-</sup> ne développent pas spontanément de leucémie, n'ayant pas subi de mutagenèse (Dang et al., 2015), et peuvent même vivre jusqu'à 2-3 ans dans un environnement sans pathogène (courbe de survie A issue d'Heltemes-Harris et al., 2011 ; courbe de survie B issue de Martín-Lorenzo et al., 2015 ; confirmées indépendamment par notre colonie en zootechnie SPF, observations non publiées).

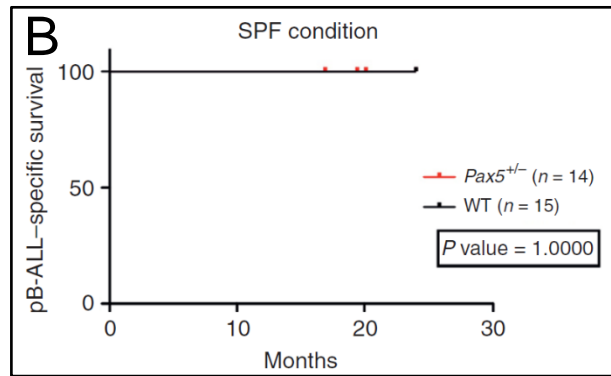
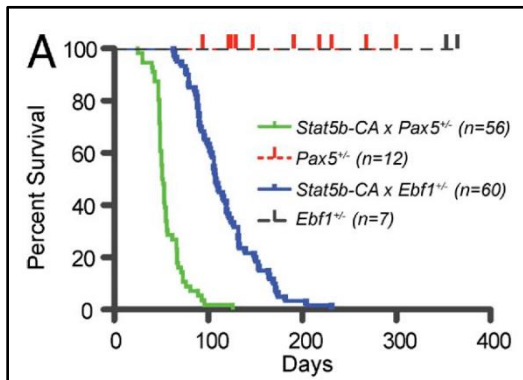


Figure A issue d'Heltemes-Harris et al., 2011. Figure B issue de Martín-Lorenzo et al., 2015.

## Références bibliographiques

- Adams, B., Dörfler, P., Aguzzi, A., Kozmik, Z., Urbánek, P., Maurer-Fogy, I., & Busslinger, M. (1992). Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS, and adult testis. *Genes & Development*, 6(9), 1589-1607. <https://doi.org/10.1101/gad.6.9.1589>
- Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C. T., Bryder, D., Yang, L., Borge, O.-J., Thoren, L. A. M., Anderson, K., Sitnicka, E., Sasaki, Y., Sigvardsson, M., & Jacobsen, S. E. W. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell*, 121(2), 295-306. <https://doi.org/10.1016/j.cell.2005.02.013>
- Åhsberg, J., Tsapogas, P., Qian, H., Zetterblad, J., Zandi, S., Månsson, R., Jönsson, J.-I., & Sigvardsson, M. (2010). Interleukin-7-induced Stat-5 acts in synergy with Flt-3 signaling to stimulate expansion of hematopoietic progenitor cells. *The Journal of Biological Chemistry*, 285(47), 36275-36284. <https://doi.org/10.1074/jbc.M110.155531>
- Alexander, T. B., Gu, Z., Iacobucci, I., Dickerson, K., Choi, J. K., Xu, B., Payne-Turner, D., Yoshihara, H., Loh, M. L., Horan, J., Buldini, B., Basso, G., Elitzur, S., de Haas, V., Zwaan, C. M., Yeoh, A., Reinhardt, D., Tomizawa, D., Kiyokawa, N., ... Mullighan, C. G. (2018). The genetic basis and cell of origin of mixed phenotype acute leukaemia. *Nature*, 562(7727), 373-379. <https://doi.org/10.1038/s41586-018-0436-0>
- Alvarez, J. V., & Frank, D. A. (2004). Genome-wide analysis of STAT target genes : Elucidating the mechanism of STAT-mediated oncogenesis. *Cancer Biology & Therapy*, 3(11), 1045-1050. <https://doi.org/10.4161/cbt.3.11.1172>
- Amin, R. H., & Schlissel, M. S. (2008). Foxo1 directly regulates the transcription of recombination-activating genes during B cell development. *Nature Immunology*, 9(6), 613-622. <https://doi.org/10.1038/ni.1612>
- An, Q., Wright, S. L., Konn, Z. J., Matheson, E., Minto, L., Moorman, A. V., Parker, H., Griffiths, M., Ross, F. M., Davies, T., Hall, A. G., Harrison, C. J., Irving, J. A., & Strefford, J. C. (2008). Variable breakpoints target PAX5 in patients with dicentric chromosomes : A model for the basis of unbalanced translocations in cancer. *Proceedings of the National Academy of Sciences of the United States of America*, 105(44), 17050-17054. <https://doi.org/10.1073/pnas.0803494105>
- Andersson, A. K., Ma, J., Wang, J., Chen, X., Gedman, A. L., Dang, J., Nakitandwe, J., Holmfeldt, L., Parker, M., Easton, J., Huether, R., Kriwacki, R., Rusch, M., Wu, G., Li, Y., Mulder, H., Raimondi, S., Pounds, S., Kang, G., ... St. Jude Children's Research Hospital–Washington University Pediatric Cancer Genome Project. (2015). The

- landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nature Genetics*, 47(4), 330-337. <https://doi.org/10.1038/ng.3230>
- Apuzzo, S., & Gros, P. (2007). Cooperative interactions between the two DNA binding domains of Pax3 : Helix 2 of the paired domain is in the proximity of the amino terminus of the homeodomain. *Biochemistry*, 46(11), 2984-2993. <https://doi.org/10.1021/bi062107q>
- Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. M., Bloomfield, C. D., Cazzola, M., & Vardiman, J. W. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*, 127(20), 2391-2405. <https://doi.org/10.1182/blood-2016-03-643544>
- Aricò, M., Schrappe, M., Hunger, S. P., Carroll, W. L., Conter, V., Galimberti, S., Manabe, A., Saha, V., Baruchel, A., Vettenranta, K., Horibe, K., Benoit, Y., Pieters, R., Escherich, G., Silverman, L. B., Pui, C.-H., & Valsecchi, M. G. (2010). Clinical Outcome of Children With Newly Diagnosed Philadelphia Chromosome–Positive Acute Lymphoblastic Leukemia Treated Between 1995 and 2005. *Journal of Clinical Oncology*, 28(31), 4755-4761. <https://doi.org/10.1200/JCO.2010.30.1325>
- Auer, F., Rüschemdorf, F., Gombert, M., Husemann, P., Ginzl, S., Izraeli, S., Harit, M., Weintraub, M., Weinstein, O. Y., Lerer, I., Stepensky, P., Borkhardt, A., & Hauer, J. (2014). Inherited susceptibility to pre B-ALL caused by germline transmission of PAX5 c.547G>A. *Leukemia*, 28(5), 1136-1138. <https://doi.org/10.1038/leu.2013.363>
- Bai, S., Li, B., Zhang, H., Jonas, J. B., Zhao, B., Shen, L., & Wang, Y. (2011). Pax6 regulates proliferation and apoptosis of human retinoblastoma cells. *Investigative Ophthalmology & Visual Science*, 52(7), 4560-4570. <https://doi.org/10.1167/iovs.10-5487>
- Balczarek, K. A., Lai, Z. C., & Kumar, S. (1997). Evolution of functional diversification of the paired box (Pax) DNA-binding domains. *Molecular Biology and Evolution*, 14(8), 829-842. <https://doi.org/10.1093/oxfordjournals.molbev.a025824>
- Balling, R., Deutsch, U., & Gruss, P. (1988). Undulated, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of Pax 1. *Cell*, 55(3), 531-535. [https://doi.org/10.1016/0092-8674\(88\)90039-6](https://doi.org/10.1016/0092-8674(88)90039-6)
- Banerjee, A., Northrup, D., Boukarabila, H., Jacobsen, S. E. W., & Allman, D. (2013). Transcriptional repression of Gata3 is essential for early B cell commitment. *Immunity*, 38(5), 930-942. <https://doi.org/10.1016/j.immuni.2013.01.014>
- Bao, E. L., Cheng, A. N., & Sankaran, V. G. (2019). The genetics of human hematopoiesis and its disruption in disease. *EMBO Molecular Medicine*, 11(8), e10316. <https://doi.org/10.15252/emmm.201910316>
- Barbosa, T. C., Lopes, B. A., Blunck, C. B., Mansur, M. B., Deyl, A. V. S., Emerenciano, M., & Pombo-de-Oliveira, M. S. (2018). A novel PAX5 rearrangement in TCF3-PBX1 acute

- lymphoblastic leukemia: A case report. *BMC Medical Genomics*, 11(1), 122. <https://doi.org/10.1186/s12920-018-0444-9>
- Barlev, N. A., Emelyanov, A. V., Castagnino, P., Zegerman, P., Bannister, A. J., Sepulveda, M. A., Robert, F., Tora, L., Kouzarides, T., Birshtein, B. K., & Berger, S. L. (2003). A novel human Ada2 homologue functions with Gcn5 or Brg1 to coactivate transcription. *Molecular and Cellular Biology*, 23(19), 6944-6957. <https://doi.org/10.1128/mcb.23.19.6944-6957.2003>
- Barr, F. G. (1997). Chromosomal translocations involving paired box transcription factors in human cancer. *The International Journal of Biochemistry & Cell Biology*, 29(12), 1449-1461. [https://doi.org/10.1016/s1357-2725\(97\)00095-2](https://doi.org/10.1016/s1357-2725(97)00095-2)
- Barrington-Trimis, J. L., Cockburn, M., Metayer, C., Gauderman, W. J., Wiemels, J., & McKean-Cowdin, R. (2015). Rising rates of acute lymphoblastic leukemia in Hispanic children: Trends in incidence from 1992 to 2011. *Blood*, 125(19), 3033-3034. <https://doi.org/10.1182/blood-2015-03-634006>
- Bassan, R., & Hoelzer, D. (2011). Modern therapy of acute lymphoblastic leukemia. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 29(5), 532-543. <https://doi.org/10.1200/JCO.2010.30.1382>
- Bastian, L., Schroeder, M. P., Eckert, C., Schlee, C., Sanchez, J. O., Kämpf, S., Wagner, D. L., Schulze, V., Isaakidis, K., Lázaro-Navarro, J., Hänzelmann, S., James, A. R., Ekici, A., Burmeister, T., Schwartz, S., Schrappe, M., Horstmann, M., Vosberg, S., Krebs, S., ... Baldus, C. D. (2019). PAX5 biallelic genomic alterations define a novel subgroup of B-cell precursor acute lymphoblastic leukemia. *Leukemia*, 33(8), 1895-1909. <https://doi.org/10.1038/s41375-019-0430-z>
- Baumann Kubetzko, F. B., Di Paolo, C., Maag, C., Meier, R., Schäfer, B. W., Betts, D. R., Stahel, R. A., & Himmelmann, A. (2004). The PAX5 oncogene is expressed in N-type neuroblastoma cells and increases tumorigenicity of a S-type cell line. *Carcinogenesis*, 25(10), 1839-1846. <https://doi.org/10.1093/carcin/bgh190>
- Bene, M. C., Castoldi, G., Knapp, W., Ludwig, W. D., Matutes, E., Orfao, A., & van't Veer, M. B. (1995). Proposals for the immunological classification of acute leukemias. European Group for the Immunological Characterization of Leukemias (EGIL). *Leukemia*, 9(10), 1783-1786.
- Beneforti, L., Dander, E., Bresolin, S., Bueno, C., Acunzo, D., Bertagna, M., Ford, A., Gentner, B., Kronnie, G. T., Vergani, P., Menéndez, P., Biondi, A., D'Amico, G., Palmi, C., & Cazzaniga, G. (2020). Pro-inflammatory cytokines favor the emergence of ETV6-RUNX1-positive pre-leukemic cells in a model of mesenchymal niche. *British Journal of Haematology*, 190(2), 262-273. <https://doi.org/10.1111/bjh.16523>



- Bennicelli, J. L., Advani, S., Schäfer, B. W., & Barr, F. G. (1999). PAX3 and PAX7 exhibit conserved cis-acting transcription repression domains and utilize a common gain of function mechanism in alveolar rhabdomyosarcoma. *Oncogene*, *18*(30), 4348-4356. <https://doi.org/10.1038/sj.onc.1202812>
- Bennicelli, J. L., Edwards, R. H., & Barr, F. G. (1996). Mechanism for transcriptional gain of function resulting from chromosomal translocation in alveolar rhabdomyosarcoma. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(11), 5455-5459. <https://doi.org/10.1073/pnas.93.11.5455>
- Blake, J. A., & Ziman, M. R. (2014). Pax genes : Regulators of lineage specification and progenitor cell maintenance. *Development*, *141*(4), 737-751. <https://doi.org/10.1242/dev.091785>
- Blokzijl, F., de Ligt, J., Jager, M., Sasselli, V., Roerink, S., Sasaki, N., Huch, M., Boymans, S., Kuijk, E., Prins, P., Nijman, I. J., Martincorena, I., Mokry, M., Wiegerinck, C. L., Middendorp, S., Sato, T., Schwank, G., Nieuwenhuis, E. E. S., Verstegen, M. M. A., ... van Boxtel, R. (2016). Tissue-specific mutation accumulation in human adult stem cells during life. *Nature*, *538*(7624), 260-264. <https://doi.org/10.1038/nature19768>
- Boisset, J.-C., & Robin, C. (2010). Imaging the founder of adult hematopoiesis in the mouse embryo aorta. *Cell Cycle (Georgetown, Tex.)*, *9*(13), 2489-2490. <https://doi.org/10.4161/cc.9.13.12319>
- Boller, S., & Grosschedl, R. (2014). The regulatory network of B-cell differentiation : A focused view of early B-cell factor 1 function. *Immunological Reviews*, *261*(1), 102-115. <https://doi.org/10.1111/imr.12206>
- Bopp, D., Burri, M., Baumgartner, S., Frigerio, G., & Noll, M. (1986). Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. *Cell*, *47*(6), 1033-1040. [https://doi.org/10.1016/0092-8674\(86\)90818-4](https://doi.org/10.1016/0092-8674(86)90818-4)
- Bose, S., Deininger, M., Gora-Tybor, J., Goldman, J. M., & Melo, J. V. (1998). The presence of typical and atypical BCR-ABL fusion genes in leukocytes of normal individuals : Biologic significance and implications for the assessment of minimal residual disease. *Blood*, *92*(9), 3362-3367.
- Bouchard, M., Souabni, A., Mandler, M., Neubüser, A., & Busslinger, M. (2002). Nephric lineage specification by Pax2 and Pax8. *Genes & Development*, *16*(22), 2958-2970. <https://doi.org/10.1101/gad.240102>
- Boudjadi, S., Chatterjee, B., Sun, W., Vemu, P., & Barr, F. G. (2018). The expression and function of PAX3 in development and disease. *Gene*, *666*, 145-157. <https://doi.org/10.1016/j.gene.2018.04.087>

- Bousquet, M., Broccardo, C., Quelen, C., Meggetto, F., Kuhlein, E., Delsol, G., Dastugue, N., & Brousset, P. (2007). A novel PAX5-ELN fusion protein identified in B-cell acute lymphoblastic leukemia acts as a dominant negative on wild-type PAX5. *Blood*, *109*(8), 3417-3423. <https://doi.org/10.1182/blood-2006-05-025221>
- Boutter, J., Huang, Y., Marovca, B., Vonderheit, A., Grotzer, M. A., Eckert, C., Cario, G., Wollscheid, B., Horvath, P., Bornhauser, B. C., & Bourquin, J.-P. (2014). Image-based RNA interference screening reveals an individual dependence of acute lymphoblastic leukemia on stromal cysteine support. *Oncotarget*, *5*(22), 11501-11512. <https://doi.org/10.18632/oncotarget.2572>
- Brodeur, G. M. (2003). Neuroblastoma: Biological insights into a clinical enigma. *Nature Reviews. Cancer*, *3*(3), 203-216. <https://doi.org/10.1038/nrc1014>
- Brun, T., Duhamel, D. L., Hu He, K. H., Wollheim, C. B., & Gauthier, B. R. (2007). The transcription factor PAX4 acts as a survival gene in INS-1E insulinoma cells. *Oncogene*, *26*(29), 4261-4271. <https://doi.org/10.1038/sj.onc.1210205>
- Buescher, J. L., & El-Hodiri, H. M. (2019). The C-terminus of the retinal homeobox (rax) gene product modulates transcription in a context-dependent manner. *Molecular Vision*, *25*, 165-173.
- Bürglin, T. R. (1994). A *Caenorhabditis elegans* prospero homologue defines a novel domain. *Trends in Biochemical Sciences*, *19*(2), 70-71. [https://doi.org/10.1016/0968-0004\(94\)90035-3](https://doi.org/10.1016/0968-0004(94)90035-3)
- Bürglin, T. R., & Affolter, M. (2016). Homeodomain proteins: An update. *Chromosoma*, *125*, 497-521. <https://doi.org/10.1007/s00412-015-0543-8>
- Burri, M., Tromvoukis, Y., Bopp, D., Frigerio, G., & Noll, M. (1989). Conservation of the paired domain in metazoans and its structure in three isolated human genes. *The EMBO Journal*, *8*(4), 1183-1190.
- Busslinger, M., Klix, N., Pfeffer, P., Graninger, P. G., & Kozmik, Z. (1996). Deregulation of PAX-5 by translocation of the Emu enhancer of the IgH locus adjacent to two alternative PAX-5 promoters in a diffuse large-cell lymphoma. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(12), 6129-6134. <https://doi.org/10.1073/pnas.93.12.6129>
- Calame, K. L., Lin, K.-I., & Tunyaplin, C. (2003). Regulatory mechanisms that determine the development and function of plasma cells. *Annual Review of Immunology*, *21*, 205-230. <https://doi.org/10.1146/annurev.immunol.21.120601.141138>

- Cambier, J. C., Gauld, S. B., Merrell, K. T., & Vilen, B. J. (2007). B-cell anergy: From transgenic models to naturally occurring anergic B cells? *Nature Reviews. Immunology*, 7(8), 633-643. <https://doi.org/10.1038/nri2133>
- Campo, E., Swerdlow, S. H., Harris, N. L., Pileri, S., Stein, H., & Jaffe, E. S. (2011). The 2008 WHO classification of lymphoid neoplasms and beyond: Evolving concepts and practical applications. *Blood*, 117(19), 5019-5032. <https://doi.org/10.1182/blood-2011-01-293050>
- Carotta, S., Willis, S. N., Hasbold, J., Inouye, M., Pang, S. H. M., Emslie, D., Light, A., Chopin, M., Shi, W., Wang, H., Morse, H. C., Tarlinton, D. M., Corcoran, L. M., Hodgkin, P. D., & Nutt, S. L. (2014). The transcription factors IRF8 and PU.1 negatively regulate plasma cell differentiation. *The Journal of Experimental Medicine*, 211(11), 2169-2181. <https://doi.org/10.1084/jem.20140425>
- Castellanos, A., Pintado, B., Weruaga, E., Arévalo, R., López, A., Orfao, A., & Sánchez-García, I. (1997). A BCR-ABL(p190) fusion gene made by homologous recombination causes B-cell acute lymphoblastic leukemias in chimeric mice with independence of the endogenous bcr product. *Blood*, 90(6), 2168-2174.
- Castro, P., Rebocho, A. P., Soares, R. J., Magalhães, J., Roque, L., Trovisco, V., Vieira de Castro, I., Cardoso-de-Oliveira, M., Fonseca, E., Soares, P., & Sobrinho-Simões, M. (2006). PAX8-PPARgamma rearrangement is frequently detected in the follicular variant of papillary thyroid carcinoma. *The Journal of Clinical Endocrinology and Metabolism*, 91(1), 213-220. <https://doi.org/10.1210/jc.2005-1336>
- Cazzaniga, G., Bisanti, L., Randi, G., Deandrea, S., Bungaro, S., Pregliasco, F., Perotti, D., Spreafico, F., Masera, G., Valsecchi, M. G., Biondi, A., & Greaves, M. (2017). Possible role of pandemic AH1N1 swine flu virus in a childhood leukemia cluster. *Leukemia*, 31(8), 1819-1821. <https://doi.org/10.1038/leu.2017.127>
- Cazzaniga, G., Daniotti, M., Tosi, S., Giudici, G., Aloisi, A., Pogliani, E., Kearney, L., & Biondi, A. (2001). The paired box domain gene PAX5 is fused to ETV6/TEL in an acute lymphoblastic leukemia case. *Cancer Research*, 61(12), 4666-4670.
- Cazzaniga, G., van Delft, F. W., Lo Nigro, L., Ford, A. M., Score, J., Iacobucci, I., Mirabile, E., Taj, M., Colman, S. M., Biondi, A., & Greaves, M. (2011). Developmental origins and impact of BCR-ABL1 fusion and IKZF1 deletions in monozygotic twins with Ph+ acute lymphoblastic leukemia. *Blood*, 118(20), 5559-5564. <https://doi.org/10.1182/blood-2011-07-366542>
- Cazzaniga, V., Bugarin, C., Bardini, M., Giordan, M., te Kronnie, G., Basso, G., Biondi, A., Fazio, G., & Cazzaniga, G. (2015). LCK over-expression drives STAT5 oncogenic

- signaling in PAX5 translocated BCP-ALL patients. *Oncotarget*, 6(3), 1569-1581. <https://doi.org/10.18632/oncotarget.2807>
- Chalepakis, G., Jones, F. S., Edelman, G. M., & Gruss, P. (1994). Pax-3 contains domains for transcription activation and transcription inhibition. *Proceedings of the National Academy of Sciences of the United States of America*, 91(26), 12745-12749. <https://doi.org/10.1073/pnas.91.26.12745>
- Chan, L. N., Murakami, M. A., Robinson, M. E., Caeser, R., Sadras, T., Lee, J., Cosgun, K. N., Kume, K., Khairnar, V., Xiao, G., Ahmed, M. A., Aghania, E., Deb, G., Hurtz, C., Shojaee, S., Hong, C., Pölönen, P., Nix, M. A., Chen, Z., ... Müschen, M. (2020). Signalling input from divergent pathways subverts B cell transformation. *Nature*, 583(7818), 845-851. <https://doi.org/10.1038/s41586-020-2513-4>
- Chan, L. N., & Müschen, M. (2017). B-cell identity as a metabolic barrier against malignant transformation. *Experimental Hematology*, 53, 1-6. <https://doi.org/10.1016/j.exphem.2017.06.004>
- Cheng, H., Zheng, Z., & Cheng, T. (2020). New paradigms on hematopoietic stem cell differentiation. *Protein & Cell*, 11(1), 34-44. <https://doi.org/10.1007/s13238-019-0633-0>
- Chi, N., & Epstein, J. A. (2002). Getting your Pax straight : Pax proteins in development and disease. *Trends in Genetics: TIG*, 18(1), 41-47. [https://doi.org/10.1016/s0168-9525\(01\)02594-x](https://doi.org/10.1016/s0168-9525(01)02594-x)
- Churchman, M. L., Qian, M., Te Kronnie, G., Zhang, R., Yang, W., Zhang, H., Lana, T., Tedrick, P., Baskin, R., Verbist, K., Peters, J. L., Devidas, M., Larsen, E., Moore, I. M., Gu, Z., Qu, C., Yoshihara, H., Porter, S. N., Pruett-Miller, S. M., ... Mullighan, C. G. (2018). Germline Genetic IKZF1 Variation and Predisposition to Childhood Acute Lymphoblastic Leukemia. *Cancer Cell*, 33(5), 937-948.e8. <https://doi.org/10.1016/j.ccell.2018.03.021>
- Clappier, E., Auclerc, M. F., Rapion, J., Bakkus, M., Caye, A., Khemiri, A., Giroux, C., Hernandez, L., Kabongo, E., Savola, S., Leblanc, T., Yakouben, K., Plat, G., Costa, V., Ferster, A., Girard, S., Fenneteau, O., Cayuela, J. M., Sigaux, F., ... Cavé, H. (2014). An intragenic ERG deletion is a marker of an oncogenic subtype of B-cell precursor acute lymphoblastic leukemia with a favorable outcome despite frequent IKZF1 deletions. *Leukemia*, 28(1), 70-77. <https://doi.org/10.1038/leu.2013.277>
- Cobaleda, C., Jochum, W., & Busslinger, M. (2007). Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature*, 449(7161), 473-477. <https://doi.org/10.1038/nature06159>

- Cobaleda, C., Schebesta, A., Delogu, A., & Busslinger, M. (2007). Pax5 : The guardian of B cell identity and function. *Nature Immunology*, 8(5), 463-470. <https://doi.org/10.1038/ni1454>
- Cooper, S. T., & Hanson, I. M. (2005). A screen for proteins that interact with PAX6 : C-terminal mutations disrupt interaction with HOMER3, DNCL1 and TRIM11. *BMC Genetics*, 6, 43. <https://doi.org/10.1186/1471-2156-6-43>
- Corces, M. R., Buenrostro, J. D., Wu, B., Greenside, P. G., Chan, S. M., Koenig, J. L., Snyder, M. P., Pritchard, J. K., Kundaje, A., Greenleaf, W. J., Majeti, R., & Chang, H. Y. (2016). Lineage-specific and single cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nature genetics*, 48(10), 1193-1203. <https://doi.org/10.1038/ng.3646>
- Corfe, S. A., & Paige, C. J. (2012). The many roles of IL-7 in B cell development; mediator of survival, proliferation and differentiation. *Seminars in Immunology*, 24(3), 198-208. <https://doi.org/10.1016/j.smim.2012.02.001>
- Coyaud, E., Struski, S., Dastugue, N., Brousset, P., Broccardo, C., & Bradtke, J. (2010). PAX5-AUTS2 fusion resulting from t(7;9)(q11.2;p13.2) can now be classified as recurrent in B cell acute lymphoblastic leukemia. *Leukemia Research*, 34(12), e323-325. <https://doi.org/10.1016/j.leukres.2010.07.035>
- Coyaud, E., Struski, S., Prade, N., Familiades, J., Eichner, R., Quelen, C., Bousquet, M., Mugneret, F., Talmant, P., Pages, M.-P., Lefebvre, C., Penther, D., Lippert, E., Nadal, N., Taviaux, S., Poppe, B., Luquet, I., Baranger, L., Eclache, V., ... Broccardo, C. (2010). Wide diversity of PAX5 alterations in B-ALL : A Groupe Francophone de Cytogenetique Hematologique study. *Blood*, 115(15), 3089-3097. <https://doi.org/10.1182/blood-2009-07-234229>
- Cozma, D., Yu, D., Hodawadekar, S., Azvolinsky, A., Grande, S., Tobias, J. W., Metzgar, M. H., Paterson, J., Erikson, J., Marafioti, T., Monroe, J. G., Atchison, M. L., & Thomas-Tikhonenko, A. (2007). B cell activator PAX5 promotes lymphomagenesis through stimulation of B cell receptor signaling. *The Journal of Clinical Investigation*, 117(9), 2602-2610. <https://doi.org/10.1172/JCI30842>
- Crawford, L. W., Foley, J. F., & Elmore, S. A. (2010). Histology atlas of the developing mouse hepatobiliary system with emphasis on embryonic days 9.5-18.5. *Toxicologic Pathology*, 38(6), 872-906. <https://doi.org/10.1177/0192623310374329>
- Cresson, C., Péron, S., Jamrog, L., Rouquié, N., Prade, N., Dubois, M., Hébrard, S., Lagarde, S., Gerby, B., Mancini, S. J. C., Cogné, M., Delabesse, E., Delpy, L., & Broccardo, C. (2018). PAX5A and PAX5B isoforms are both efficient to drive B cell differentiation. *Oncotarget*, 9(67), 32841-32854. <https://doi.org/10.18632/oncotarget.26003>

- Curran, E., & Stock, W. (2015). How I treat acute lymphoblastic leukemia in older adolescents and young adults. *Blood*, *125*(24), 3702-3710. <https://doi.org/10.1182/blood-2014-11-551481>
- Curto, G. G., Gard, C., & Ribes, V. (2015). Structures and properties of PAX linked regulatory networks architecting and pacing the emergence of neuronal diversity. *Seminars in Cell & Developmental Biology*, *44*, 75-86. <https://doi.org/10.1016/j.semcd.2015.09.010>
- Cvekl, A., Kashanchi, F., Brady, J. N., & Piatigorsky, J. (1999). Pax-6 interactions with TATA-box-binding protein and retinoblastoma protein. *Investigative Ophthalmology & Visual Science*, *40*(7), 1343-1350.
- Czerny, T., & Busslinger, M. (1995). DNA-binding and transactivation properties of Pax-6: Three amino acids in the paired domain are responsible for the different sequence recognition of Pax-6 and BSAP (Pax-5). *Molecular and Cellular Biology*, *15*(5), 2858-2871. <https://doi.org/10.1128/mcb.15.5.2858>
- Czerny, T., Schaffner, G., & Busslinger, M. (1993). DNA sequence recognition by Pax proteins: Bipartite structure of the paired domain and its binding site. *Genes & Development*, *7*(10), 2048-2061. <https://doi.org/10.1101/gad.7.10.2048>
- Dahl, E., Koseki, H., & Balling, R. (1997). Pax genes and organogenesis. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *19*(9), 755-765. <https://doi.org/10.1002/bies.950190905>
- Dang, J., Wei, L., de Ridder, J., Su, X., Rust, A. G., Roberts, K. G., Payne-Turner, D., Cheng, J., Ma, J., Qu, C., Wu, G., Song, G., Huether, R. G., Schulman, B., Janke, L., Zhang, J., Downing, J. R., van der Weyden, L., Adams, D. J., & Mullighan, C. G. (2015). PAX5 is a tumor suppressor in mouse mutagenesis models of acute lymphoblastic leukemia. *Blood*, *125*(23), 3609-3617. <https://doi.org/10.1182/blood-2015-02-626127>
- Dartsch, N., Schulte, D., Hägerling, R., Kiefer, F., & Vestweber, D. (2014). Fusing VE-Cadherin to  $\alpha$ -Catenin Impairs Fetal Liver Hematopoiesis and Lymph but Not Blood Vessel Formation. *Molecular and Cellular Biology*, *34*(9), 1634-1648. <https://doi.org/10.1128/MCB.01526-13>
- Decker, T., Pasca di Magliano, M., McManus, S., Sun, Q., Bonifer, C., Tagoh, H., & Busslinger, M. (2009). Stepwise activation of enhancer and promoter regions of the B cell commitment gene Pax5 in early lymphopoiesis. *Immunity*, *30*(4), 508-520. <https://doi.org/10.1016/j.immuni.2009.01.012>
- Dege, C., & Hagman, J. (2013). Activation of Aicda gene transcription by Pax5 in plasmacytoma cells. *Immunologic Research*, *55*(1-3), 155-161. <https://doi.org/10.1007/s12026-012-8357-8>

- DeKoter, R. P., Lee, H.-J., & Singh, H. (2002). PU.1 regulates expression of the interleukin-7 receptor in lymphoid progenitors. *Immunity*, 16(2), 297-309. [https://doi.org/10.1016/s1074-7613\(02\)00269-8](https://doi.org/10.1016/s1074-7613(02)00269-8)
- DeKoter, R. P., & Singh, H. (2000). Regulation of B lymphocyte and macrophage development by graded expression of PU.1. *Science (New York, N.Y.)*, 288(5470), 1439-1441. <https://doi.org/10.1126/science.288.5470.1439>
- Delogu, A., Schebesta, A., Sun, Q., Aschenbrenner, K., Perlot, T., & Busslinger, M. (2006). Gene Repression by Pax5 in B Cells Is Essential for Blood Cell Homeostasis and Is Reversed in Plasma Cells. *Immunity*, 24(3), 269-281. <https://doi.org/10.1016/j.immuni.2006.01.012>
- Den Boer, M. L., van Slegtenhorst, M., De Menezes, R. X., Cheok, M. H., Buijs-Gladdines, J. G. C. A. M., Peters, S. T. C. J. M., Van Zutven, L. J. C. M., Beverloo, H. B., Van der Spek, P. J., Escherich, G., Horstmann, M. A., Janka-Schaub, G. E., Kamps, W. A., Evans, W. E., & Pieters, R. (2009). A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: A genome-wide classification study. *The Lancet. Oncology*, 10(2), 125-134. [https://doi.org/10.1016/S1470-2045\(08\)70339-5](https://doi.org/10.1016/S1470-2045(08)70339-5)
- Denk, D., Bradtke, J., König, M., & Strehl, S. (2014). PAX5 fusion genes in t(7;9)(q11.2;p13) leukemia: A case report and review of the literature. *Molecular Cytogenetics*, 7(1), 13. <https://doi.org/10.1186/1755-8166-7-13>
- Denk, D., Nebral, K., Bradtke, J., Pass, G., Möricke, A., Attarbaschi, A., & Strehl, S. (2012). PAX5-AUTS2: A recurrent fusion gene in childhood B-cell precursor acute lymphoblastic leukemia. *Leukemia Research*, 36(8), e178-181. <https://doi.org/10.1016/j.leukres.2012.04.015>
- Desplan, C. (1997). Fonction des gènes Pax: Synergie de liaison à l'ADN entre le domaine paired et l'homéodomaine. *médecine/sciences*, 13(2), 147. <https://doi.org/10.4267/10608/330>
- Dias, S., Månsson, R., Gurbuxani, S., Sigvardsson, M., & Kee, B. L. (2008). E2A proteins promote development of lymphoid-primed multipotent progenitors. *Immunity*, 29(2), 217-227. <https://doi.org/10.1016/j.immuni.2008.05.015>
- Dias, S., Silva, H., Cumano, A., & Vieira, P. (2005). Interleukin-7 is necessary to maintain the B cell potential in common lymphoid progenitors. *The Journal of Experimental Medicine*, 201(6), 971-979. <https://doi.org/10.1084/jem.20042393>
- Dong, H. Y., Liu, W., Cohen, P., Mahle, C. E., & Zhang, W. (2005). B-cell specific activation protein encoded by the PAX-5 gene is commonly expressed in merkel cell carcinoma and small cell carcinomas. *The American Journal of Surgical Pathology*, 29(5), 687-692. <https://doi.org/10.1097/01.pas.0000155162.33044.4f>

- Dörfler, P., & Busslinger, M. (1996). C-terminal activating and inhibitory domains determine the transactivation potential of BSAP (Pax-5), Pax-2 and Pax-8. *The EMBO Journal*, 15(8), 1971-1982.
- Dressler, G. R., Deutsch, U., Chowdhury, K., Nornes, H. O., & Gruss, P. (1990). Pax2, a new murine paired-box-containing gene and its expression in the developing excretory system. *Development (Cambridge, England)*, 109(4), 787-795.
- Dumon, S., Santos, S. C., Debierre-Grockiego, F., Gouilleux-Gruart, V., Cocault, L., Boucheron, C., Mollat, P., Gisselbrecht, S., & Gouilleux, F. (1999). IL-3 dependent regulation of Bcl-xL gene expression by STAT5 in a bone marrow derived cell line. *Oncogene*, 18(29), 4191-4199. <https://doi.org/10.1038/sj.onc.1202796>
- Duployez, N., Jamrog, L. A., Fregona, V., Hamelle, C., Fenwarth, L., Lejeune, S., Helevaut, N., Geffroy, S., Caillault, A., Marceau-Renaut, A., Poulain, S., Roche-Lestienne, C., Largeaud, L., Prade, N., Dufrechou, S., Hébrard, S., Berthon, C., Nelken, B., Fernandes, J., ... Broccardo, C. (2021). Germline PAX5 mutation predisposes to familial B-cell precursor acute lymphoblastic leukemia. *Blood*, 137(10), 1424-1428. <https://doi.org/10.1182/blood.2020005756>
- Duque-Afonso, J., Feng, J., Scherer, F., Lin, C.-H., Wong, S. H. K., Wang, Z., Iwasaki, M., & Cleary, M. L. (2015). Comparative genomics reveals multistep pathogenesis of E2A-PBX1 acute lymphoblastic leukemia. *The Journal of Clinical Investigation*, 125(9), 3667-3680. <https://doi.org/10.1172/JCI81158>
- Duy, C., Yu, J. J., Nahar, R., Swaminathan, S., Kweon, S.-M., Polo, J. M., Valls, E., Klemm, L., Shojaee, S., Cerchietti, L., Schuh, W., Jäck, H.-M., Hurtz, C., Ramezani-Rad, P., Herzog, S., Jumaa, H., Koeffler, H. P., de Alborán, I. M., Melnick, A. M., ... Müschen, M. (2010). BCL6 is critical for the development of a diverse primary B cell repertoire. *The Journal of Experimental Medicine*, 207(6), 1209-1221. <https://doi.org/10.1084/jem.20091299>
- Dzierzak, E., & Philipsen, S. (2013). Erythropoiesis : Development and Differentiation. *Cold Spring Harbor Perspectives in Medicine*, 3(4), a011601. <https://doi.org/10.1101/cshperspect.a011601>
- Eberhard, D., & Busslinger, M. (1999). The partial homeodomain of the transcription factor Pax-5 (BSAP) is an interaction motif for the retinoblastoma and TATA-binding proteins. *Cancer Research*, 59(7 Suppl), 1716s-1724s; discussion 1724s-1725s.
- Eberhard, D., Jiménez, G., Heavey, B., & Busslinger, M. (2000). Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. *The EMBO Journal*, 19(10), 2292-2303. <https://doi.org/10.1093/emboj/19.10.2292>



- Eccles, M. R. (1998). The role of PAX2 in normal and abnormal development of the urinary tract. *Pediatric Nephrology (Berlin, Germany)*, 12(9), 712-720. <https://doi.org/10.1007/s004670050533>
- Emelyanov, A. V., Kovac, C. R., Sepulveda, M. A., & Birshtein, B. K. (2002). The interaction of Pax5 (BSAP) with Daxx can result in transcriptional activation in B cells. *The Journal of Biological Chemistry*, 277(13), 11156-11164. <https://doi.org/10.1074/jbc.M111763200>
- Epstein, D. J., Malo, D., Vekemans, M., & Gros, P. (1991). Molecular characterization of a deletion encompassing the splotch mutation on mouse chromosome 1. *Genomics*, 10(1), 89-93. [https://doi.org/10.1016/0888-7543\(91\)90488-z](https://doi.org/10.1016/0888-7543(91)90488-z)
- Epstein, J., Cai, J., Glaser, T., Jepeal, L., & Maas, R. (1994). Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes. *The Journal of Biological Chemistry*, 269(11), 8355-8361.
- Fairfax, K. A., Kallies, A., Nutt, S. L., & Tarlinton, D. M. (2008). Plasma cell development : From B-cell subsets to long-term survival niches. *Seminars in Immunology*, 20(1), 49-58. <https://doi.org/10.1016/j.smim.2007.12.002>
- Familiades, J., Bousquet, M., Lafage-Pochitaloff, M., Béné, M.-C., Beldjord, K., De Vos, J., Dastugue, N., Coyaud, E., Struski, S., Quelen, C., Prade-Houdellier, N., Dobbelsstein, S., Cayuela, J.-M., Soulier, J., Grardel, N., Preudhomme, C., Cavé, H., Blanchet, O., Lhéritier, V., ... Delabesse, E. (2009). PAX5 mutations occur frequently in adult B-cell progenitor acute lymphoblastic leukemia and PAX5 haploinsufficiency is associated with BCR-ABL1 and TCF3-PBX1 fusion genes : A GRAALL study. *Leukemia*, 23(11), 1989-1998. <https://doi.org/10.1038/leu.2009.135>
- Favor, J., Sandulache, R., Neuhäuser-Klaus, A., Pretsch, W., Chatterjee, B., Senft, E., Wurst, W., Blanquet, V., Grimes, P., Spörle, R., & Schughart, K. (1996). The mouse Pax2(1Neu) mutation is identical to a human PAX2 mutation in a family with renal-coloboma syndrome and results in developmental defects of the brain, ear, eye, and kidney. *Proceedings of the National Academy of Sciences of the United States of America*, 93(24), 13870-13875. <https://doi.org/10.1073/pnas.93.24.13870>
- Fazio, G., Cazzaniga, V., Palmi, C., Galbiati, M., Giordan, M., te Kronnie, G., Rolink, A., Biondi, A., & Cazzaniga, G. (2013). PAX5/ETV6 alters the gene expression profile of precursor B cells with opposite dominant effect on endogenous PAX5. *Leukemia*, 27(4), Article 4. <https://doi.org/10.1038/leu.2012.281>
- Fazio, G., Palmi, C., Rolink, A., Biondi, A., & Cazzaniga, G. (2008). PAX5/TEL acts as a transcriptional repressor causing down-modulation of CD19, enhances migration to

- CXCL12, and confers survival advantage in pre-B1 cells. *Cancer Research*, 68(1), 181-189. <https://doi.org/10.1158/0008-5472.CAN-07-2778>
- Feichtinger, R. G., Mucha, B. E., Hengel, H., Orfi, Z., Makowski, C., Dort, J., D'Anjou, G., Nguyen, T. T. M., Buchert, R., Juenger, H., Freisinger, P., Baumeister, S., Schoser, B., Ahting, U., Keimer, R., Nguyen, C.-T. E., Fabre, P., Gauthier, J., Miguet, M., ... Dumont, N. A. (2019). Biallelic variants in the transcription factor PAX7 are a new genetic cause of myopathy. *Genetics in Medicine*, 21(11), Article 11. <https://doi.org/10.1038/s41436-019-0532-z>
- Fielding, A. K. (2015). Treatment of Philadelphia Chromosome–Positive Acute Lymphoblastic Leukemia in Adults : A Broader Range of Options, Improved Outcomes, and More Therapeutic Dilemmas. *American Society of Clinical Oncology Educational Book*, 35, e352-e359. [https://doi.org/10.14694/EdBook\\_AM.2015.35.e352](https://doi.org/10.14694/EdBook_AM.2015.35.e352)
- Fleming, H. E., & Paige, C. J. (2001). Pre-B cell receptor signaling mediates selective response to IL-7 at the pro-B to pre-B cell transition via an ERK/MAP kinase-dependent pathway. *Immunity*, 15(4), 521-531. [https://doi.org/10.1016/s1074-7613\(01\)00216-3](https://doi.org/10.1016/s1074-7613(01)00216-3)
- Forman, S. J., & Rowe, J. M. (2013). The myth of the second remission of acute leukemia in the adult. *Blood*, 121(7), 1077-1082. <https://doi.org/10.1182/blood-2012-08-234492>
- Fortschegger, K., Anderl, S., Denk, D., & Strehl, S. (2014). Functional heterogeneity of PAX5 chimeras reveals insight for leukemia development. *Molecular Cancer Research: MCR*, 12(4), 595-606. <https://doi.org/10.1158/1541-7786.MCR-13-0337>
- Francioso, F., Carinci, F., Tosi, L., Scapoli, L., Pezzetti, F., Passerella, E., Evangelisti, R., Pastore, A., Pelucchi, S., Piattelli, A., Rubini, C., Fioroni, M., Carinci, P., & Volinia, S. (2002). Identification of differentially expressed genes in human salivary gland tumors by DNA microarrays. *Molecular Cancer Therapeutics*, 1(7), 533-538.
- Francis, S. S., Selvin, S., Yang, W., Buffler, P. A., & Wiemels, J. L. (2012). Unusual space-time patterning of the Fallon, Nevada leukemia cluster: Evidence of an infectious etiology. *Chemico-Biological Interactions*, 196(3), 102-109. <https://doi.org/10.1016/j.cbi.2011.02.019>
- Fregona, V., Bayet, M., & Gerby, B. (2021). Oncogene-Induced Reprogramming in Acute Lymphoblastic Leukemia : Towards Targeted Therapy of Leukemia-Initiating Cells. *Cancers*, 13(21), 5511. <https://doi.org/10.3390/cancers13215511>
- Frey, N. V., & Luger, S. M. (2015). How I treat adults with relapsed or refractory Philadelphia chromosome–negative acute lymphoblastic leukemia. *Blood*, 126(5), 589-596. <https://doi.org/10.1182/blood-2014-09-551937>

- Frigerio, G., Burri, M., Bopp, D., Baumgartner, S., & Noll, M. (1986). Structure of the segmentation gene paired and the Drosophila PRD gene set as part of a gene network. *Cell*, 47(5), 735-746. [https://doi.org/10.1016/0092-8674\(86\)90516-7](https://doi.org/10.1016/0092-8674(86)90516-7)
- Fujitani, Y., Kajimoto, Y., Yasuda, T., Matsuoka, T. A., Kaneto, H., Umayahara, Y., Fujita, N., Watada, H., Miyazaki, J. I., Yamasaki, Y., & Hori, M. (1999). Identification of a portable repression domain and an E1A-responsive activation domain in Pax4 : A possible role of Pax4 as a transcriptional repressor in the pancreas. *Molecular and Cellular Biology*, 19(12), 8281-8291. <https://doi.org/10.1128/mcb.19.12.8281>
- Fulbright, J. M., Raman, S., McClellan, W. S., & August, K. J. (2011). Late Effects of Childhood Leukemia Therapy. *Current Hematologic Malignancy Reports*, 6(3), 195-205. <https://doi.org/10.1007/s11899-011-0094-x>
- Furukawa, T., Kozak, C. A., & Cepko, C. L. (1997). Rax, a novel paired-type homeobox gene, shows expression in the anterior neural fold and developing retina. *Proceedings of the National Academy of Sciences of the United States of America*, 94(7), 3088-3093. <https://doi.org/10.1073/pnas.94.7.3088>
- Fuxa, M., & Busslinger, M. (2007). Reporter Gene Insertions Reveal a Strictly B Lymphoid-Specific Expression Pattern of Pax5 in Support of Its B Cell Identity Function. *The Journal of Immunology*, 178(5), 3031-3037. <https://doi.org/10.4049/jimmunol.178.5.3031>
- Fuxa, M., Skok, J., Souabni, A., Salvagiotto, G., Roldan, E., & Busslinger, M. (2004). Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes & Development*, 18(4), 411-422. <https://doi.org/10.1101/gad.291504>
- Gauthier, L., Rossi, B., Roux, F., Termine, E., & Schiff, C. (2002). Galectin-1 is a stromal cell ligand of the pre-B cell receptor (BCR) implicated in synapse formation between pre-B and stromal cells and in pre-BCR triggering. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20), 13014-13019. <https://doi.org/10.1073/pnas.202323999>
- Gehring, W. J., Affolter, M., & Bürglin, T. (1994). Homeodomain proteins. *Annual Review of Biochemistry*, 63, 487-526. <https://doi.org/10.1146/annurev.bi.63.070194.002415>
- Gehring, W. J., Müller, M., Affolter, M., Percival-Smith, A., Billeter, M., Qian, Y. Q., Otting, G., & Wüthrich, K. (1990). The structure of the homeodomain and its functional implications. *Trends in Genetics: TIG*, 6(10), 323-329. [https://doi.org/10.1016/0168-9525\(90\)90253-3](https://doi.org/10.1016/0168-9525(90)90253-3)

- Geier, J. K., & Schlissel, M. S. (2006). Pre-BCR signals and the control of Ig gene rearrangements. *Seminars in Immunology*, 18(1), 31-39. <https://doi.org/10.1016/j.smim.2005.11.001>
- Gellert, M. (2002). V(D)J recombination : RAG proteins, repair factors, and regulation. *Annual Review of Biochemistry*, 71, 101-132. <https://doi.org/10.1146/annurev.biochem.71.090501.150203>
- Gerber, J.-K., Richter, T., Kremmer, E., Adamski, J., Höfler, H., Balling, R., & Peters, H. (2002). Progressive loss of PAX9 expression correlates with increasing malignancy of dysplastic and cancerous epithelium of the human oesophagus. *The Journal of Pathology*, 197(3), 293-297. <https://doi.org/10.1002/path.1115>
- Gibson, W., Green, A., Bullard, R. S., Eaddy, A. C., & Donald, C. D. (2007). Inhibition of PAX2 expression results in alternate cell death pathways in prostate cancer cells differing in p53 status. *Cancer Letters*, 248(2), 251-261. <https://doi.org/10.1016/j.canlet.2006.08.007>
- Gilham, C., Peto, J., Simpson, J., Roman, E., Eden, T. O. B., Greaves, M. F., Alexander, F. E., & UKCCS Investigators. (2005). Day care in infancy and risk of childhood acute lymphoblastic leukaemia : Findings from UK case-control study. *BMJ (Clinical Research Ed.)*, 330(7503), 1294. <https://doi.org/10.1136/bmj.38428.521042.8F>
- Glaser, T., Jepeal, L., Edwards, J. G., Young, S. R., Favor, J., & Maas, R. L. (1994). PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. *Nature Genetics*, 7(4), 463-471. <https://doi.org/10.1038/ng0894-463>
- Gnarra, J. R., & Dressler, G. R. (1995). Expression of Pax-2 in human renal cell carcinoma and growth inhibition by antisense oligonucleotides. *Cancer Research*, 55(18), 4092-4098.
- Goebel, P., Janney, N., Valenzuela, J. R., Romanow, W. J., Murre, C., & Feeney, A. J. (2001). Localized gene-specific induction of accessibility to V(D)J recombination induced by E2A and early B cell factor in nonlymphoid cells. *The Journal of Experimental Medicine*, 194(5), 645-656. <https://doi.org/10.1084/jem.194.5.645>
- Goetz, C. A., Harmon, I. R., O'Neil, J. J., Burchill, M. A., & Farrar, M. A. (2004). STAT5 activation underlies IL7 receptor-dependent B cell development. *Journal of Immunology (Baltimore, Md.: 1950)*, 172(8), 4770-4778. <https://doi.org/10.4049/jimmunol.172.8.4770>
- Goode, D. K., & Elgar, G. (2009). The PAX258 gene subfamily : A comparative perspective. *Developmental Dynamics: An Official Publication of the American Association of Anatomists*, 238(12), 2951-2974. <https://doi.org/10.1002/dvdy.22146>

- Gosmain, Y., Cheyssac, C., Masson, M. H., Guérardel, A., Poisson, C., & Philippe, J. (2012). Pax6 is a key component of regulated glucagon secretion. *Endocrinology*, *153*(9), 4204-4215. <https://doi.org/10.1210/en.2012-1425>
- Gosmain, Y., Katz, L. S., Masson, M. H., Cheyssac, C., Poisson, C., & Philippe, J. (2012). Pax6 Is Crucial for  $\beta$ -Cell Function, Insulin Biosynthesis, and Glucose-Induced Insulin Secretion. *Molecular Endocrinology*, *26*(4), 696-709. <https://doi.org/10.1210/me.2011-1256>
- Grawunder, U., Leu, T. M., Schatz, D. G., Werner, A., Rolink, A. G., Melchers, F., & Winkler, T. H. (1995). Down-regulation of RAG1 and RAG2 gene expression in preB cells after functional immunoglobulin heavy chain rearrangement. *Immunity*, *3*(5), 601-608. [https://doi.org/10.1016/1074-7613\(95\)90131-0](https://doi.org/10.1016/1074-7613(95)90131-0)
- Greaves, M. (2018). A causal mechanism for childhood acute lymphoblastic leukaemia. *Nature Reviews. Cancer*, *18*(8), 471-484. <https://doi.org/10.1038/s41568-018-0015-6>
- Gregory Powell, J., Wang, X., Allard, B. L., Sahin, M., Wang, X.-L., Hay, I. D., Hiddinga, H. J., Deshpande, S. S., Kroll, T. G., Grebe, S. K. G., Eberhardt, N. L., & Mclver, B. (2004). The PAX8/PPARgamma fusion oncoprotein transforms immortalized human thyrocytes through a mechanism probably involving wild-type PPARgamma inhibition. *Oncogene*, *23*(20), 3634-3641. <https://doi.org/10.1038/sj.onc.1207399>
- Grootens, J., Ungerstedt, J. S., Ekoff, M., Rönnerberg, E., Klimkowska, M., Amini, R.-M., Arock, M., Söderlund, S., Mattsson, M., Nilsson, G., & Dahlin, J. S. (2019). Single-cell analysis reveals the KIT D816V mutation in haematopoietic stem and progenitor cells in systemic mastocytosis. *EBioMedicine*, *43*, 150-158. <https://doi.org/10.1016/j.ebiom.2019.03.089>
- Grupp, S. A., Kalos, M., Barrett, D., Aplenc, R., Porter, D. L., Rheingold, S. R., Teachey, D. T., Chew, A., Hauck, B., Wright, J. F., Milone, M. C., Levine, B. L., & June, C. H. (2013). Chimeric antigen receptor-modified T cells for acute lymphoid leukemia. *The New England Journal of Medicine*, *368*(16), 1509-1518. <https://doi.org/10.1056/NEJMoa1215134>
- Gu, Z., Churchman, M. L., Roberts, K. G., Moore, I., Zhou, X., Nakitandwe, J., Hagiwara, K., Pelletier, S., Gingras, S., Berns, H., Payne-Turner, D., Hill, A., Iacobucci, I., Shi, L., Pounds, S., Cheng, C., Pei, D., Qu, C., Newman, S., ... Mullighan, C. G. (2019). PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nature Genetics*, *51*(2), 296-307. <https://doi.org/10.1038/s41588-018-0315-5>
- Gu, Z., Churchman, M., Roberts, K., Li, Y., Liu, Y., Harvey, R. C., McCastlain, K., Reshmi, S. C., Payne-Turner, D., Iacobucci, I., Shao, Y., Chen, I.-M., Valentine, M., Pei, D., Mungall, K. L., Mungall, A. J., Ma, Y., Moore, R., Marra, M., ... Mullighan, C. G. (2016).

- Genomic analyses identify recurrent MEF2D fusions in acute lymphoblastic leukaemia. *Nature Communications*, 7, 13331. <https://doi.org/10.1038/ncomms13331>
- Hagman, J., & Lukin, K. (2006). Transcription factors drive B cell development. *Current Opinion in Immunology*, 18(2), 127-134. <https://doi.org/10.1016/j.coi.2006.01.007>
- Hardy, R. R., Carmack, C. E., Shinton, S. A., Kemp, J. D., & Hayakawa, K. (1991). Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *The Journal of Experimental Medicine*, 173(5), 1213-1225. <https://doi.org/10.1084/jem.173.5.1213>
- Hardy, R. R., & Hayakawa, K. (2001). B cell development pathways. *Annual Review of Immunology*, 19, 595-621. <https://doi.org/10.1146/annurev.immunol.19.1.595>
- Hardy, R. R., Kincade, P. W., & Dorshkind, K. (2007). The protean nature of cells in the B lymphocyte lineage. *Immunity*, 26(6), 703-714. <https://doi.org/10.1016/j.immuni.2007.05.013>
- Harrison, C. J., & Schwab, C. (2016). Constitutional abnormalities of chromosome 21 predispose to iAMP21-acute lymphoblastic leukaemia. *European Journal of Medical Genetics*, 59(3), 162-165. <https://doi.org/10.1016/j.ejmg.2016.01.006>
- Hauser, J., Grundström, C., Kumar, R., & Grundström, T. (2016). Regulated localization of an AID complex with E2A, PAX5 and IRF4 at the Igh locus. *Molecular Immunology*, 80, 78-90. <https://doi.org/10.1016/j.molimm.2016.10.014>
- He, T., Hong, S. Y., Huang, L., Xue, W., Yu, Z., Kwon, H., Kirk, M., Ding, S., Su, K., & Zhang, Z. (2011). Histone acetyltransferase p300 acetylates Pax5 and strongly enhances Pax5-mediated transcriptional activity. *The Journal of Biological Chemistry*, 286(16), 14137-14145. <https://doi.org/10.1074/jbc.M110.176289>
- Heath, C. W., & Hasterlik, R. J. (1990). Leukemia among children in a suburban community. 1963. *CA: A Cancer Journal for Clinicians*, 40(1), 27-50. <https://doi.org/10.3322/canjclin.40.1.29>
- Hein, D., Dreisig, K., Metzler, M., Izraeli, S., Schmiegelow, K., Borkhardt, A., & Fischer, U. (2019). The preleukemic TCF3-PBX1 gene fusion can be generated in utero and is present in ≈0.6% of healthy newborns. *Blood*, 134(16), 1355-1358. <https://doi.org/10.1182/blood.2019002215>
- Heltemes-Harris, L. M., Willette, M. J. L., Ramsey, L. B., Qiu, Y. H., Neeley, E. S., Zhang, N., Thomas, D. A., Koeuth, T., Baechler, E. C., Kornblau, S. M., & Farrar, M. A. (2011). Ebf1 or Pax5 haploinsufficiency synergizes with STAT5 activation to initiate acute lymphoblastic leukemia. *The Journal of Experimental Medicine*, 208(6), 1135-1149. <https://doi.org/10.1084/jem.20101947>

- Hennighausen, L., & Robinson, G. W. (2008). Interpretation of cytokine signaling through the transcription factors STAT5A and STAT5B. *Genes & Development*, 22(6), 711-721. <https://doi.org/10.1101/gad.1643908>
- Herglotz, J., Unrau, L., Hauschildt, F., Fischer, M., Kriebitzsch, N., Alawi, M., Indenbirken, D., Spohn, M., Müller, U., Ziegler, M., Schuh, W., Jäck, H.-M., & Stocking, C. (2016). Essential control of early B-cell development by Mef2 transcription factors. *Blood*, 127(5), 572-581. <https://doi.org/10.1182/blood-2015-04-643270>
- Hermans, A., Heisterkamp, N., von Linden, M., van Baal, S., Meijer, D., van der Plas, D., Wiedemann, L. M., Groffen, J., Bootsma, D., & Grosveld, G. (1987). Unique fusion of bcr and c-abl genes in Philadelphia chromosome positive acute lymphoblastic leukemia. *Cell*, 51(1), 33-40. [https://doi.org/10.1016/0092-8674\(87\)90007-9](https://doi.org/10.1016/0092-8674(87)90007-9)
- Herzog, S., Reth, M., & Jumaa, H. (2009). Regulation of B-cell proliferation and differentiation by pre-B-cell receptor signalling. *Nature Reviews. Immunology*, 9(3), 195-205. <https://doi.org/10.1038/nri2491>
- Hesslein, D. G., & Schatz, D. G. (2001). Factors and forces controlling V(D)J recombination. *Advances in Immunology*, 78, 169-232. [https://doi.org/10.1016/s0065-2776\(01\)78004-2](https://doi.org/10.1016/s0065-2776(01)78004-2)
- Hill, R. E., Favor, J., Hogan, B. L., Ton, C. C., Saunders, G. F., Hanson, I. M., Prosser, J., Jordan, T., Hastie, N. D., & van Heyningen, V. (1991). Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature*, 354(6354), 522-525. <https://doi.org/10.1038/354522a0>
- Hirokawa, S., Sato, H., Kato, I., & Kudo, A. (2003). EBF-regulating Pax5 transcription is enhanced by STAT5 in the early stage of B cells. *European Journal of Immunology*, 33(7), 1824-1829. <https://doi.org/10.1002/eji.200323974>
- Hofmeister, R., Khaled, A. R., Benbernou, N., Rajnavolgyi, E., Muegge, K., & Durum, S. K. (1999). Interleukin-7: Physiological roles and mechanisms of action. *Cytokine & Growth Factor Reviews*, 10(1), 41-60. [https://doi.org/10.1016/s1359-6101\(98\)00025-2](https://doi.org/10.1016/s1359-6101(98)00025-2)
- Hollenbach, A. D., Sublett, J. E., McPherson, C. J., & Grosveld, G. (1999). The Pax3-FKHR oncoprotein is unresponsive to the Pax3-associated repressor hDaxx. *The EMBO Journal*, 18(13), 3702-3711. <https://doi.org/10.1093/emboj/18.13.3702>
- Holmes, M. L., Carotta, S., Corcoran, L. M., & Nutt, S. L. (2006). Repression of Flt3 by Pax5 is crucial for B-cell lineage commitment. *Genes & Development*, 20(8), 933-938. <https://doi.org/10.1101/gad.1396206>
- Holmfeldt, L., Wei, L., Diaz-Flores, E., Walsh, M., Zhang, J., Ding, L., Payne-Turner, D., Churchman, M., Andersson, A., Chen, S.-C., McCastlain, K., Becksfort, J., Ma, J., Wu,

- G., Patel, S. N., Heatley, S. L., Phillips, L. A., Song, G., Easton, J., ... Mullighan, C. G. (2013). The genomic landscape of hypodiploid acute lymphoblastic leukemia. *Nature Genetics*, *45*(3), 242-252. <https://doi.org/10.1038/ng.2532>
- Hong, D., Gupta, R., Ancliff, P., Atzberger, A., Brown, J., Soneji, S., Green, J., Colman, S., Piacibello, W., Buckle, V., Tsuzuki, S., Greaves, M., & Enver, T. (2008). Initiating and Cancer-Propagating Cells in TEL-AML1-Associated Childhood Leukemia. *Science*, *319*(5861), 336-339. <https://doi.org/10.1126/science.1150648>
- Hong, E. E., Erickson, H., Lutz, R. J., Whiteman, K. R., Jones, G., Kovtun, Y., Blanc, V., & Lambert, J. M. (2015). Design of Coltuximab Ravtansine, a CD19-Targeting Antibody-Drug Conjugate (ADC) for the Treatment of B-Cell Malignancies : Structure-Activity Relationships and Preclinical Evaluation. *Molecular Pharmaceutics*, *12*(6), 1703-1716. <https://doi.org/10.1021/acs.molpharmaceut.5b00175>
- Hormann, F. M., Hoogkamer, A. Q., Beverloo, H. B., Boeree, A., Dingjan, I., Wattel, M. M., Stam, R. W., Escherich, G., Pieters, R., den Boer, M. L., & Boer, J. M. (2019). NUTM1 is a recurrent fusion gene partner in B-cell precursor acute lymphoblastic leukemia associated with increased expression of genes on chromosome band 10p12.31-12.2. *Haematologica*, *104*(10), e455-e459. <https://doi.org/10.3324/haematol.2018.206961>
- Hoth, C. F., Milunsky, A., Lipsky, N., Sheffer, R., Clarren, S. K., & Baldwin, C. T. (1993). Mutations in the paired domain of the human PAX3 gene cause Klein-Waardenburg syndrome (WS-III) as well as Waardenburg syndrome type I (WS-I). *American Journal of Human Genetics*, *52*(3), 455-462.
- Huang, C., Geng, H., Boss, I., Wang, L., & Melnick, A. (2014). Cooperative transcriptional repression by BCL6 and BACH2 in germinal center B-cell differentiation. *Blood*, *123*(7), 1012-1020. <https://doi.org/10.1182/blood-2013-07-518605>
- Huettner, C. S., Koschmieder, S., Iwasaki, H., Iwasaki-Arai, J., Radomska, H. S., Akashi, K., & Tenen, D. G. (2003). Inducible expression of BCR/ABL using human CD34 regulatory elements results in a megakaryocytic myeloproliferative syndrome. *Blood*, *102*(9), 3363-3370. <https://doi.org/10.1182/blood-2003-03-0768>
- Huettner, C. S., Zhang, P., Van Etten, R. A., & Tenen, D. G. (2000). Reversibility of acute B-cell leukaemia induced by BCR-ABL1. *Nature Genetics*, *24*(1), 57-60. <https://doi.org/10.1038/71691>
- Hunger, S. P., Lu, X., Devidas, M., Camitta, B. M., Gaynon, P. S., Winick, N. J., Reaman, G. H., & Carroll, W. L. (2012). Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005 : A report from the children's oncology group. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *30*(14), 1663-1669. <https://doi.org/10.1200/JCO.2011.37.8018>



- Hunger, S. P., & Mullighan, C. G. (2015). Acute Lymphoblastic Leukemia in Children. *The New England Journal of Medicine*, 373(16), 1541-1552. <https://doi.org/10.1056/NEJMra1400972>
- Huret, J.-L., Ahmad, M., Arsaban, M., Bernheim, A., Cigna, J., Desangles, F., Guignard, J.-C., Jacquemot-Perbal, M.-C., Labarussias, M., Leberre, V., Malo, A., Morel-Pair, C., Mossafa, H., Potier, J.-C., Texier, G., Viguié, F., Yau Chun Wan-Senon, S., Zasadzinski, A., & Dessen, P. (2013). Atlas of genetics and cytogenetics in oncology and haematology in 2013. *Nucleic Acids Research*, 41(Database issue), D920-924. <https://doi.org/10.1093/nar/gks1082>
- Iacobucci, I., Lonetti, A., Paoloni, F., Papayannidis, C., Ferrari, A., Storlazzi, C. T., Vignetti, M., Cilloni, D., Messa, F., Guadagnuolo, V., Paolini, S., Elia, L., Messina, M., Vitale, A., Meloni, G., Soverini, S., Pane, F., Baccarani, M., Foà, R., & Martinelli, G. (2010). The PAX5 gene is frequently rearranged in BCR-ABL1-positive acute lymphoblastic leukemia but is not associated with outcome. A report on behalf of the GIMEMA Acute Leukemia Working Party. *Haematologica*, 95(10), 1683-1690. <https://doi.org/10.3324/haematol.2009.020792>
- Iacobucci, I., & Mullighan, C. G. (2017). Genetic Basis of Acute Lymphoblastic Leukemia. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 35(9), 975-983. <https://doi.org/10.1200/JCO.2016.70.7836>
- Igarashi, H., Gregory, S. C., Yokota, T., Sakaguchi, N., & Kincade, P. W. (2002). Transcription from the RAG1 locus marks the earliest lymphocyte progenitors in bone marrow. *Immunity*, 17(2), 117-130. [https://doi.org/10.1016/s1074-7613\(02\)00366-7](https://doi.org/10.1016/s1074-7613(02)00366-7)
- Igarashi, K., Ochiai, K., Itoh-Nakadai, A., & Muto, A. (2014). Orchestration of plasma cell differentiation by Bach2 and its gene regulatory network. *Immunological Reviews*, 261(1), 116-125. <https://doi.org/10.1111/imr.12201>
- Iida, S., Rao, P. H., Nallasivam, P., Hibshoosh, H., Butler, M., Louie, D. C., Dyomin, V., Ohno, H., Chaganti, R. S., & Dalla-Favera, R. (1996). The t(9;14)(p13;q32) chromosomal translocation associated with lymphoplasmacytoid lymphoma involves the PAX-5 gene. *Blood*, 88(11), 4110-4117.
- Imamura, Y., Oda, A., Katahira, T., Bundo, K., Pike, K. A., Ratcliffe, M. J. H., & Kitamura, D. (2009). BLNK binds active H-Ras to promote B cell receptor-mediated capping and ERK activation. *The Journal of Biological Chemistry*, 284(15), 9804-9813. <https://doi.org/10.1074/jbc.M809051200>
- Inaba, H., Greaves, M., & Mullighan, C. G. (2013). Acute lymphoblastic leukaemia. *Lancet (London, England)*, 381(9881), 1943-1955. [https://doi.org/10.1016/S0140-6736\(12\)62187-4](https://doi.org/10.1016/S0140-6736(12)62187-4)

- Inagaki, Y., Hayakawa, F., Hirano, D., Kojima, Y., Morishita, T., Yasuda, T., Naoe, T., & Kiyoi, H. (2016). PAX5 tyrosine phosphorylation by SYK co-operatively functions with its serine phosphorylation to cancel the PAX5-dependent repression of BLIMP1: A mechanism for antigen-triggered plasma cell differentiation. *Biochemical and Biophysical Research Communications*, *475*(2), 176-181. <https://doi.org/10.1016/j.bbrc.2016.05.067>
- Inlay, M. A., Bhattacharya, D., Sahoo, D., Serwold, T., Seita, J., Karsunky, H., Plevritis, S. K., Dill, D. L., & Weissman, I. L. (2009). Ly6d marks the earliest stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. *Genes & Development*, *23*(20), 2376-2381. <https://doi.org/10.1101/gad.1836009>
- Iritani, B. M., Forbush, K. A., Farrar, M. A., & Perlmutter, R. M. (1997). Control of B cell development by Ras-mediated activation of Raf. *The EMBO Journal*, *16*(23), 7019-7031. <https://doi.org/10.1093/emboj/16.23.7019>
- Itoh-Nakadai, A., Hikota, R., Muto, A., Kometani, K., Watanabe-Matsui, M., Sato, Y., Kobayashi, M., Nakamura, A., Miura, Y., Yano, Y., Tashiro, S., Sun, J., Ikawa, T., Ochiai, K., Kurosaki, T., & Igarashi, K. (2014). The transcription repressors Bach2 and Bach1 promote B cell development by repressing the myeloid program. *Nature Immunology*, *15*(12), 1171-1180. <https://doi.org/10.1038/ni.3024>
- Jabbour, E., O'Brien, S., Konopleva, M., & Kantarjian, H. (2015). New insights into the pathophysiology and therapy of adult acute lymphoblastic leukemia. *Cancer*, *121*(15), 2517-2528. <https://doi.org/10.1002/cncr.29383>
- Jacobsen, J. A., Bartom, E. T., Sigvardsson, M., & Kee, B. L. (2020). Ezh2 Represses Transcription of Innate Lymphoid Genes in B Lymphocyte Progenitors and Maintains the B-2 Cell Fate. *The Journal of Immunology*, *204*(7), 1760-1769. <https://doi.org/10.4049/jimmunol.1901188>
- Jamrog, L., Chemin, G., Fregona, V., Coster, L., Pasquet, M., Oudinet, C., Rouquié, N., Prade, N., Lagarde, S., Cresson, C., Hébrard, S., Nguyen Huu, N. S., Bousquet, M., Quelen, C., Brousset, P., Mancini, S. J. C., Delabesse, E., Khamlichi, A. A., Gerby, B., & Broccardo, C. (2018). PAX5-ELN oncoprotein promotes multistep B-cell acute lymphoblastic leukemia in mice. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(41), 10357-10362. <https://doi.org/10.1073/pnas.1721678115>
- Jennings, B. H., & Ish-Horowicz, D. (2008). The Groucho/TLE/Grg family of transcriptional co-repressors. *Genome Biology*, *9*(1), 205. <https://doi.org/10.1186/gb-2008-9-1-205>
- Jennings, B. H., Pickles, L. M., Wainwright, S. M., Roe, S. M., Pearl, L. H., & Ish-Horowicz, D. (2006). Molecular Recognition of Transcriptional Repressor Motifs by the WD Domain

- of the Groucho/TLE Corepressor. *Molecular Cell*, 22(5), 645-655. <https://doi.org/10.1016/j.molcel.2006.04.024>
- Jensen, C. T., Åhsberg, J., Sommarin, M. N. E., Strid, T., Somasundaram, R., Okuyama, K., Ungerbäck, J., Kupari, J., Airaksinen, M. S., Lang, S., Bryder, D., Soneji, S., Karlsson, G., & Sigvardsson, M. (2018). Dissection of progenitor compartments resolves developmental trajectories in B-lymphopoiesis. *The Journal of Experimental Medicine*, 215(7), 1947-1963. <https://doi.org/10.1084/jem.20171384>
- Jensen, C. T., Lang, S., Somasundaram, R., Soneji, S., & Sigvardsson, M. (2016). Identification of Stage-Specific Surface Markers in Early B Cell Development Provides Novel Tools for Identification of Progenitor Populations. *Journal of Immunology (Baltimore, Md.: 1950)*, 197(5), 1937-1944. <https://doi.org/10.4049/jimmunol.1600297>
- Jordan, T., Hanson, I., Zaletayev, D., Hodgson, S., Prosser, J., Seawright, A., Hastie, N., & van Heyningen, V. (1992). The human PAX6 gene is mutated in two patients with aniridia. *Nature Genetics*, 1(5), 328-332. <https://doi.org/10.1038/ng0892-328>
- Jun, S., & Desplan, C. (1996). Cooperative interactions between paired domain and homeodomain. *Development (Cambridge, England)*, 122(9), 2639-2650.
- Jurado, S., Kostanova, D., & Busslinger, M. (2015). Rapid formation of B cell leukemia in mice expressing the PAX5-JAK2 fusion protein. *Experimental Hematology*, 43(9), S45. <https://doi.org/10.1016/j.exphem.2015.06.049>
- Kallies, A., Hasbold, J., Fairfax, K., Pridans, C., Emslie, D., McKenzie, B. S., Lew, A. M., Corcoran, L. M., Hodgkin, P. D., Tarlinton, D. M., & Nutt, S. L. (2007). Initiation of Plasma-Cell Differentiation Is Independent of the Transcription Factor Blimp-1. *Immunity*, 26(5), 555-566. <https://doi.org/10.1016/j.immuni.2007.04.007>
- Kallies, A., & Nutt, S. L. (2007). Terminal differentiation of lymphocytes depends on Blimp-1. *Current Opinion in Immunology*, 19(2), 156-162. <https://doi.org/10.1016/j.coi.2007.01.003>
- Kanteti, R., Nallasura, V., Loganathan, S., Tretiakova, M., Kroll, T., Krishnaswamy, S., Faoro, L., Cagle, P., Husain, A. N., Vokes, E. E., Lang, D., & Salgia, R. (2009). PAX5 is expressed in small-cell lung cancer and positively regulates c-Met transcription. *Laboratory Investigation; a Journal of Technical Methods and Pathology*, 89(3), 301-314. <https://doi.org/10.1038/labinvest.2008.168>
- Kantner, H.-P., Warsch, W., Delogu, A., Bauer, E., Esterbauer, H., Casanova, E., Sexl, V., & Stoiber, D. (2013). ETV6/RUNX1 induces reactive oxygen species and drives the accumulation of DNA damage in B cells. *Neoplasia (New York, N.Y.)*, 15(11), 1292-1300. <https://doi.org/10.1593/neo.131310>

- Kawamata, N., Ogawa, S., Zimmermann, M., Niebuhr, B., Stocking, C., Sanada, M., Hemminki, K., Yamamoto, G., Nannya, Y., Koehler, R., Flohr, T., Miller, C. W., Harbott, J., Ludwig, W.-D., Stanulla, M., Schrappe, M., Bartram, C. R., & Koeffler, H. P. (2008). Cloning of genes involved in chromosomal translocations by high-resolution single nucleotide polymorphism genomic microarray. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(33), 11921-11926. <https://doi.org/10.1073/pnas.0711039105>
- Kawamata, N., Pennella, M. A., Woo, J. L., Berk, A. J., & Koeffler, H. P. (2012). Dominant-negative mechanism of leukemogenic PAX5 fusions. *Oncogene*, *31*(8), 966-977. <https://doi.org/10.1038/onc.2011.291>
- Kawamata, N., Pennella, M., Woo, J., Berk, A., & Koeffler, H. P. (2010). Dominant Negative Form of PAX5 Is Generated by Multimerization of Its DNA Binding Domain. *Blood*, *116*(21), 25-25. <https://doi.org/10.1182/blood.V116.21.25.25>
- Kikuchi, H., Nakayama, M., Kuribayashi, F., Mimuro, H., Imajoh-Ohmi, S., Nishitoh, H., & Nakayama, T. (2015). Histone acetyltransferase PCAF is involved in transactivation of Bcl-6 and Pax5 genes in immature B cells. *Biochemical and Biophysical Research Communications*, *467*(3), 509-513. <https://doi.org/10.1016/j.bbrc.2015.10.011>
- Kikuchi, H., Nakayama, M., Kuribayashi, F., Mimuro, H., Imajoh-Ohmi, S., Nishitoh, H., Takami, Y., & Nakayama, T. (2015). Paired box gene 5 isoforms A and B have different functions in transcriptional regulation of B cell development-related genes in immature B cells. *Microbiology and Immunology*, *59*(7), 426-431. <https://doi.org/10.1111/1348-0421.12272>
- Kikuchi, H., Nakayama, M., Takami, Y., Kuribayashi, F., & Nakayama, T. (2011). Possible involvement of Helios in controlling the immature B cell functions via transcriptional regulation of protein kinase Cs. *Results in Immunology*, *1*(1), 88-94. <https://doi.org/10.1016/j.rinim.2011.11.002>
- Kikuchi, H., Nakayama, M., Takami, Y., Kuribayashi, F., & Nakayama, T. (2012). EBF1 acts as a powerful repressor of Blimp-1 gene expression in immature B cells. *Biochemical and Biophysical Research Communications*, *422*(4), 780-785. <https://doi.org/10.1016/j.bbrc.2012.05.099>
- Kikuchi, H., Yamashita, K., Nakayama, M., Toyonaga, K., Tsuneyoshi, I., Takasaki, M., & Nakayama, T. (2009). Lacking of Aiolos accelerates pre-mature B cell apoptosis mediated by BCR signaling through elevation in cytochrome c release. *Biochimica Et Biophysica Acta*, *1793*(7), 1304-1314. <https://doi.org/10.1016/j.bbamcr.2009.04.016>
- Kim, M., Choi, J. E., She, C. J., Hwang, S. M., Shin, H. Y., Ahn, H. S., Yoon, S.-S., Kim, B. K., Park, M. H., & Lee, D. S. (2011). PAX5 deletion is common and concurrently occurs

- with CDKN2A deletion in B-lineage acute lymphoblastic leukemia. *Blood Cells, Molecules & Diseases*, 47(1), 62-66. <https://doi.org/10.1016/j.bcmed.2011.04.003>
- Kimura, S., & Mullighan, C. G. (2020). Molecular markers in ALL : Clinical implications. *Best Practice & Research. Clinical Haematology*, 33(3), 101193. <https://doi.org/10.1016/j.beha.2020.101193>
- Klaus, A., & Robin, C. (2017). Embryonic hematopoiesis under microscopic observation. *Developmental Biology*, 428(2), 318-327. <https://doi.org/10.1016/j.ydbio.2017.03.008>
- Klein, U., Casola, S., Cattoretti, G., Shen, Q., Lia, M., Mo, T., Ludwig, T., Rajewsky, K., & Dalla-Favera, R. (2006). Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nature Immunology*, 7(7), 773-782. <https://doi.org/10.1038/ni1357>
- Kolhe, R., Reid, M. D., Lee, J. R., Cohen, C., & Ramalingam, P. (2013). Immunohistochemical expression of PAX5 and TdT by Merkel cell carcinoma and pulmonary small cell carcinoma : A potential diagnostic pitfall but useful discriminatory marker. *International Journal of Clinical and Experimental Pathology*, 6(2), 142-147.
- Kometani, K., & Kurosaki, T. (2015). Differentiation and maintenance of long-lived plasma cells. *Current Opinion in Immunology*, 33, 64-69. <https://doi.org/10.1016/j.coi.2015.01.017>
- Kometani, K., Nakagawa, R., Shinnakasu, R., Kaji, T., Rybouchkin, A., Moriyama, S., Furukawa, K., Koseki, H., Takemori, T., & Kurosaki, T. (2013). Repression of the transcription factor Bach2 contributes to predisposition of IgG1 memory B cells toward plasma cell differentiation. *Immunity*, 39(1), 136-147. <https://doi.org/10.1016/j.immuni.2013.06.011>
- Kondo, M., Weissman, I. L., & Akashi, K. (1997). Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell*, 91(5), 661-672. [https://doi.org/10.1016/s0092-8674\(00\)80453-5](https://doi.org/10.1016/s0092-8674(00)80453-5)
- Korsmeyer, S. J. (1992). Chromosomal translocations in lymphoid malignancies reveal novel proto-oncogenes. *Annual Review of Immunology*, 10, 785-807. <https://doi.org/10.1146/annurev.iy.10.040192.004033>
- Kosik, P., Skorvaga, M., Durdik, M., Jakl, L., Nikitina, E., Markova, E., Kozics, K., Horvathova, E., & Belyaev, I. (2017). Low numbers of pre-leukemic fusion genes are frequently present in umbilical cord blood without affecting DNA damage response. *Oncotarget*, 8(22), 35824-35834. <https://doi.org/10.18632/oncotarget.16211>
- Kovac, C. R., Emelyanov, A., Singh, M., Ashouian, N., & Birshtein, B. K. (2000). BSAP (Pax5)-importin alpha 1 (Rch1) interaction identifies a nuclear localization sequence. *The*

- Journal of Biological Chemistry*, 275(22), 16752-16757.  
<https://doi.org/10.1074/jbc.M001551200>
- Kozmik, Z., Sure, U., Rüedi, D., Busslinger, M., & Aguzzi, A. (1995). Deregulated expression of PAX5 in medulloblastoma. *Proceedings of the National Academy of Sciences of the United States of America*, 92(12), 5709-5713. <https://doi.org/10.1073/pnas.92.12.5709>
- Kozmik, Z., Wang, S., Dörfler, P., Adams, B., & Busslinger, M. (1992). The promoter of the CD19 gene is a target for the B-cell-specific transcription factor BSAP. *Molecular and Cellular Biology*, 12(6), 2662-2672. <https://doi.org/10.1128/mcb.12.6.2662>
- Kroll, T. G., Sarraf, P., Pecciarini, L., Chen, C. J., Mueller, E., Spiegelman, B. M., & Fletcher, J. A. (2000). PAX8-PPARgamma1 fusion oncogene in human thyroid carcinoma [corrected]. *Science (New York, N.Y.)*, 289(5483), 1357-1360. <https://doi.org/10.1126/science.289.5483.1357>
- Kuiper, R. P., Schoenmakers, E. F. P. M., van Reijmersdal, S. V., Hehir-Kwa, J. Y., van Kessel, A. G., van Leeuwen, F. N., & Hoogerbrugge, P. M. (2007). High-resolution genomic profiling of childhood ALL reveals novel recurrent genetic lesions affecting pathways involved in lymphocyte differentiation and cell cycle progression. *Leukemia*, 21(6), 1258-1266. <https://doi.org/10.1038/sj.leu.2404691>
- Laffleur, B., Denis-Lagache, N., Péron, S., Sirac, C., Moreau, J., & Cogné, M. (2014). AID-induced remodeling of immunoglobulin genes and B cell fate. *Oncotarget*, 5(5), 1118-1131. <https://doi.org/10.18632/oncotarget.1546>
- Lang, D., Mascarenhas, J. B., Powell, S. K., Halegoua, J., Nelson, M., & Ruggeri, B. A. (2008). PAX6 is expressed in pancreatic adenocarcinoma and is downregulated during induction of terminal differentiation. *Molecular Carcinogenesis*, 47(2), 148-156. <https://doi.org/10.1002/mc.20375>
- Largeaud, L., Jamrog, L., Fregona, V., Hamelle, C., Dufrechou, S., Prade, N., Hebrard, S., Sellam, E., Bayet, M., Didier, C., Gerby, B., Delabesse, E., Pasquet, M., & Broccardo, C. (2021). New Approach to Decipher GATA2 Deficiency Syndrome : Focus on the Recurrent GATA2 R396Q Mutation. *Blood*, 138(Supplement 1), 27. <https://doi.org/10.1182/blood-2021-148307>
- Lasorella, A., Benezra, R., & Iavarone, A. (2014). The ID proteins : Master regulators of cancer stem cells and tumour aggressiveness. *Nature Reviews Cancer*, 14(2), Article 2. <https://doi.org/10.1038/nrc3638>
- Lazarus, A. H., Kawachi, K., Rapoport, M. J., & Delovitch, T. L. (1993). Antigen-induced B lymphocyte activation involves the p21ras and ras.GAP signaling pathway. *The Journal of Experimental Medicine*, 178(5), 1765-1769. <https://doi.org/10.1084/jem.178.5.1765>

- Le Guyader-Peyrou, S., Belot, A., Maynadié, M., Binder-Foucard, F., Remontet, L., Troussard, X., Bossard, N., & Monnereau, A. (2016). Cancer incidence in France over the 1980–2012 period: Hematological malignancies. *Revue d'Épidémiologie et de Santé Publique*, *64*(2), 103-112. <https://doi.org/10.1016/j.respe.2015.12.017>
- Lee, S.-T., Ji, Y., Kim, H.-J., Ki, C.-S., Jung, C. W., Kim, J.-W., & Kim, S.-H. (2012). Sequential array comparative genomic hybridization analysis identifies copy number changes during blastic transformation of chronic myeloid leukemia. *Leukemia Research*, *36*(4), 418-421. <https://doi.org/10.1016/j.leukres.2011.12.021>
- Lemmers, B., Arnoulet, C., Fossat, C., Chambost, H., Sainty, D., Gabert, J., & Schiff, C. (2000). Fine characterization of childhood and adult acute lymphoblastic leukemia (ALL) by a proB and preB surrogate light chain-specific mAb and a proposal for a new B cell ALL classification. *Leukemia*, *14*(12), 2103-2111. <https://doi.org/10.1038/sj.leu.2401932>
- Lengline, E., Beldjord, K., Dombret, H., Soulier, J., Boissel, N., & Clappier, E. (2013). Successful tyrosine kinase inhibitor therapy in a refractory B-cell precursor acute lymphoblastic leukemia with EBF1-PDGFRB fusion. *Haematologica*, *98*(11), e146-148. <https://doi.org/10.3324/haematol.2013.095372>
- Ley, T. J., Ding, L., Walter, M. J., McLellan, M. D., Lamprecht, T., Larson, D. E., Kandoth, C., Payton, J. E., Baty, J., Welch, J., Harris, C. C., Lichti, C. F., Townsend, R. R., Fulton, R. S., Dooling, D. J., Koboldt, D. C., Schmidt, H., Zhang, Q., Osborne, J. R., ... Wilson, R. K. (2010). DNMT3A mutations in acute myeloid leukemia. *The New England Journal of Medicine*, *363*(25), 2424-2433. <https://doi.org/10.1056/NEJMoa1005143>
- Li, J.-F., Dai, Y.-T., Lilljebjörn, H., Shen, S.-H., Cui, B.-W., Bai, L., Liu, Y.-F., Qian, M.-X., Kubota, Y., Kiyoi, H., Matsumura, I., Miyazaki, Y., Olsson, L., Tan, A. M., Ariffin, H., Chen, J., Takita, J., Yasuda, T., Mano, H., ... Huang, J.-Y. (2018). Transcriptional landscape of B cell precursor acute lymphoblastic leukemia based on an international study of 1,223 cases. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(50), E11711-E11720. <https://doi.org/10.1073/pnas.1814397115>
- Li, L.-X., Goetz, C. A., Katerndahl, C. D. S., Sakaguchi, N., & Farrar, M. A. (2010). A Flt3- and Ras-dependent pathway primes B cell development by inducing a state of IL-7 responsiveness. *Journal of Immunology (Baltimore, Md.: 1950)*, *184*(4), 1728-1736. <https://doi.org/10.4049/jimmunol.0903023>
- Li, Y., Gupta, G., Molofsky, A., Xie, Y., Shihabi, N., McCormick, J., & Jaffe, E. S. (2018). B Lymphoblastic Leukemia/Lymphoma with Burkitt-like morphology and IGH/MYC rearrangement: Report of three cases in adult patients. *The American journal of surgical pathology*, *42*(2), 269-276. <https://doi.org/10.1097/PAS.0000000000000982>

- Li, Y., Schwab, C., Ryan, S., Papaemmanuil, E., Robinson, H. M., Jacobs, P., Moorman, A. V., Dyer, S., Borrow, J., Griffiths, M., Heerema, N. A., Carroll, A. J., Talley, P., Bown, N., Telford, N., Ross, F. M., Gaunt, L., McNally, R. J. Q., Young, B. D., ... Harrison, C. J. (2014). Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia. *Nature*, *508*(7494), 98-102. <https://doi.org/10.1038/nature13115>
- Lieber, M. R., Yu, K., & Raghavan, S. C. (2006). Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. *DNA Repair*, *5*(9-10), 1234-1245. <https://doi.org/10.1016/j.dnarep.2006.05.013>
- Liew, E., Atenafu, E. G., Schimmer, A. D., Yee, K. W. L., Schuh, A. C., Minden, M. D., Gupta, V., & Brandwein, J. M. (2012). Outcomes of adult patients with relapsed acute lymphoblastic leukemia following frontline treatment with a pediatric regimen. *Leukemia Research*, *36*(12), 1517-1520. <https://doi.org/10.1016/j.leukres.2012.08.019>
- Lilljebjörn, H., Henningsson, R., Hyrenius-Wittsten, A., Olsson, L., Orsmark-Pietras, C., von Palffy, S., Askmyr, M., Rissler, M., Schrappe, M., Cario, G., Castor, A., Pronk, C. J. H., Behrendtz, M., Mitelman, F., Johansson, B., Paulsson, K., Andersson, A. K., Fontes, M., & Fioretos, T. (2016). Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia. *Nature Communications*, *7*, 11790. <https://doi.org/10.1038/ncomms11790>
- Lim, J. Y.-S., Bhatia, S., Robison, L. L., & Yang, J. J. (2014). Genomics of racial and ethnic disparities in childhood acute lymphoblastic leukemia. *Cancer*, *120*(7), 955-962. <https://doi.org/10.1002/cncr.28531>
- Lin, K.-I., Angelin-Duclos, C., Kuo, T. C., & Calame, K. (2002). Blimp-1-dependent repression of Pax-5 is required for differentiation of B cells to immunoglobulin M-secreting plasma cells. *Molecular and Cellular Biology*, *22*(13), 4771-4780. <https://doi.org/10.1128/mcb.22.13.4771-4780.2002>
- Lin, L., Gerth, A. J., & Peng, S. L. (2004). Active inhibition of plasma cell development in resting B cells by microphthalmia-associated transcription factor. *The Journal of Experimental Medicine*, *200*(1), 115-122. <https://doi.org/10.1084/jem.20040612>
- Lin, Y. C., Jhunjhunwala, S., Benner, C., Heinz, S., Welinder, E., Mansson, R., Sigvardsson, M., Hagman, J., Espinoza, C. A., Dutkowski, J., Ideker, T., Glass, C. K., & Murre, C. (2010). A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nature Immunology*, *11*(7), 635-643. <https://doi.org/10.1038/ni.1891>
- Linderson, Y., Eberhard, D., Malin, S., Johansson, A., Busslinger, M., & Pettersson, S. (2004). Corecruitment of the Grg4 repressor by PU.1 is critical for Pax5-mediated repression



- of B-cell-specific genes. *EMBO Reports*, 5(3), 291-296. <https://doi.org/10.1038/sj.embor.7400089>
- Linnekin, D. (1999). Early signaling pathways activated by c-Kit in hematopoietic cells. *The International Journal of Biochemistry & Cell Biology*, 31(10), 1053-1074. [https://doi.org/10.1016/s1357-2725\(99\)00078-3](https://doi.org/10.1016/s1357-2725(99)00078-3)
- Liu, Y., Easton, J., Shao, Y., Maciaszek, J., Wang, Z., Wilkinson, M. R., McCastlain, K., Edmonson, M., Pounds, S. B., Shi, L., Zhou, X., Ma, X., Sioson, E., Li, Y., Rusch, M., Gupta, P., Pei, D., Cheng, C., Smith, M. A., ... Mullighan, C. G. (2017). The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nature Genetics*, 49(8), 1211-1218. <https://doi.org/10.1038/ng.3909>
- Lord, C. J., Tutt, A. N. J., & Ashworth, A. (2015). Synthetic lethality and cancer therapy: Lessons learned from the development of PARP inhibitors. *Annual Review of Medicine*, 66, 455-470. <https://doi.org/10.1146/annurev-med-050913-022545>
- Lorenzo, P. I., Juárez-Vicente, F., Cobo-Vuilleumier, N., García-Domínguez, M., & Gauthier, B. R. (2017). The Diabetes-Linked Transcription Factor PAX4: From Gene to Functional Consequences. *Genes*, 8(3), Article 3. <https://doi.org/10.3390/genes8030101>
- Lowen, M., Scott, G., & Zwollo, P. (2001). Functional analyses of two alternative isoforms of the transcription factor Pax-5. *The Journal of Biological Chemistry*, 276(45), 42565-42574. <https://doi.org/10.1074/jbc.M106536200>
- Lu, Z., Xie, J., Wu, G., Shen, J., Collins, R., Chen, W., Kang, X., Luo, M., Zou, Y., Huang, L. J.-S., Amatruda, J. F., Slone, T., Winick, N., Scherer, P. E., & Zhang, C. C. (2017). Fasting selectively blocks development of acute lymphoblastic leukemia via leptin-receptor upregulation. *Nature Medicine*, 23(1), 79-90. <https://doi.org/10.1038/nm.4252>
- Lukes, J., Potuckova, E., Sramkova, L., Stary, J., Starkova, J., Trka, J., Votava, F., Zuna, J., & Zaliova, M. (2018). Two novel fusion genes, AIF1L-ETV6 and ABL1-AIF1L, result together with ETV6-ABL1 from a single chromosomal rearrangement in acute lymphoblastic leukemia with prenatal origin. *Genes, Chromosomes & Cancer*, 57(9), 471-477. <https://doi.org/10.1002/gcc.6>
- Ma, X., Edmonson, M., Yergeau, D., Muzny, D. M., Hampton, O. A., Rusch, M., Song, G., Easton, J., Harvey, R. C., Wheeler, D. A., Ma, J., Doddapaneni, H., Vadodaria, B., Wu, G., Nagahawatte, P., Carroll, W. L., Chen, I.-M., Gastier-Foster, J. M., Relling, M. V., ... Zhang, J. (2015). Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia. *Nature Communications*, 6, 6604. <https://doi.org/10.1038/ncomms7604>

- Macchia, P. E., Lapi, P., Krude, H., Pirro, M. T., Missero, C., Chiovato, L., Souabni, A., Baserga, M., Tassi, V., Pinchera, A., Fenzi, G., Grüters, A., Busslinger, M., & Di Lauro, R. (1998). PAX8 mutations associated with congenital hypothyroidism caused by thyroid dysgenesis. *Nature Genetics*, *19*(1), 83-86. <https://doi.org/10.1038/ng0598-83>
- Magnaghi, P., Roberts, C., Lorain, S., Lipinski, M., & Scambler, P. J. (1998). HIRA, a mammalian homologue of *Saccharomyces cerevisiae* transcriptional co-repressors, interacts with Pax3. *Nature Genetics*, *20*(1), 74-77. <https://doi.org/10.1038/1739>
- Mahajan, P., Leavey, P. J., & Galindo, R. L. (2015). PAX genes in childhood oncogenesis : Developmental biology gone awry? *Oncogene*, *34*(21), 2681-2689. <https://doi.org/10.1038/onc.2014.209>
- Mailhos, C., André, S., Mollereau, B., Goriely, A., Hemmati-Brivanlou, A., & Desplan, C. (1998). *Drosophila* Goosecoid requires a conserved heptapeptide for repression of paired-class homeoprotein activators. *Development (Cambridge, England)*, *125*(5), 937-947.
- Malin, S., McManus, S., Cobaleda, C., Novatchkova, M., Delogu, A., Bouillet, P., Strasser, A., & Busslinger, M. (2010). Role of STAT5 in controlling cell survival and immunoglobulin gene recombination during pro-B cell development. *Nature Immunology*, *11*(2), 171-179. <https://doi.org/10.1038/ni.1827>
- Månsson, R., Hultquist, A., Luc, S., Yang, L., Anderson, K., Kharazi, S., Al-Hashmi, S., Liuba, K., Thorén, L., Adolfsson, J., Buza-Vidas, N., Qian, H., Soneji, S., Enver, T., Sigvardsson, M., & Jacobsen, S. E. W. (2007). Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity*, *26*(4), 407-419. <https://doi.org/10.1016/j.immuni.2007.02.013>
- Mansson, R., Welinder, E., Åhsberg, J., Lin, Y. C., Benner, C., Glass, C. K., Lucas, J. S., Sigvardsson, M., & Murre, C. (2012). Positive intergenic feedback circuitry, involving EBF1 and FOXO1, orchestrates B-cell fate. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(51), 21028-21033. <https://doi.org/10.1073/pnas.1211427109>
- Mansson, R., Zandi, S., Welinder, E., Tsapogas, P., Sakaguchi, N., Bryder, D., & Sigvardsson, M. (2010). Single-cell analysis of the common lymphoid progenitor compartment reveals functional and molecular heterogeneity. *Blood*, *115*(13), 2601-2609. <https://doi.org/10.1182/blood-2009-08-236398>
- Marques, A. R., Espadinha, C., Catarino, A. L., Moniz, S., Pereira, T., Sobrinho, L. G., & Leite, V. (2002). Expression of PAX8-PPAR gamma 1 rearrangements in both follicular thyroid carcinomas and adenomas. *The Journal of Clinical Endocrinology and Metabolism*, *87*(8), 3947-3952. <https://doi.org/10.1210/jcem.87.8.8756>

- Marshall, A. J., Fleming, H. E., Wu, G. E., & Paige, C. J. (1998). Modulation of the IL-7 dose-response threshold during pro-B cell differentiation is dependent on pre-B cell receptor expression. *Journal of Immunology (Baltimore, Md.: 1950)*, *161*(11), 6038-6045.
- Martin, C. H., Aifantis, I., Scimone, M. L., von Andrian, U. H., Reizis, B., von Boehmer, H., & Gounari, F. (2003). Efficient thymic immigration of B220+ lymphoid-restricted bone marrow cells with T precursor potential. *Nature Immunology*, *4*(9), 866-873. <https://doi.org/10.1038/ni965>
- Martín-Lorenzo, A., Auer, F., Chan, L. N., García-Ramírez, I., González-Herrero, I., Rodríguez-Hernández, G., Bartenhagen, C., Dugas, M., Gombert, M., Ginzl, S., Blanco, O., Orfao, A., Alonso-López, D., Rivas, J. D. L., García-Cenador, M. B., García-Criado, F. J., Müschen, M., Sánchez-García, I., Borkhardt, A., ... Hauer, J. (2018). Loss of Pax5 Exploits Sca1-BCR-ABLp190 Susceptibility to Confer the Metabolic Shift Essential for pB-ALL. *Cancer Research*, *78*(10), 2669-2679. <https://doi.org/10.1158/0008-5472.CAN-17-3262>
- Martín-Lorenzo, A., Hauer, J., Vicente-Dueñas, C., Auer, F., González-Herrero, I., García-Ramírez, I., Ginzl, S., Thiele, R., Constantinescu, S. N., Bartenhagen, C., Dugas, M., Gombert, M., Schäfer, D., Blanco, O., Mayado, A., Orfao, A., Alonso-López, D., Rivas, J. D. L., Cobaleda, C., ... Borkhardt, A. (2015). Infection Exposure Is a Causal Factor in B-cell Precursor Acute Lymphoblastic Leukemia as a Result of Pax5-Inherited Susceptibility. *Cancer Discovery*, *5*(12), 1328-1343. <https://doi.org/10.1158/2159-8290.CD-15-0892>
- Matsumura, I., Kitamura, T., Wakao, H., Tanaka, H., Hashimoto, K., Albanese, C., Downward, J., Pestell, R. G., & Kanakura, Y. (1999). Transcriptional regulation of the cyclin D1 promoter by STAT5: Its involvement in cytokine-dependent growth of hematopoietic cells. *The EMBO Journal*, *18*(5), 1367-1377. <https://doi.org/10.1093/emboj/18.5.1367>
- Maude, S. L., Teachey, D. T., Porter, D. L., & Grupp, S. A. (2015). CD19-targeted chimeric antigen receptor T-cell therapy for acute lymphoblastic leukemia. *Blood*, *125*(26), 4017-4023. <https://doi.org/10.1182/blood-2014-12-580068>
- Mayes, D. A., Hu, Y., Teng, Y., Siegel, E., Wu, X., Panda, K., Tan, F., Yung, W. K. A., & Zhou, Y.-H. (2006). PAX6 suppresses the invasiveness of glioblastoma cells and the expression of the matrix metalloproteinase-2 gene. *Cancer Research*, *66*(20), 9809-9817. <https://doi.org/10.1158/0008-5472.CAN-05-3877>
- Mayran, A., Pelletier, A., & Drouin, J. (2015). Pax factors in transcription and epigenetic remodelling. *Seminars in Cell & Developmental Biology*, *44*, 135-144. <https://doi.org/10.1016/j.semcdb.2015.07.007>

- McCarthy, N. (2010). Leukaemia: MLL makes friends and influences. *Nature Reviews. Cancer*, 10(8), 529. <https://doi.org/10.1038/nrc2904>
- Medina, K. L., Tangen, S. N., Seaburg, L. M., Thapa, P., Gwin, K. A., & Shapiro, V. S. (2013). Separation of plasmacytoid dendritic cells from B-cell-biased lymphoid progenitor (BLP) and Pre-pro B cells using PDCA-1. *PLoS One*, 8(10), e78408. <https://doi.org/10.1371/journal.pone.0078408>
- Medvedovic, J., Ebert, A., Tagoh, H., & Busslinger, M. (2011). Pax5: A master regulator of B cell development and leukemogenesis. *Advances in Immunology*, 111, 179-206. <https://doi.org/10.1016/B978-0-12-385991-4.00005-2>
- Melchers, F. (2005). B cell development and its deregulation to transformed states at the pre-B cell receptor-expressing pre-BII cell stage. *Current Topics in Microbiology and Immunology*, 294, 1-17. [https://doi.org/10.1007/3-540-29933-5\\_1](https://doi.org/10.1007/3-540-29933-5_1)
- Melchers, F., Haasner, D., Grawunder, U., Kalberer, C., Karasuyama, H., Winkler, T., & Rolink, A. G. (1994). Roles of IgH and L chains and of surrogate H and L chains in the development of cells of the B lymphocyte lineage. *Annual Review of Immunology*, 12, 209-225. <https://doi.org/10.1146/annurev.iy.12.040194.001233>
- Melchers, F., Rolink, A., Grawunder, U., Winkler, T. H., Karasuyama, H., Ghia, P., & Andersson, J. (1995). Positive and negative selection events during B lymphopoiesis. *Current Opinion in Immunology*, 7(2), 214-227. [https://doi.org/10.1016/0952-7915\(95\)80006-9](https://doi.org/10.1016/0952-7915(95)80006-9)
- Melchers, F., Strasser, A., Bauer, S. R., Kudo, A., Thalmann, P., & Rolink, A. (1991). B cell development in fetal liver. *Advances in Experimental Medicine and Biology*, 292, 201-205. [https://doi.org/10.1007/978-1-4684-5943-2\\_22](https://doi.org/10.1007/978-1-4684-5943-2_22)
- Melchers, F., ten Boekel, E., Seidl, T., Kong, X. C., Yamagami, T., Onishi, K., Shimizu, T., Rolink, A. G., & Andersson, J. (2000). Repertoire selection by pre-B-cell receptors and B-cell receptors, and genetic control of B-cell development from immature to mature B cells. *Immunological Reviews*, 175, 33-46.
- Méndez-Ferrer, S., Bonnet, D., Steensma, D. P., Hasserjian, R. P., Ghobrial, I. M., Gribben, J. G., Andreeff, M., & Krause, D. S. (2020). Bone marrow niches in haematological malignancies. *Nature Reviews. Cancer*, 20(5), 285-298. <https://doi.org/10.1038/s41568-020-0245-2>
- Miles, C., Sanchez, M. J., Sinclair, A., & Dzierzak, E. (1997). Expression of the Ly-6E.1 (Sca-1) transgene in adult hematopoietic stem cells and the developing mouse embryo. *Development (Cambridge, England)*, 124(2), 537-547.
- Miller, D. J., Hayward, D. C., Reece-Hoyes, J. S., Scholten, I., Catmull, J., Gehring, W. J., Callaerts, P., Larsen, J. E., & Ball, E. E. (2000). Pax gene diversity in the basal

- cnidarian *Acropora millepora* (Cnidaria, Anthozoa) : Implications for the evolution of the Pax gene family. *Proceedings of the National Academy of Sciences*, 97(9), 4475-4480. <https://doi.org/10.1073/pnas.97.9.4475>
- Milne, C. D., Fleming, H. E., & Paige, C. J. (2004). IL-7 does not prevent pro-B/pre-B cell maturation to the immature/sIgM(+) stage. *European Journal of Immunology*, 34(10), 2647-2655. <https://doi.org/10.1002/eji.200425400>
- Milne, C. D., & Paige, C. J. (2006). IL-7 : A key regulator of B lymphopoiesis. *Seminars in Immunology*, 18(1), 20-30. <https://doi.org/10.1016/j.smim.2005.10.003>
- Miyai, T., Takano, J., Endo, T. A., Kawakami, E., Agata, Y., Motomura, Y., Kubo, M., Kashima, Y., Suzuki, Y., Kawamoto, H., & Ikawa, T. (2018). Three-step transcriptional priming that drives the commitment of multipotent progenitors toward B cells. *Genes & Development*, 32(2), 112-126. <https://doi.org/10.1101/gad.309575.117>
- Miyamoto, T., Kakizawa, T., Ichikawa, K., Nishio, S., Kajikawa, S., & Hashizume, K. (2001). Expression of dominant negative form of PAX4 in human insulinoma. *Biochemical and Biophysical Research Communications*, 282(1), 34-40. <https://doi.org/10.1006/bbrc.2001.4552>
- Montaño, A., Ordoñez, J. L., Alonso-Pérez, V., Hernández-Sánchez, J., Santos, S., González, T., Benito, R., García-Tuñón, I., & Hernández-Rivas, J. M. (2020). ETV6/RUNX1 Fusion Gene Abrogation Decreases the Oncogenicity of Tumour Cells in a Preclinical Model of Acute Lymphoblastic Leukaemia. *Cells*, 9(1). <https://doi.org/10.3390/cells9010215>
- Morgenstern, D. A., Hasan, F., Gibson, S., Winyard, P., Sebire, N. J., & Anderson, J. (2010). PAX5 Expression in Nonhematopoietic Tissues : Reappraisal of Previous Studies. *American Journal of Clinical Pathology*, 133(3), 407-415. <https://doi.org/10.1309/AJCPZPQN0LUGKMME>
- Moriyama, T., Metzger, M. L., Wu, G., Nishii, R., Qian, M., Devidas, M., Yang, W., Cheng, C., Cao, X., Quinn, E., Raimondi, S., Gastier-Foster, J. M., Raetz, E., Larsen, E., Martin, P. L., Bowman, W. P., Winick, N., Komada, Y., Wang, S., ... Yang, J. J. (2015). Germline genetic variation in ETV6 and risk of childhood acute lymphoblastic leukaemia : A systematic genetic study. *The Lancet. Oncology*, 16(16), 1659-1666. [https://doi.org/10.1016/S1470-2045\(15\)00369-1](https://doi.org/10.1016/S1470-2045(15)00369-1)
- Moriyama, T., Relling, M. V., & Yang, J. J. (2015). Inherited genetic variation in childhood acute lymphoblastic leukemia. *Blood*, 125(26), 3988-3995. <https://doi.org/10.1182/blood-2014-12-580001>

- Mullighan, C. G. (2012). The molecular genetic makeup of acute lymphoblastic leukemia. *Hematology. American Society of Hematology. Education Program, 2012*, 389-396. <https://doi.org/10.1182/asheducation-2012.1.389>
- Mullighan, C. G. (2019). How advanced are we in targeting novel subtypes of ALL? *Best Practice & Research. Clinical Haematology*, 32(4), 101095. <https://doi.org/10.1016/j.beha.2019.101095>
- Mullighan, C. G., & Downing, J. R. (2009). Genome-wide profiling of genetic alterations in acute lymphoblastic leukemia: Recent insights and future directions. *Leukemia*, 23(7), 1209-1218. <https://doi.org/10.1038/leu.2009.18>
- Mullighan, C. G., Goorha, S., Radtke, I., Miller, C. B., Coustan-Smith, E., Dalton, J. D., Girtman, K., Mathew, S., Ma, J., Pounds, S. B., Su, X., Pui, C.-H., Relling, M. V., Evans, W. E., Shurtleff, S. A., & Downing, J. R. (2007). Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*, 446(7137), 758-764. <https://doi.org/10.1038/nature05690>
- Mullighan, C. G., Phillips, L. A., Su, X., Ma, J., Miller, C. B., Shurtleff, S. A., & Downing, J. R. (2008). Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science (New York, N.Y.)*, 322(5906), 1377-1380. <https://doi.org/10.1126/science.1164266>
- Mullighan, C. G., Su, X., Zhang, J., Radtke, I., Phillips, L. A. A., Miller, C. B., Ma, J., Liu, W., Cheng, C., Schulman, B. A., Harvey, R. C., Chen, I.-M., Clifford, R. J., Carroll, W. L., Reaman, G., Bowman, W. P., Devidas, M., Gerhard, D. S., Yang, W., ... Children's Oncology Group. (2009). Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *The New England Journal of Medicine*, 360(5), 470-480. <https://doi.org/10.1056/NEJMoa0808253>
- Mullighan, C. G., Zhang, J., Harvey, R. C., Collins-Underwood, J. R., Schulman, B. A., Phillips, L. A., Tasian, S. K., Loh, M. L., Su, X., Liu, W., Devidas, M., Atlas, S. R., Chen, I.-M., Clifford, R. J., Gerhard, D. S., Carroll, W. L., Reaman, G. H., Smith, M., Downing, J. R., ... Willman, C. L. (2009). JAK mutations in high-risk childhood acute lymphoblastic leukemia. *Proceedings of the National Academy of Sciences of the United States of America*, 106(23), 9414-9418. <https://doi.org/10.1073/pnas.0811761106>
- Muratovska, A., Zhou, C., He, S., Goodyer, P., & Eccles, M. R. (2003). Paired-Box genes are frequently expressed in cancer and often required for cancer cell survival. *Oncogene*, 22(39), 7989-7997. <https://doi.org/10.1038/sj.onc.1206766>
- Murre, C. (2018). « Big bang » of B-cell development revealed. *Genes & Development*, 32(2), 93-95. <https://doi.org/10.1101/gad.311357.118>

- Muto, A., Ochiai, K., Kimura, Y., Itoh-Nakadai, A., Calame, K. L., Ikebe, D., Tashiro, S., & Igarashi, K. (2010). Bach2 represses plasma cell gene regulatory network in B cells to promote antibody class switch. *The EMBO Journal*, *29*(23), 4048-4061. <https://doi.org/10.1038/emboj.2010.257>
- Nagata, K., Nakamura, T., Kitamura, F., Kuramochi, S., Taki, S., Campbell, K. S., & Karasuyama, H. (1997). The Ig alpha/Igbeta heterodimer on mu-negative proB cells is competent for transducing signals to induce early B cell differentiation. *Immunity*, *7*(4), 559-570. [https://doi.org/10.1016/s1074-7613\(00\)80377-5](https://doi.org/10.1016/s1074-7613(00)80377-5)
- Nahar, R., Ramezani-Rad, P., Mossner, M., Duy, C., Cerchiatti, L., Geng, H., Dovat, S., Jumaa, H., Ye, B. H., Melnick, A., & Müschen, M. (2011). Pre-B cell receptor-mediated activation of BCL6 induces pre-B cell quiescence through transcriptional repression of MYC. *Blood*, *118*(15), 4174-4178. <https://doi.org/10.1182/blood-2011-01-331181>
- Nakayama, J., Yamamoto, M., Hayashi, K., Satoh, H., Bundo, K., Kubo, M., Goitsuka, R., Farrar, M. A., & Kitamura, D. (2009). BLNK suppresses pre-B-cell leukemogenesis through inhibition of JAK3. *Blood*, *113*(7), 1483-1492. <https://doi.org/10.1182/blood-2008-07-166355>
- Nebral, K., Denk, D., Attarbaschi, A., König, M., Mann, G., Haas, O. A., & Strehl, S. (2009). Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia. *Leukemia*, *23*(1), 134-143. <https://doi.org/10.1038/leu.2008.306>
- Nebral, K., König, M., Harder, L., Siebert, R., Haas, O. A., & Strehl, S. (2007). Identification of PML as novel PAX5 fusion partner in childhood acute lymphoblastic leukaemia. *British Journal of Haematology*, *139*(2), 269-274. <https://doi.org/10.1111/j.1365-2141.2007.06731.x>
- Nechanitzky, R., Akbas, D., Scherer, S., Györy, I., Hoyler, T., Ramamoorthy, S., Diefenbach, A., & Grosschedl, R. (2013). Transcription factor EBF1 is essential for the maintenance of B cell identity and prevention of alternative fates in committed cells. *Nature Immunology*, *14*(8), 867-875. <https://doi.org/10.1038/ni.2641>
- Nera, K.-P., Kohonen, P., Narvi, E., Peippo, A., Mustonen, L., Terho, P., Koskela, K., Buerstedde, J.-M., & Lassila, O. (2006). Loss of Pax5 promotes plasma cell differentiation. *Immunity*, *24*(3), 283-293. <https://doi.org/10.1016/j.immuni.2006.02.003>
- Nishana, M., & Raghavan, S. C. (2012). Role of recombination activating genes in the generation of antigen receptor diversity and beyond. *Immunology*, *137*(4), 271-281. <https://doi.org/10.1111/imm.12009>
- Nishii, R., Baskin-Doerfler, R., Yang, W., Oak, N., Zhao, X., Yang, W., Hoshitsuki, K., Bloom, M., Verbist, K. C., Burns, M. A., Li, Z., Lin, T.-N., Qian, M., Moriyama, T., Gastier-Foster, J. M., Rabin, K. R., Raetz, E., Mullighan, C. G., Pui, C.-H., ... Yang, J. J. (2020).

- Molecular Basis of ETV6-Mediated Predisposition to Childhood Acute Lymphoblastic Leukemia. *Blood*. <https://doi.org/10.1182/blood.2020006164>
- Noetzli, L., Lo, R. W., Lee-Sherick, A. B., Callaghan, M., Noris, P., Savoia, A., Rajpurkar, M., Jones, K., Gowan, K., Balduini, C., Pecci, A., Gnan, C., De Rocco, D., Doubek, M., Li, L., Lu, L., Leung, R., Landolt-Marticorena, C., Hunger, S., ... Di Paola, J. (2015). Germline mutations in ETV6 are associated with thrombocytopenia, red cell macrocytosis and predisposition to lymphoblastic leukemia. *Nature Genetics*, *47*(5), 535-538. <https://doi.org/10.1038/ng.3253>
- Noll, M. (1993). Evolution and role of Pax genes. *Current Opinion in Genetics & Development*, *3*(4), 595-605. [https://doi.org/10.1016/0959-437x\(93\)90095-7](https://doi.org/10.1016/0959-437x(93)90095-7)
- Nornes, H. O., Dressler, G. R., Knapik, E. W., Deutsch, U., & Gruss, P. (1990). Spatially and temporally restricted expression of Pax2 during murine neurogenesis. *Development (Cambridge, England)*, *109*(4), 797-809.
- Nutt, S. L., Heavey, B., Rolink, A. G., & Busslinger, M. (1999). Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature*, *401*(6753), 556-562. <https://doi.org/10.1038/44076>
- Nutt, S. L., Hodgkin, P. D., Tarlinton, D. M., & Corcoran, L. M. (2015). The generation of antibody-secreting plasma cells. *Nature Reviews. Immunology*, *15*(3), 160-171. <https://doi.org/10.1038/nri3795>
- Nutt, S. L., & Kee, B. L. (2007). The transcriptional regulation of B cell lineage commitment. *Immunity*, *26*(6), 715-725. <https://doi.org/10.1016/j.immuni.2007.05.010>
- Nutt, S. L., Taubenheim, N., Hasbold, J., Corcoran, L. M., & Hodgkin, P. D. (2011). The genetic network controlling plasma cell differentiation. *Seminars in Immunology*, *23*(5), 341-349. <https://doi.org/10.1016/j.smim.2011.08.010>
- Ochiai, K., Katoh, Y., Ikura, T., Hoshikawa, Y., Noda, T., Karasuyama, H., Tashiro, S., Muto, A., & Igarashi, K. (2006). Plasmacytic transcription factor Blimp-1 is repressed by Bach2 in B cells. *The Journal of Biological Chemistry*, *281*(50), 38226-38234. <https://doi.org/10.1074/jbc.M607592200>
- Ochiai, K., Maienschein-Cline, M., Simonetti, G., Chen, J., Rosenthal, R., Brink, R., Chong, A. S., Klein, U., Dinner, A. R., Singh, H., & Sciammas, R. (2013). Transcriptional regulation of germinal center B and plasma cell fates by dynamical control of IRF4. *Immunity*, *38*(5), 918-929. <https://doi.org/10.1016/j.immuni.2013.04.009>
- Ochiai, K., Muto, A., Tanaka, H., Takahashi, S., & Igarashi, K. (2008). Regulation of the plasma cell transcription factor Blimp-1 gene by Bach2 and Bcl6. *International Immunology*, *20*(3), 453-460. <https://doi.org/10.1093/intimm/dxn005>



- Oettinger, M. A., Schatz, D. G., Gorka, C., & Baltimore, D. (1990). RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science (New York, N.Y.)*, *248*(4962), 1517-1523. <https://doi.org/10.1126/science.2360047>
- Ofverholm, I., Tran, A. N., Heyman, M., Zachariadis, V., Nordenskjöld, M., Nordgren, A., & Barbany, G. (2013). Impact of IKZF1 deletions and PAX5 amplifications in pediatric B-cell precursor ALL treated according to NOPHO protocols. *Leukemia*, *27*(9), 1936-1939. <https://doi.org/10.1038/leu.2013.92>
- Ohnishi, K., & Melchers, F. (2003). The nonimmunoglobulin portion of lambda5 mediates cell-autonomous pre-B cell receptor signaling. *Nature Immunology*, *4*(9), 849-856. <https://doi.org/10.1038/ni959>
- Ohno, H., Furukawa, T., Fukuhara, S., Zong, S. Q., Kamesaki, H., Shows, T. B., Le Beau, M. M., McKeithan, T. W., Kawakami, T., & Honjo, T. (1990). Molecular analysis of a chromosomal translocation, t(9;14)(p13;q32), in a diffuse large-cell lymphoma cell line expressing the Ki-1 antigen. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(2), 628-632. <https://doi.org/10.1073/pnas.87.2.628>
- O'Neil, J., & Look, A. T. (2007). Mechanisms of transcription factor deregulation in lymphoid cell transformation. *Oncogene*, *26*(47), 6838-6849. <https://doi.org/10.1038/sj.onc.1210766>
- Oppezzo, P., Dumas, G., Lalanne, A. I., Payelle-Brogard, B., Magnac, C., Pritsch, O., Dighiero, G., & Vuillier, F. (2005). Different isoforms of BSAP regulate expression of AID in normal and chronic lymphocytic leukemia B cells. *Blood*, *105*(6), 2495-2503. <https://doi.org/10.1182/blood-2004-09-3644>
- O'Riordan, M., & Grosschedl, R. (1999). Coordinate regulation of B cell differentiation by the transcription factors EBF and E2A. *Immunity*, *11*(1), 21-31. [https://doi.org/10.1016/s1074-7613\(00\)80078-3](https://doi.org/10.1016/s1074-7613(00)80078-3)
- Orkin, S. H., & Zon, L. I. (2008). Hematopoiesis : An evolving paradigm for stem cell biology. *Cell*, *132*(4), 631-644. <https://doi.org/10.1016/j.cell.2008.01.025>
- Ortmann, C. A., Kent, D. G., Nangalia, J., Silber, Y., Wedge, D. C., Grinfeld, J., Baxter, E. J., Massie, C. E., Papaemmanuil, E., Menon, S., Godfrey, A. L., Dimitropoulou, D., Guglielmelli, P., Bellosillo, B., Besses, C., Döhner, K., Harrison, C. N., Vassiliou, G. S., Vannucchi, A., ... Green, A. R. (2015). Effect of mutation order on myeloproliferative neoplasms. *The New England Journal of Medicine*, *372*(7), 601-612. <https://doi.org/10.1056/NEJMoa1412098>
- Ozcan, A., de la Roza, G., Ro, J. Y., Shen, S. S., & Truong, L. D. (2012). PAX2 and PAX8 expression in primary and metastatic renal tumors : A comprehensive comparison.

- Archives of Pathology & Laboratory Medicine*, 136(12), 1541-1551.  
<https://doi.org/10.5858/arpa.2012-0072-OA>
- Paixão-Côrtés, V. R., Salzano, F. M., & Bortolini, M. C. (2015). Origins and evolvability of the PAX family. *Seminars in Cell & Developmental Biology*, 44, 64-74.  
<https://doi.org/10.1016/j.semcdb.2015.08.014>
- Papaemmanuil, E., Hosking, F. J., Vijayakrishnan, J., Price, A., Olver, B., Sheridan, E., Kinsey, S. E., Lightfoot, T., Roman, E., Irving, J. A. E., Allan, J. M., Tomlinson, I. P., Taylor, M., Greaves, M., & Houlston, R. S. (2009). Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nature Genetics*, 41(9), 1006-1010. <https://doi.org/10.1038/ng.430>
- Park, D. J., Kwon, A., Cho, B.-S., Kim, H.-J., Hwang, K.-A., Kim, M., & Kim, Y. (2020). Characteristics of DNMT3A mutations in acute myeloid leukemia. *Blood Research*, 55(1), 17-26. <https://doi.org/10.5045/br.2020.55.1.17>
- Park, E., Chen, J., Moore, A., Mangolini, M., Santoro, A., Boyd, J. R., Schjerven, H., Ecker, V., Buchner, M., Williamson, J. C., Lehner, P. J., Gasparoli, L., Williams, O., Bloehdorn, J., Stilgenbauer, S., Leitges, M., Egle, A., Schmidt-Supprian, M., Fietze, S., & Ringshausen, I. (2020). Stromal cell protein kinase C-β inhibition enhances chemosensitivity in B cell malignancies and overcomes drug resistance. *Science Translational Medicine*, 12(526). <https://doi.org/10.1126/scitranslmed.aax9340>
- Parker, M. J., Licence, S., Erlandsson, L., Galler, G. R., Chakalova, L., Osborne, C. S., Morgan, G., Fraser, P., Jumaa, H., Winkler, T. H., Skok, J., & Mårtensson, I.-L. (2005). The pre-B-cell receptor induces silencing of VpreB and lambda5 transcription. *The EMBO Journal*, 24(22), 3895-3905. <https://doi.org/10.1038/sj.emboj.7600850>
- Pasquet, M., Bellanné-Chantelot, C., Tavitian, S., Prade, N., Beaupain, B., Larochelle, O., Petit, A., Rohrllich, P., Ferrand, C., Van Den Neste, E., Poirel, H. A., Lamy, T., Ouachée-Chardin, M., Mansat-De Mas, V., Corre, J., Récher, C., Plat, G., Bachelerie, F., Donadieu, J., & Delabesse, E. (2013). High frequency of GATA2 mutations in patients with mild chronic neutropenia evolving to MonoMac syndrome, myelodysplasia, and acute myeloid leukemia. *Blood*, 121(5), 822-829. <https://doi.org/10.1182/blood-2012-08-447367>
- Passet, M., Boissel, N., Sigaux, F., Saillard, C., Bargetzi, M., Ba, I., Thomas, X., Graux, C., Chalandon, Y., Leguay, T., Lengliné, E., Konopacki, J., Quentin, S., Delabesse, E., Lafage-Pochitaloff, M., Pastoret, C., Gardel, N., Asnafi, V., Lhéritier, V., ... Group for Research on Adult ALL (GRAALL). (2019). PAX5 P80R mutation identifies a novel subtype of B-cell precursor acute lymphoblastic leukemia with favorable outcome. *Blood*, 133(3), 280-284. <https://doi.org/10.1182/blood-2018-10-882142>

- Perez-Andreu, V., Roberts, K. G., Harvey, R. C., Yang, W., Cheng, C., Pei, D., Xu, H., Gastier-Foster, J., E, S., Lim, J. Y.-S., Chen, I.-M., Fan, Y., Devidas, M., Borowitz, M. J., Smith, C., Neale, G., Burchard, E. G., Torgerson, D. G., Klusmann, F. A., ... Yang, J. J. (2013). Inherited GATA3 variants are associated with Ph-like childhood acute lymphoblastic leukemia and risk of relapse. *Nature Genetics*, *45*(12), 1494-1498. <https://doi.org/10.1038/ng.2803>
- Pérez-Caro, M., Cobaleda, C., González-Herrero, I., Vicente-Dueñas, C., Bermejo-Rodríguez, C., Sánchez-Beato, M., Orfao, A., Pintado, B., Flores, T., Sánchez-Martín, M., Jiménez, R., Piris, M. A., & Sánchez-García, I. (2009). Cancer induction by restriction of oncogene expression to the stem cell compartment. *The EMBO Journal*, *28*(1), 8-20. <https://doi.org/10.1038/emboj.2008.253>
- Pérez-Caro, M., Gutierrez-Cianca, N., González-Herrero, I., López-Hernández, I., Flores, T., Orfao, A., Sánchez-Martín, M., Gutiérrez-Adán, A., Pintado, B., & Sánchez-García, I. (2007). Sustained leukaemic phenotype after inactivation of BCR-ABLp190 in mice. *Oncogene*, *26*(12), 1702-1713. <https://doi.org/10.1038/sj.onc.1209968>
- Péron, S., Laffleur, B., Denis-Lagache, N., Cook-Moreau, J., Tinguely, A., Delpy, L., Denizot, Y., Pinaud, E., & Cogné, M. (2012). AID-driven deletion causes immunoglobulin heavy chain locus suicide recombination in B cells. *Science (New York, N.Y.)*, *336*(6083), 931-934. <https://doi.org/10.1126/science.1218692>
- Petkau, G., & Turner, M. (2019). Signalling circuits that direct early B-cell development. *The Biochemical Journal*, *476*(5), 769-778. <https://doi.org/10.1042/BCJ20180565>
- Pfeffer, P. L., Bouchard, M., & Busslinger, M. (2000). Pax2 and homeodomain proteins cooperatively regulate a 435 bp enhancer of the mouse Pax5 gene at the midbrain-hindbrain boundary. *Development (Cambridge, England)*, *127*(5), 1017-1028.
- Pieters, R., de Groot-Kruseman, H., Van der Velden, V., Fiocco, M., van den Berg, H., de Bont, E., Egeler, R. M., Hoogerbrugge, P., Kaspers, G., Van der Schoot, E., De Haas, V., & Van Dongen, J. (2016). Successful Therapy Reduction and Intensification for Childhood Acute Lymphoblastic Leukemia Based on Minimal Residual Disease Monitoring: Study ALL10 From the Dutch Childhood Oncology Group. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, *34*(22), 2591-2601. <https://doi.org/10.1200/JCO.2015.64.6364>
- Pietras, E. M., Reynaud, D., Kang, Y.-A., Carlin, D., Calero-Nieto, F. J., Leavitt, A. D., Stuart, J. M., Göttgens, B., & Passegué, E. (2015). Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem Cell*, *17*(1), 35-46. <https://doi.org/10.1016/j.stem.2015.05.003>

- Plachov, D., Chowdhury, K., Walther, C., Simon, D., Guenet, J. L., & Gruss, P. (1990). Pax8, a murine paired box gene expressed in the developing excretory system and thyroid gland. *Development (Cambridge, England)*, 110(2), 643-651.
- Placzkowski, K. A., Reddi, H. V., Grebe, S. K. G., Eberhardt, N. L., & McIver, B. (2008, octobre 29). *The Role of the PAX8/PPAR Fusion Oncogene in Thyroid Cancer* [Review Article]. PPAR Research; Hindawi. <https://doi.org/10.1155/2008/672829>
- Pongubala, J. M. R., Northrup, D. L., Lancki, D. W., Medina, K. L., Treiber, T., Bertolino, E., Thomas, M., Grosschedl, R., Allman, D., & Singh, H. (2008). Transcription factor EBF restricts alternative lineage options and promotes B cell fate commitment independently of Pax5. *Nature Immunology*, 9(2), 203-215. <https://doi.org/10.1038/ni1555>
- Poppe, B., De Paepe, P., Michaux, L., Dastugue, N., Bastard, C., Herens, C., Moreau, E., Cavazzini, F., Yigit, N., Van Limbergen, H., De Paepe, A., Praet, M., De Wolf-Peeters, C., Wlodarska, I., & Speleman, F. (2005). PAX5/IGH rearrangement is a recurrent finding in a subset of aggressive B-NHL with complex chromosomal rearrangements. *Genes, Chromosomes & Cancer*, 44(2), 218-223. <https://doi.org/10.1002/gcc.20214>
- Possot, C., Schmutz, S., Chea, S., Boucontet, L., Louise, A., Cumano, A., & Golub, R. (2011). Notch signaling is necessary for adult, but not fetal, development of ROR $\gamma$ t(+) innate lymphoid cells. *Nature Immunology*, 12(10), 949-958. <https://doi.org/10.1038/ni.2105>
- Pridans, C., Holmes, M. L., Polli, M., Wettenhall, J. M., Dakic, A., Corcoran, L. M., Smyth, G. K., & Nutt, S. L. (2008). Identification of Pax5 target genes in early B cell differentiation. *Journal of Immunology (Baltimore, Md.: 1950)*, 180(3), 1719-1728. <https://doi.org/10.4049/jimmunol.180.3.1719>
- Pui, C.-H., Yang, J. J., Hunger, S. P., Pieters, R., Schrappe, M., Biondi, A., Vora, A., Baruchel, A., Silverman, L. B., Schmiegelow, K., Escherich, G., Horibe, K., Benoit, Y. C. M., Izraeli, S., Yeoh, A. E. J., Liang, D.-C., Downing, J. R., Evans, W. E., Relling, M. V., & Mullighan, C. G. (2015). Childhood Acute Lymphoblastic Leukemia : Progress Through Collaboration. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 33(27), 2938-2948. <https://doi.org/10.1200/JCO.2014.59.1636>
- Qian, M., Cao, X., Devidas, M., Yang, W., Cheng, C., Dai, Y., Carroll, A., Heerema, N. A., Zhang, H., Moriyama, T., Gastier-Foster, J. M., Xu, H., Raetz, E., Larsen, E., Winick, N., Bowman, W. P., Martin, P. L., Mardis, E. R., Fulton, R., ... Yang, J. J. (2018). TP53 Germline Variations Influence the Predisposition and Prognosis of B-Cell Acute Lymphoblastic Leukemia in Children. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 36(6), 591-599. <https://doi.org/10.1200/JCO.2017.75.5215>

- Rath, M. F., Bailey, M. J., Kim, J.-S., Ho, A. K., Gaildrat, P., Coon, S. L., Møller, M., & Klein, D. C. (2009). Developmental and Diurnal Dynamics of Pax4 Expression in the Mammalian Pineal Gland: Nocturnal Down-Regulation Is Mediated by Adrenergic-Cyclic Adenosine 3',5'-Monophosphate Signaling. *Endocrinology*, *150*(2), 803-811. <https://doi.org/10.1210/en.2008-0882>
- Reimold, A. M., Iwakoshi, N. N., Manis, J., Vallabhajosyula, P., Szomolanyi-Tsuda, E., Gravallesse, E. M., Friend, D., Grusby, M. J., Alt, F., & Glimcher, L. H. (2001). Plasma cell differentiation requires the transcription factor XBP-1. *Nature*, *412*(6844), 300-307. <https://doi.org/10.1038/35085509>
- Reimold, A. M., Ponath, P. D., Li, Y. S., Hardy, R. R., David, C. S., Strominger, J. L., & Glimcher, L. H. (1996). Transcription factor B cell lineage-specific activator protein regulates the gene for human X-box binding protein 1. *The Journal of Experimental Medicine*, *183*(2), 393-401. <https://doi.org/10.1084/jem.183.2.393>
- Revilla-I-Domingo, R., Bilic, I., Vilagos, B., Tagoh, H., Ebert, A., Tamir, I. M., Smeenk, L., Trupke, J., Sommer, A., Jaritz, M., & Busslinger, M. (2012). The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *The EMBO Journal*, *31*(14), 3130-3146. <https://doi.org/10.1038/emboj.2012.155>
- Reynaud, D., Demarco, I. A., Reddy, K. L., Schjerven, H., Bertolino, E., Chen, Z., Smale, S. T., Winandy, S., & Singh, H. (2008). Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nature Immunology*, *9*(8), 927-936. <https://doi.org/10.1038/ni.1626>
- Rickert, R. C. (2013). New insights into pre-BCR and BCR signalling with relevance to B cell malignancies. *Nature Reviews. Immunology*, *13*(8), 578-591. <https://doi.org/10.1038/nri3487>
- Ries, L. a. G., Smith, M. A., Gurney, J. G., Linet, M., Tamra, T., Young, J. L., & Bunin, G. R. (1999). Cancer incidence and survival among children and adolescents : United States SEER Program 1975-1995. *Cancer Incidence and Survival among Children and Adolescents: United States SEER Program 1975-1995*. <https://www.cabdirect.org/cabdirect/abstract/20083240999>
- Ritz-Laser, B., Estreicher, A., Gauthier, B. R., Mamin, A., Edlund, H., & Philippe, J. (2002). The pancreatic beta-cell-specific transcription factor Pax-4 inhibits glucagon gene expression through Pax-6. *Diabetologia*, *45*(1), 97-107. <https://doi.org/10.1007/s125-002-8249-9>
- Roberts, K. G. (2018). Why and how to treat Ph-like ALL? *Best Practice & Research. Clinical Haematology*, *31*(4), 351-356. <https://doi.org/10.1016/j.beha.2018.09.003>

- Roberts, K. G., Gu, Z., Payne-Turner, D., McCastlain, K., Harvey, R. C., Chen, I.-M., Pei, D., Iacobucci, I., Valentine, M., Pounds, S. B., Shi, L., Li, Y., Zhang, J., Cheng, C., Rambaldi, A., Tosi, M., Spinelli, O., Radich, J. P., Minden, M. D., ... Mullighan, C. G. (2017). High Frequency and Poor Outcome of Philadelphia Chromosome-Like Acute Lymphoblastic Leukemia in Adults. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 35(4), 394-401. <https://doi.org/10.1200/JCO.2016.69.0073>
- Roberts, K. G., Li, Y., Payne-Turner, D., Harvey, R. C., Yang, Y.-L., Pei, D., McCastlain, K., Ding, L., Lu, C., Song, G., Ma, J., Becksfort, J., Rusch, M., Chen, S.-C., Easton, J., Cheng, J., Boggs, K., Santiago-Morales, N., Iacobucci, I., ... Mullighan, C. G. (2014). Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. *The New England Journal of Medicine*, 371(11), 1005-1015. <https://doi.org/10.1056/NEJMoa1403088>
- Roberts, K. G., Morin, R. D., Zhang, J., Hirst, M., Zhao, Y., Su, X., Chen, S.-C., Payne-Turner, D., Churchman, M. L., Harvey, R. C., Chen, X., Kasap, C., Yan, C., Becksfort, J., Finney, R. P., Teachey, D. T., Maude, S. L., Tse, K., Moore, R., ... Mullighan, C. G. (2012). Genetic alterations activating kinase and cytokine receptor signaling in high-risk acute lymphoblastic leukemia. *Cancer Cell*, 22(2), 153-166. <https://doi.org/10.1016/j.ccr.2012.06.005>
- Roberts, K. G., Pei, D., Campana, D., Payne-Turner, D., Li, Y., Cheng, C., Sandlund, J. T., Jeha, S., Easton, J., Becksfort, J., Zhang, J., Coustan-Smith, E., Raimondi, S. C., Leung, W. H., Relling, M. V., Evans, W. E., Downing, J. R., Mullighan, C. G., & Pui, C.-H. (2014). Outcomes of children with BCR-ABL1-like acute lymphoblastic leukemia treated with risk-directed therapy based on the levels of minimal residual disease. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 32(27), 3012-3020. <https://doi.org/10.1200/JCO.2014.55.4105>
- Robichaud, G. A., Nardini, M., Laflamme, M., Cuperlovic-Culf, M., & Ouellette, R. J. (2004). Human Pax-5 C-terminal isoforms possess distinct transactivation properties and are differentially modulated in normal and malignant B cells. *The Journal of Biological Chemistry*, 279(48), 49956-49963. <https://doi.org/10.1074/jbc.M407171200>
- Robson, E. J. D., He, S.-J., & Eccles, M. R. (2006). A PANorama of PAX genes in cancer and development. *Nature Reviews. Cancer*, 6(1), 52-62. <https://doi.org/10.1038/nrc1778>
- Rodríguez-Hernández, G., Opitz, F. V., Delgado, P., Walter, C., Álvarez-Prado, Á. F., González-Herrero, I., Auer, F., Fischer, U., Janssen, S., Bartenhagen, C., Raboso-Gallego, J., Casado-García, A., Orfao, A., Blanco, O., Alonso-López, D., Rivas, J. D. L., Tena-Dávila, S. G. de, Müschen, M., Dugas, M., ... Borkhardt, A. (2019). Infectious

- stimuli promote malignant B-cell acute lymphoblastic leukemia in the absence of AID. *Nature Communications*, 10. <https://doi.org/10.1038/s41467-019-13570-y>
- Roessler, S., Györy, I., Imhof, S., Spivakov, M., Williams, R. R., Busslinger, M., Fisher, A. G., & Grosschedl, R. (2007). Distinct promoters mediate the regulation of Ebf1 gene expression by interleukin-7 and Pax5. *Molecular and Cellular Biology*, 27(2), 579-594. <https://doi.org/10.1128/MCB.01192-06>
- Rothenberg, E. V. (2014). Transcriptional control of early T and B cell developmental choices. *Annual Review of Immunology*, 32, 283-321. <https://doi.org/10.1146/annurev-immunol-032712-100024>
- Rowley, J. D. (2001). Chromosome translocations : Dangerous liaisons revisited. *Nature Reviews Cancer*, 1(3), Article 3. <https://doi.org/10.1038/35106108>
- Rozanski, C. H., Utley, A., Carlson, L. M., Farren, M. R., Murray, M., Russell, L. M., Nair, J. R., Yang, Z., Brady, W., Garrett-Sinha, L. A., Schoenberger, S. P., Green, J. M., Boise, L. H., & Lee, K. P. (2015). CD28 Promotes Plasma Cell Survival, Sustained Antibody Responses, and BLIMP-1 Upregulation through Its Distal PYAP Proline Motif. *The Journal of Immunology Author Choice*, 194(10), 4717-4728. <https://doi.org/10.4049/jimmunol.1402260>
- Rumfelt, L. L., Zhou, Y., Rowley, B. M., Shinton, S. A., & Hardy, R. R. (2006). Lineage specification and plasticity in CD19- early B cell precursors. *The Journal of Experimental Medicine*, 203(3), 675-687. <https://doi.org/10.1084/jem.20052444>
- Russell, L. J., Capasso, M., Vater, I., Akasaka, T., Bernard, O. A., Calasanz, M. J., Chandrasekaran, T., Chapiro, E., Gesk, S., Griffiths, M., Guttery, D. S., Haferlach, C., Harder, L., Heidenreich, O., Irving, J., Kearney, L., Nguyen-Khac, F., Machado, L., Minto, L., ... Harrison, C. J. (2009). Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. *Blood*, 114(13), 2688-2698. <https://doi.org/10.1182/blood-2009-03-208397>
- Russell, L. J., De Castro, D. G., Griffiths, M., Telford, N., Bernard, O., Panzer-Grümayer, R., Heidenreich, O., Moorman, A. V., & Harrison, C. J. (2009). A novel translocation, t(14;19)(q32;p13), involving IGH@ and the cytokine receptor for erythropoietin. *Leukemia*, 23(3), 614-617. <https://doi.org/10.1038/leu.2008.250>
- Sadras, T., Chan, L. N., Xiao, G., & Müschen, M. (2020). Metabolic Gatekeepers of Pathological B Cell Activation. *Annual Review of Pathology*. <https://doi.org/10.1146/annurev-pathol-061020-050135>
- Saito, M., Gao, J., Basso, K., Kitagawa, Y., Smith, P. M., Bhagat, G., Pernis, A., Pasqualucci, L., & Dalla-Favera, R. (2007). A signaling pathway mediating downregulation of BCL6

- in germinal center B cells is blocked by BCL6 gene alterations in B cell lymphoma. *Cancer Cell*, 12(3), 280-292. <https://doi.org/10.1016/j.ccr.2007.08.011>
- Sant, M., Allemani, C., Tereanu, C., De Angelis, R., Capocaccia, R., Visser, O., Marcos-Gragera, R., Maynadié, M., Simonetti, A., Lutz, J.-M., Berrino, F., & and the HAEMACARE Working Group. (2010). Incidence of hematologic malignancies in Europe by morphologic subtype : Results of the HAEMACARE project. *Blood*, 116(19), 3724-3734. <https://doi.org/10.1182/blood-2010-05-282632>
- Schäfer, D., Olsen, M., Lähnemann, D., Stanulla, M., Slany, R., Schmiegelow, K., Borkhardt, A., & Fischer, U. (2018). Five percent of healthy newborns have an ETV6-RUNX1 fusion as revealed by DNA-based GIPFEL screening. *Blood*, 131(7), 821-826. <https://doi.org/10.1182/blood-2017-09-808402>
- Scharer, C. D., Barwick, B. G., Guo, M., Bally, A. P. R., & Boss, J. M. (2018). Plasma cell differentiation is controlled by multiple cell division-coupled epigenetic programs. *Nature Communications*, 9(1), 1698. <https://doi.org/10.1038/s41467-018-04125-8>
- Schebesta, A., McManus, S., Salvaggio, G., Delogu, A., Busslinger, G. A., & Busslinger, M. (2007). Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration, and immune function. *Immunity*, 27(1), 49-63. <https://doi.org/10.1016/j.immuni.2007.05.019>
- Scheuermann, R. H., & Racila, E. (1995). CD19 antigen in leukemia and lymphoma diagnosis and immunotherapy. *Leukemia & Lymphoma*, 18(5-6), 385-397. <https://doi.org/10.3109/10428199509059636>
- Schindler, J. W., Van Buren, D., Foudi, A., Krejci, O., Qin, J., Orkin, S. H., & Hock, H. (2009). TEL-AML1 corrupts hematopoietic stem cells to persist in the bone marrow and initiate leukemia. *Cell Stem Cell*, 5(1), 43-53. <https://doi.org/10.1016/j.stem.2009.04.019>
- Schinnerl, D., Fortschegger, K., Kauer, M., Marchante, J. R. M., Kofler, R., Den Boer, M. L., & Strehl, S. (2015). The role of the Janus-faced transcription factor PAX5-JAK2 in acute lymphoblastic leukemia. *Blood*, 125(8), 1282-1291. <https://doi.org/10.1182/blood-2014-04-570960>
- Schinnerl, D., Mejstrikova, E., Schumich, A., Zaliouva, M., Fortschegger, K., Nebral, K., Attarbaschi, A., Fiser, K., Kauer, M. O., Popitsch, N., Haslinger, S., Inthal, A., Buldini, B., Basso, G., Bourquin, J.-P., Gaipa, G., Brüggemann, M., Feuerstein, T., Maurer-Granofszky, M., ... Strehl, S. (2019). CD371 cell surface expression : A unique feature of DUX4-rearranged acute lymphoblastic leukemia. *Haematologica*, 104(8), e352-e355. <https://doi.org/10.3324/haematol.2018.214353>



- Scholl, F. A., Kamarashev, J., Murmann, O. V., Geertsen, R., Dummer, R., & Schäfer, B. W. (2001). PAX3 is expressed in human melanomas and contributes to tumor cell survival. *Cancer Research*, *61*(3), 823-826.
- Schwab, C. J., Jones, L. R., Morrison, H., Ryan, S. L., Yigittop, H., Schouten, J. P., & Harrison, C. J. (2010). Evaluation of multiplex ligation-dependent probe amplification as a method for the detection of copy number abnormalities in B-cell precursor acute lymphoblastic leukemia. *Genes, Chromosomes & Cancer*, *49*(12), 1104-1113. <https://doi.org/10.1002/gcc.20818>
- Schwab, C., Nebral, K., Chilton, L., Leschi, C., Waanders, E., Boer, J. M., Žaliová, M., Sutton, R., Öfverholm, I. I., Ohki, K., Yamashita, Y., Groeneveld-Krentz, S., Froňková, E., Bakkus, M., Tchinda, J., Barbosa, T. da C., Fazio, G., Mlynarski, W., Pastorczak, A., ... Strehl, S. (2017). Intragenic amplification of PAX5: A novel subgroup in B-cell precursor acute lymphoblastic leukemia? *Blood Advances*, *1*(19), 1473-1477. <https://doi.org/10.1182/bloodadvances.2017006734>
- Sciammas, R., Shaffer, A. L., Schatz, J. H., Zhao, H., Staudt, L. M., & Singh, H. (2006). Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. *Immunity*, *25*(2), 225-236. <https://doi.org/10.1016/j.immuni.2006.07.009>
- Scott, E. W., Fisher, R. C., Olson, M. C., Kehrl, E. W., Simon, M. C., & Singh, H. (1997). PU.1 functions in a cell-autonomous manner to control the differentiation of multipotential lymphoid-myeloid progenitors. *Immunity*, *6*(4), 437-447. [https://doi.org/10.1016/s1074-7613\(00\)80287-3](https://doi.org/10.1016/s1074-7613(00)80287-3)
- Scouten, W. T., Patel, A., Terrell, R., Burch, H. B., Bernet, V. J., Tuttle, R. M., & Francis, G. L. (2004). Cytoplasmic localization of the paired box gene, Pax-8, is found in pediatric thyroid cancer and may be associated with a greater risk of recurrence. *Thyroid: Official Journal of the American Thyroid Association*, *14*(12), 1037-1046. <https://doi.org/10.1089/thy.2004.14.1037>
- Sexl, V., Piekorz, R., Moriggl, R., Rohrer, J., Brown, M. P., Bunting, K. D., Rothhammer, K., Roussel, M. F., & Ihle, J. N. (2000). Stat5a/b contribute to interleukin 7-induced B-cell precursor expansion, but abl- and bcr/abl-induced transformation are independent of stat5. *Blood*, *96*(6), 2277-2283.
- Shaffer, A. L., Emre, N. C. T., Romesser, P. B., & Staudt, L. M. (2009). IRF4: Immunity. Malignancy! Therapy? *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *15*(9), 2954-2961. <https://doi.org/10.1158/1078-0432.CCR-08-1845>

- Shaffer, A. L., Lin, K. I., Kuo, T. C., Yu, X., Hurt, E. M., Rosenwald, A., Giltzane, J. M., Yang, L., Zhao, H., Calame, K., & Staudt, L. M. (2002). Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity*, 17(1), 51-62. [https://doi.org/10.1016/s1074-7613\(02\)00335-7](https://doi.org/10.1016/s1074-7613(02)00335-7)
- Shah, S., Schrader, K. A., Waanders, E., Timms, A. E., Vijai, J., Miething, C., Wechsler, J., Yang, J., Hayes, J., Klein, R. J., Zhang, J., Wei, L., Wu, G., Rusch, M., Nagahawatte, P., Ma, J., Chen, S.-C., Song, G., Cheng, J., ... Offit, K. (2013). A recurrent germline PAX5 mutation confers susceptibility to pre-B cell acute lymphoblastic leukemia. *Nature Genetics*, 45(10), 1226-1231. <https://doi.org/10.1038/ng.2754>
- Shapiro-Shelef, M., & Calame, K. (2005). Regulation of plasma-cell development. *Nature Reviews. Immunology*, 5(3), 230-242. <https://doi.org/10.1038/nri1572>
- Sherborne, A. L., Hosking, F. J., Prasad, R. B., Kumar, R., Koehler, R., Vijayakrishnan, J., Papaemmanuil, E., Bartram, C. R., Stanulla, M., Schrappe, M., Gast, A., Dobbins, S. E., Ma, Y., Sheridan, E., Taylor, M., Kinsey, S. E., Lightfoot, T., Roman, E., Irving, J. A. E., ... Houlston, R. S. (2010). Variation in CDKN2A at 9p21.3 influences childhood acute lymphoblastic leukemia risk. *Nature Genetics*, 42(6), 492-494. <https://doi.org/10.1038/ng.585>
- Shimajiri, Y., Sanke, T., Furuta, H., Hanabusa, T., Nakagawa, T., Fujitani, Y., Kajimoto, Y., Takasu, N., & Nanjo, K. (2001). A missense mutation of Pax4 gene (R121W) is associated with type 2 diabetes in Japanese. *Diabetes*, 50(12), 2864-2869. <https://doi.org/10.2337/diabetes.50.12.2864>
- Shurtleff, S. A., Buijs, A., Behm, F. G., Rubnitz, J. E., Raimondi, S. C., Hancock, M. L., Chan, G. C., Pui, C. H., Grosveld, G., & Downing, J. R. (1995). TEL/AML1 fusion resulting from a cryptic t(12;21) is the most common genetic lesion in pediatric ALL and defines a subgroup of patients with an excellent prognosis. *Leukemia*, 9(12), 1985-1989.
- Sigvardsson, M. (2018). Molecular Regulation of Differentiation in Early B-Lymphocyte Development. *International Journal of Molecular Sciences*, 19(7). <https://doi.org/10.3390/ijms19071928>
- Sigvardsson, M., O'Riordan, M., & Grosschedl, R. (1997). EBF and E47 collaborate to induce expression of the endogenous immunoglobulin surrogate light chain genes. *Immunity*, 7(1), 25-36. [https://doi.org/10.1016/s1074-7613\(00\)80507-5](https://doi.org/10.1016/s1074-7613(00)80507-5)
- Sivakamasundari, V., Kraus, P., & Lufkin, T. (2018). Regulatory Functions of Pax1 and Pax9 in Mammalian Cells. *Gene Expression and Regulation in Mammalian Cells - Transcription Toward the Establishment of Novel Therapeutics*. <https://doi.org/10.5772/intechopen.71920>

- Slayton, W. B., Schultz, K. R., Kairalla, J. A., Devidas, M., Mi, X., Pulsipher, M. A., Chang, B. H., Mullighan, C., Iacobucci, I., Silverman, L. B., Borowitz, M. J., Carroll, A. J., Heerema, N. A., Gastier-Foster, J. M., Wood, B. L., Mizrahy, S. L., Merchant, T., Brown, V. I., Sieger, L., ... Hunger, S. P. (2018). Dasatinib Plus Intensive Chemotherapy in Children, Adolescents, and Young Adults With Philadelphia Chromosome–Positive Acute Lymphoblastic Leukemia: Results of Children’s Oncology Group Trial AALL0622. *Journal of Clinical Oncology*, 36(22), 2306-2314. <https://doi.org/10.1200/JCO.2017.76.7228>
- Smeenk, L., Fischer, M., Jurado, S., Jaritz, M., Azaryan, A., Werner, B., Roth, M., Zuber, J., Stanulla, M., den Boer, M. L., Mullighan, C. G., Strehl, S., & Busslinger, M. (2017). Molecular role of the PAX5-ETV6 oncoprotein in promoting B-cell acute lymphoblastic leukemia. *The EMBO Journal*, 36(6), 718-735. <https://doi.org/10.15252/emj.201695495>
- Smith, S. B., Ee, H. C., Conners, J. R., & German, M. S. (1999). Paired-homeodomain transcription factor PAX4 acts as a transcriptional repressor in early pancreatic development. *Molecular and Cellular Biology*, 19(12), 8272-8280. <https://doi.org/10.1128/mcb.19.12.8272>
- Souabni, A., Cobaleda, C., Schebesta, M., & Busslinger, M. (2002). Pax5 promotes B lymphopoiesis and blocks T cell development by repressing Notch1. *Immunity*, 17(6), 781-793. [https://doi.org/10.1016/s1074-7613\(02\)00472-7](https://doi.org/10.1016/s1074-7613(02)00472-7)
- Souabni, A., Jochum, W., & Busslinger, M. (2007). Oncogenic role of Pax5 in the T-lymphoid lineage upon ectopic expression from the immunoglobulin heavy-chain locus. *Blood*, 109(1), 281-289. <https://doi.org/10.1182/blood-2006-03-009670>
- Spooner, C. J., Cheng, J. X., Pujadas, E., Laslo, P., & Singh, H. (2009). A recurrent network involving the transcription factors PU.1 and Gfi1 orchestrates innate and adaptive immune cell fates. *Immunity*, 31(4), 576-586. <https://doi.org/10.1016/j.immuni.2009.07.011>
- Srinivasan, L., Sasaki, Y., Calado, D. P., Zhang, B., Paik, J. H., DePinho, R. A., Kutok, J. L., Kearney, J. F., Otipoby, K. L., & Rajewsky, K. (2009). PI3 kinase signals BCR-dependent mature B cell survival. *Cell*, 139(3), 573-586. <https://doi.org/10.1016/j.cell.2009.08.041>
- Stapleton, P., Weith, A., Urbánek, P., Kozmik, Z., & Busslinger, M. (1993). Chromosomal localization of seven PAX genes and cloning of a novel family member, PAX-9. *Nature Genetics*, 3(4), 292-298. <https://doi.org/10.1038/ng0493-292>

- Stockton, D. W., Das, P., Goldenberg, M., D'Souza, R. N., & Patel, P. I. (2000). Mutation of PAX9 is associated with oligodontia. *Nature Genetics*, 24(1), 18-19. <https://doi.org/10.1038/71634>
- Stuart, E. T., Kioussi, C., & Gruss, P. (1994). Mammalian Pax genes. *Annual Review of Genetics*, 28, 219-236. <https://doi.org/10.1146/annurev.ge.28.120194.001251>
- Su, I.-H., Basavaraj, A., Krutchinsky, A. N., Hobert, O., Ullrich, A., Chait, B. T., & Tarakhovskiy, A. (2003). Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nature Immunology*, 4(2), 124-131. <https://doi.org/10.1038/ni876>
- Swain, A., Inoue, T., Tan, K. S., Nakanishi, Y., & Sugiyama, D. (2014). Intrinsic and extrinsic regulation of mammalian hematopoiesis in the fetal liver. *Histology and Histopathology*, 29(9), 1077-1082. <https://doi.org/10.14670/HH-29.1077>
- Tagoh, H., Schebesta, A., Lefevre, P., Wilson, N., Hume, D., Busslinger, M., & Bonifer, C. (2004). Epigenetic silencing of the c-fms locus during B-lymphopoiesis occurs in discrete steps and is reversible. *The EMBO Journal*, 23(21), 4275-4285. <https://doi.org/10.1038/sj.emboj.7600421>
- Tanasi, I., Ba, I., Sirvent, N., Braun, T., Cuccuini, W., Ballerini, P., Duployez, N., Tanguy-Schmidt, A., Tamburini, J., Maury, S., Doré, E., Hemberlin, C., Duclos, C., Chevallier, P., Rousselot, P., Bonifacio, M., Cavé, H., Baruchel, A., Dombret, H., ... Clappier, E. (2019). Efficacy of tyrosine kinase inhibitors in Ph-like acute lymphoblastic leukemia harboring ABL-class rearrangements. *Blood*, 134(16), 1351-1355. <https://doi.org/10.1182/blood.2019001244>
- Tassabehji, M., Read, A. P., Newton, V. E., Harris, R., Balling, R., Gruss, P., & Strachan, T. (1992). Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature*, 355(6361), 635-636. <https://doi.org/10.1038/355635a0>
- Taub, J. W., Konrad, M. A., Ge, Y., Naber, J. M., Scott, J. S., Matherly, L. H., & Ravindranath, Y. (2002). High frequency of leukemic clones in newborn screening blood samples of children with B-precursor acute lymphoblastic leukemia. *Blood*, 99(8), 2992-2996. <https://doi.org/10.1182/blood.v99.8.2992>
- Thal, M. A., Carvalho, T. L., He, T., Kim, H.-G., Gao, H., Hagman, J., & Klug, C. A. (2009). Ebf1-mediated down-regulation of Id2 and Id3 is essential for specification of the B cell lineage. *Proceedings of the National Academy of Sciences of the United States of America*, 106(2), 552-557. <https://doi.org/10.1073/pnas.0802550106>
- Thompson, E. C., Cobb, B. S., Sabbattini, P., Meixlsperger, S., Parelho, V., Liberg, D., Taylor, B., Dillon, N., Georgopoulos, K., Jumaa, H., Smale, S. T., Fisher, A. G., & Merkenschlager, M. (2007). Ikaros DNA-binding proteins as integral components of B

- cell developmental-stage-specific regulatory circuits. *Immunity*, 26(3), 335-344. <https://doi.org/10.1016/j.immuni.2007.02.010>
- Tiacci, E., Pileri, S., Orleth, A., Pacini, R., Tabarrini, A., Frenguelli, F., Liso, A., Diverio, D., Lo-Coco, F., & Falini, B. (2004). PAX5 expression in acute leukemias: Higher B-lineage specificity than CD79a and selective association with t(8;21)-acute myelogenous leukemia. *Cancer Research*, 64(20), 7399-7404. <https://doi.org/10.1158/0008-5472.CAN-04-1865>
- Tomasetti, C., & Vogelstein, B. (2015). Cancer etiology. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. *Science (New York, N.Y.)*, 347(6217), 78-81. <https://doi.org/10.1126/science.1260825>
- Ton, C. C., Hirvonen, H., Miwa, H., Weil, M. M., Monaghan, P., Jordan, T., van Heyningen, V., Hastie, N. D., Meijers-Heijboer, H., & Drechsler, M. (1991). Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. *Cell*, 67(6), 1059-1074. [https://doi.org/10.1016/0092-8674\(91\)90284-6](https://doi.org/10.1016/0092-8674(91)90284-6)
- Torlakovic, E., Slipicevic, A., Robinson, C., DeCoteau, J. F., Alfsen, G. C., Vyberg, M., Chibbar, R., & Flørenes, V. A. (2006). Pax-5 Expression in Nonhematopoietic Tissues. *American Journal of Clinical Pathology*, 126(5), 798-804. <https://doi.org/10.1309/XEC7JMW9YRM74RN0>
- Toyonaga, K., Kikuchi, H., Yamashita, K., Nakayama, M., Chijiwa, K., & Nakayama, T. (2009). E2A participates in a fine control of pre-mature B-cell apoptosis mediated by B-cell receptor signaling via transcriptional regulation of survivin, IAP2 and caspase-8 genes. *The FEBS Journal*, 276(5), 1418-1428. <https://doi.org/10.1111/j.1742-4658.2009.06881.x>
- Treiber, T., Mandel, E. M., Pott, S., Györy, I., Firner, S., Liu, E. T., & Grosschedl, R. (2010). Early B cell factor 1 regulates B cell gene networks by activation, repression, and transcription-independent poising of chromatin. *Immunity*, 32(5), 714-725. <https://doi.org/10.1016/j.immuni.2010.04.013>
- Treisman, J., Harris, E., & Desplan, C. (1991). The paired box encodes a second DNA-binding domain in the paired homeo domain protein. *Genes & Development*, 5(4), 594-604. <https://doi.org/10.1101/gad.5.4.594>
- Tremblay, P., & Gruss, P. (1994). Pax: Genes for mice and men. *Pharmacology & Therapeutics*, 61(1-2), 205-226. [https://doi.org/10.1016/0163-7258\(94\)90063-9](https://doi.org/10.1016/0163-7258(94)90063-9)
- Treviño, L. R., Yang, W., French, D., Hunger, S. P., Carroll, W. L., Devidas, M., Willman, C., Neale, G., Downing, J., Raimondi, S. C., Pui, C.-H., Evans, W. E., & Relling, M. V. (2009). Germline genomic variants associated with childhood acute lymphoblastic leukemia. *Nature Genetics*, 41(9), 1001-1005. <https://doi.org/10.1038/ng.432>

- Tsuber, V., Kadamov, Y., Brautigam, L., Warpman Berglund, U., & Helleday, T. (2017). Mutations in Cancer Cause Gain of Cysteine, Histidine, and Tryptophan at the Expense of a Net Loss of Arginine on the Proteome Level. *Biomolecules*, 7(3). <https://doi.org/10.3390/biom7030049>
- Tsukamoto, K., Nakamura, Y., & Niikawa, N. (1994). Isolation of two isoforms of the PAX3 gene transcripts and their tissue-specific alternative expression in human adult tissues. *Human Genetics*, 93(3), 270-274. <https://doi.org/10.1007/BF00212021>
- Tsuzuki, S., Seto, M., Greaves, M., & Enver, T. (2004). Modeling first-hit functions of the t(12;21) TEL-AML1 translocation in mice. *Proceedings of the National Academy of Sciences of the United States of America*, 101(22), 8443-8448. <https://doi.org/10.1073/pnas.0402063101>
- Turki-Judeh, W., & Courey, A. J. (2012). Groucho: A corepressor with instructive roles in development. *Current Topics in Developmental Biology*, 98, 65-96. <https://doi.org/10.1016/B978-0-12-386499-4.00003-3>
- Urayama, K. Y., Buffler, P. A., Gallagher, E. R., Ayoob, J. M., & Ma, X. (2010). A meta-analysis of the association between day-care attendance and childhood acute lymphoblastic leukaemia. *International Journal of Epidemiology*, 39(3), 718-732. <https://doi.org/10.1093/ije/dyp378>
- Urbánek, P., Fetka, I., Meisler, M. H., & Busslinger, M. (1997). Cooperation of Pax2 and Pax5 in midbrain and cerebellum development. *Proceedings of the National Academy of Sciences of the United States of America*, 94(11), 5703-5708. <https://doi.org/10.1073/pnas.94.11.5703>
- Urbánek, P., Wang, Z. Q., Fetka, I., Wagner, E. F., & Busslinger, M. (1994). Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5/BSAP. *Cell*, 79(5), 901-912. [https://doi.org/10.1016/0092-8674\(94\)90079-5](https://doi.org/10.1016/0092-8674(94)90079-5)
- Utley, A., Chavel, C., Lightman, S., Holling, G. A., Cooper, J., Peng, P., Liu, W., Barwick, B. G., Gavile, C. M., Maguire, O., Murray-Dupuis, M., Rozanski, C., Jordan, M. S., Kambayashi, T., Olejniczak, S. H., Boise, L. H., & Lee, K. P. (2020). CD28 Regulates Metabolic Fitness for Long-Lived Plasma Cell Survival. *Cell Reports*, 31(12), 107815. <https://doi.org/10.1016/j.celrep.2020.107815>
- van der Weyden, L., Giotopoulos, G., Wong, K., Rust, A. G., Robles-Espinoza, C. D., Osaki, H., Huntly, B. J., & Adams, D. J. (2015). Somatic drivers of B-ALL in a model of ETV6-RUNX1; Pax5(+/-) leukemia. *BMC Cancer*, 15, 585. <https://doi.org/10.1186/s12885-015-1586-1>

- Vettermann, C., Herrmann, K., & Jäck, H.-M. (2006). Powered by pairing : The surrogate light chain amplifies immunoglobulin heavy chain signaling and pre-selects the antibody repertoire. *Seminars in Immunology*, 18(1), 44-55. <https://doi.org/10.1016/j.smim.2006.01.001>
- Vicente Dueñas, C., Janssen, S., Oldenburg, M., Auer, F., González Herrero, I., Casado-García, A., Isidro-Hernández, M., Raboso-Gallego, J., Westhoff, P., Pandyra, A. A., Hein, D., Gössling, K. L., Alonso-López, D., De Las Rivas, J., Bhatia, S., Garcia-Criado, F. J., García Cenador, M. B., Weber, A. P. M., Köhrer, K., ... Borkhardt, A. (2020). An intact gut microbiome protects genetically predisposed mice against leukemia. *Blood*. <https://doi.org/10.1182/blood.2019004381>
- Vijayakrishnan, J., & Houlston, R. S. (2010). Candidate gene association studies and risk of childhood acute lymphoblastic leukemia : A systematic review and meta-analysis. *Haematologica*, 95(8), 1405-1414. <https://doi.org/10.3324/haematol.2010.022095>
- Vijayakrishnan, J., Qian, M., Studd, J. B., Yang, W., Kinnersley, B., Law, P. J., Broderick, P., Raetz, E. A., Allan, J., Pui, C.-H., Vora, A., Evans, W. E., Moorman, A., Yeoh, A., Yang, W., Li, C., Bartram, C. R., Mullighan, C. G., Zimmerman, M., ... Yang, J. J. (2019). Identification of four novel associations for B-cell acute lymphoblastic leukaemia risk. *Nature Communications*, 10(1), 5348. <https://doi.org/10.1038/s41467-019-13069-6>
- Vijayakrishnan, J., Studd, J., Broderick, P., Kinnersley, B., Holroyd, A., Law, P. J., Kumar, R., Allan, J. M., Harrison, C. J., Moorman, A. V., Vora, A., Roman, E., Rachakonda, S., Kinsey, S. E., Sheridan, E., Thompson, P. D., Irving, J. A., Koehler, R., Hoffmann, P., ... Houlston, R. S. (2018). Genome-wide association study identifies susceptibility loci for B-cell childhood acute lymphoblastic leukemia. *Nature Communications*, 9(1), 1340. <https://doi.org/10.1038/s41467-018-03178-z>
- Vogelstein, B., & Kinzler, K. W. (1993). The multistep nature of cancer. *Trends in Genetics: TIG*, 9(4), 138-141. [https://doi.org/10.1016/0168-9525\(93\)90209-z](https://doi.org/10.1016/0168-9525(93)90209-z)
- Vogelstein, B., & Kinzler, K. W. (2015). The Path to Cancer—Three Strikes and You're Out. *The New England Journal of Medicine*, 373(20), 1895-1898. <https://doi.org/10.1056/NEJMp1508811>
- Vorobyov, E., & Horst, J. (2006). Getting the proto-Pax by the tail. *Journal of Molecular Evolution*, 63(2), 153-164. <https://doi.org/10.1007/s00239-005-0163-7>
- Wachtel, M., & Schäfer, B. W. (2015). Unpeaceful roles of mutant PAX proteins in cancer. *Seminars in Cell & Developmental Biology*, 44, 126-134. <https://doi.org/10.1016/j.semcdb.2015.09.011>

- Wakatsuki, Y., Neurath, M. F., Max, E. E., & Strober, W. (1994). The B cell-specific transcription factor BSAP regulates B cell proliferation. *The Journal of Experimental Medicine*, 179(4), 1099-1108. <https://doi.org/10.1084/jem.179.4.1099>
- Wallin, J., Mizutani, Y., Imai, K., Miyashita, N., Moriwaki, K., Taniguchi, M., Koseki, H., & Balling, R. (1993). A new Pax gene, Pax-9, maps to mouse chromosome 12. *Mammalian Genome: Official Journal of the International Mammalian Genome Society*, 4(7), 354-358. <https://doi.org/10.1007/BF00360584>
- Walter, K., Cockerill, P. N., Barlow, R., Clarke, D., Hoogenkamp, M., Follows, G. A., Richards, S. J., Cullen, M. J., Bonifer, C., & Tagoh, H. (2010). Aberrant expression of CD19 in AML with t(8;21) involves a poised chromatin structure and PAX5. *Oncogene*, 29(20), 2927-2937. <https://doi.org/10.1038/onc.2010.56>
- Walther, C., Guenet, J. L., Simon, D., Deutsch, U., Jostes, B., Goulding, M. D., Plachov, D., Balling, R., & Gruss, P. (1991). Pax: A murine multigene family of paired box-containing genes. *Genomics*, 11(2), 424-434. [https://doi.org/10.1016/0888-7543\(91\)90151-4](https://doi.org/10.1016/0888-7543(91)90151-4)
- Weinberg, O. K., & Arber, D. A. (2010). Mixed-phenotype acute leukemia : Historical overview and a new definition. *Leukemia*, 24(11), Article 11. <https://doi.org/10.1038/leu.2010.202>
- Welinder, E., Mansson, R., Mercer, E. M., Bryder, D., Sigvardsson, M., & Murre, C. (2011). The transcription factors E2A and HEB act in concert to induce the expression of FOXO1 in the common lymphoid progenitor. *Proceedings of the National Academy of Sciences of the United States of America*, 108(42), 17402-17407. <https://doi.org/10.1073/pnas.1111766108>
- Weston, B. W., Hayden, M. A., Roberts, K. G., Bowyer, S., Hsu, J., Fedoriw, G., Rao, K. W., & Mullighan, C. G. (2013). Tyrosine kinase inhibitor therapy induces remission in a patient with refractory EBF1-PDGFRB-positive acute lymphoblastic leukemia. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 31(25), e413-416. <https://doi.org/10.1200/JCO.2012.47.6770>
- Whitlock, J. A. (2006). Down syndrome and acute lymphoblastic leukaemia. *British Journal of Haematology*, 135(5), 595-602. <https://doi.org/10.1111/j.1365-2141.2006.06337.x>
- Wiemels, J. L., Cazzaniga, G., Daniotti, M., Eden, O. B., Addison, G. M., Masera, G., Saha, V., Biondi, A., & Greaves, M. F. (1999). Prenatal origin of acute lymphoblastic leukaemia in children. *Lancet (London, England)*, 354(9189), 1499-1503. [https://doi.org/10.1016/s0140-6736\(99\)09403-9](https://doi.org/10.1016/s0140-6736(99)09403-9)



- Wilcox, E. R., Rivolta, M. N., Ploplis, B., Potterf, S. B., & Fex, J. (1992). The PAX3 gene is mapped to human chromosome 2 together with a highly informative CA dinucleotide repeat. *Human Molecular Genetics*, 1(3), 215. <https://doi.org/10.1093/hmg/1.3.215-a>
- Wilson, D., Sheng, G., Lecuit, T., Dostatni, N., & Desplan, C. (1993). Cooperative dimerization of paired class homeo domains on DNA. *Genes & Development*, 7(11), 2120-2134. <https://doi.org/10.1101/gad.7.11.2120>
- Winters, A. C., & Bernt, K. M. (2017). MLL-Rearranged Leukemias—An Update on Science and Clinical Approaches. *Frontiers in Pediatrics*, 5. <https://doi.org/10.3389/fped.2017.00004>
- Winyard, P. J., Risdon, R. A., Sams, V. R., Dressler, G. R., & Woolf, A. S. (1996). The PAX2 transcription factor is expressed in cystic and hyperproliferative dysplastic epithelia in human kidney malformations. *The Journal of Clinical Investigation*, 98(2), 451-459. <https://doi.org/10.1172/JCI118811>
- Witkowski, M. T., Dolgalev, I., Evensen, N. A., Ma, C., Chambers, T., Roberts, K. G., Sreeram, S., Dai, Y., Tikhonova, A. N., Lasry, A., Qu, C., Pei, D., Cheng, C., Robbins, G. A., Pierro, J., Selvaraj, S., Mezzano, V., Daves, M., Lupo, P. J., ... Aifantis, I. (2020). Extensive Remodeling of the Immune Microenvironment in B Cell Acute Lymphoblastic Leukemia. *Cancer Cell*, 37(6), 867-882.e12. <https://doi.org/10.1016/j.ccell.2020.04.015>
- Wu, X., Xiao, Y., Zhou, Y., Zhou, Z., & Yan, W. (2019). LncRNA FOXP4-AS1 is activated by PAX5 and promotes the growth of prostate cancer by sequestering miR-3184-5p to upregulate FOXP4. *Cell Death & Disease*, 10(7), Article 7. <https://doi.org/10.1038/s41419-019-1699-6>
- Xu, H. E., Rould, M. A., Xu, W., Epstein, J. A., Maas, R. L., & Pabo, C. O. (1999). Crystal structure of the human Pax6 paired domain–DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. *Genes & Development*, 13(10), 1263-1275.
- Xu, H., Yang, W., Perez-Andreu, V., Devidas, M., Fan, Y., Cheng, C., Pei, D., Scheet, P., Burchard, E. G., Eng, C., Huntsman, S., Torgerson, D. G., Dean, M., Winick, N. J., Martin, P. L., Camitta, B. M., Bowman, W. P., Willman, C. L., Carroll, W. L., ... Yang, J. J. (2013). Novel susceptibility variants at 10p12.31-12.2 for childhood acute lymphoblastic leukemia in ethnically diverse populations. *Journal of the National Cancer Institute*, 105(10), 733-742. <https://doi.org/10.1093/jnci/djt042>
- Xu, W., Rould, M. A., Jun, S., Desplan, C., & Pabo, C. O. (1995). Crystal structure of a paired domain-DNA complex at 2.5 Å resolution reveals structural basis for Pax developmental mutations. *Cell*, 80(4), 639-650. [https://doi.org/10.1016/0092-8674\(95\)90518-9](https://doi.org/10.1016/0092-8674(95)90518-9)

- Yasuda, T., Hayakawa, F., Kurahashi, S., Sugimoto, K., Minami, Y., Tomita, A., & Naoe, T. (2012). B cell receptor-ERK1/2 signal cancels PAX5-dependent repression of BLIMP1 through PAX5 phosphorylation: A mechanism of antigen-triggering plasma cell differentiation. *Journal of Immunology (Baltimore, Md.: 1950)*, *188*(12), 6127-6134. <https://doi.org/10.4049/jimmunol.1103039>
- Yasuda, T., Sanjo, H., Pagès, G., Kawano, Y., Karasuyama, H., Pouysségur, J., Ogata, M., & Kurosaki, T. (2008). Erk kinases link pre-B cell receptor signaling to transcriptional events required for early B cell expansion. *Immunity*, *28*(4), 499-508. <https://doi.org/10.1016/j.immuni.2008.02.015>
- Yazdanparast, S., Khatami, S. R., Galehdari, H., & Jaseb, K. (2015). One missense mutation in exon 2 of the PAX5 gene in Iran. *Genetics and Molecular Research: GMR*, *14*(4), 17768-17775. <https://doi.org/10.4238/2015.December.22.1>
- Yoshida, T., Ng, S. Y.-M., Zuniga-Pflucker, J. C., & Georgopoulos, K. (2006). Early hematopoietic lineage restrictions directed by Ikaros. *Nature Immunology*, *7*(4), 382-391. <https://doi.org/10.1038/ni1314>
- Zaliova, M., Kotrova, M., Bresolin, S., Stuchly, J., Stary, J., Hrusak, O., Te Kronnie, G., Trka, J., Zuna, J., & Vaskova, M. (2017). ETV6/RUNX1-like acute lymphoblastic leukemia: A novel B-cell precursor leukemia subtype associated with the CD27/CD44 immunophenotype. *Genes, Chromosomes & Cancer*, *56*(8), 608-616. <https://doi.org/10.1002/gcc.22464>
- Zaliova, M., Stuchly, J., Winkowska, L., Musilova, A., Fiser, K., Slamova, M., Starkova, J., Vaskova, M., Hrusak, O., Sramkova, L., Stary, J., Zuna, J., & Trka, J. (2019). Genomic landscape of pediatric B-other acute lymphoblastic leukemia in a consecutive European cohort. *Haematologica*, *104*(7), 1396-1406. <https://doi.org/10.3324/haematol.2018.204974>
- Zandi, S., Mansson, R., Tsapogas, P., Zetterblad, J., Bryder, D., & Sigvardsson, M. (2008). EBF1 is essential for B-lineage priming and establishment of a transcription factor network in common lymphoid progenitors. *Journal of Immunology (Baltimore, Md.: 1950)*, *181*(5), 3364-3372. <https://doi.org/10.4049/jimmunol.181.5.3364>
- Zelent, A., Greaves, M., & Enver, T. (2004). Role of the TEL-AML1 fusion gene in the molecular pathogenesis of childhood acute lymphoblastic leukaemia. *Oncogene*, *23*(24), Article 24. <https://doi.org/10.1038/sj.onc.1207672>
- Zhang, J., Mullighan, C. G., Harvey, R. C., Wu, G., Chen, X., Edmonson, M., Buetow, K. H., Carroll, W. L., Chen, I.-M., Devidas, M., Gerhard, D. S., Loh, M. L., Reaman, G. H., Relling, M. V., Camitta, B. M., Bowman, W. P., Smith, M. A., Willman, C. L., Downing, J. R., & Hunger, S. P. (2011). Key pathways are frequently mutated in high-risk

childhood acute lymphoblastic leukemia : A report from the Children's Oncology Group. *Blood*, 118(11), 3080-3087. <https://doi.org/10.1182/blood-2011-03-341412>

Zhang, M. Y., Churpek, J. E., Keel, S. B., Walsh, T., Lee, M. K., Loeb, K. R., Gulsuner, S., Pritchard, C. C., Sanchez-Bonilla, M., Delrow, J. J., Basom, R. S., Forouhar, M., Gyurkocza, B., Schwartz, B. S., Neistadt, B., Marquez, R., Mariani, C. J., Coats, S. A., Hofmann, I., ... Shimamura, A. (2015). Germline ETV6 mutations in familial thrombocytopenia and hematologic malignancy. *Nature Genetics*, 47(2), 180-185. <https://doi.org/10.1038/ng.3177>

Zhou, Y.-H., Wu, X., Tan, F., Shi, Y.-X., Glass, T., Liu, T. J., Wathen, K., Hess, K. R., Gumin, J., Lang, F., & Yung, W. K. A. (2005). PAX6 suppresses growth of human glioblastoma cells. *Journal of Neuro-Oncology*, 71(3), 223-229. <https://doi.org/10.1007/s11060-004-1720-4>

### Références supplémentaires (non répertoriées sur la base PubMed) :

Dessen P and Huret JL. Chromosomal band 9p13. Atlas Genet Cytogenet Oncol Haematol. July 2012. URL : <http://atlasgeneticsoncology.org/Bands/9p13.html>. Atlas of Genetics and Cytogenetics in Oncology and Haematology. Available from: <http://AtlasGeneticsOncology.org>

Greaves M. Cancer: the Evolutionary Legacy. Oxford: Oxford University Press, 2000.

Howlader N, Noone AM, Krapcho M: SEER Cancer Statistics Review (CSR) 1975-2013. Bethesda, Md: National Cancer Institute, 2015. Available online. Last accessed November 25, 2020. Voir sections 28 et 29

Jackson, Graeme. (2019). Evolution and Culture Through Subsistence Strategy.

Ries LAG, Smith MA, Gurney JG, et al. Cancer incidence and survival among children and adolescents: United States SEER Program 1975-1995. Bethesda, MD: National Cancer Institute, SEER Program, 1999.

Annexe : Attestation de suivi de l'enseignement doctoral



## ECOLE DOCTORALE BIOLOGIE – SANTÉ – BIOTECHNOLOGIES

### Commission Pédagogique

DIRECTEUR  
Pr L. MAVEYRAUD

#### ATTESTATION DE SUIVI DE L'ENSEIGNEMENT DOCTORAL

COMMISSION PEDAGOGIQUE  
DR F. LARMINAT

Nom et prénom : LAURA JAMROG

Formation doctorale : Ecole Doctorale Biologie-Santé-Biotechnologies

COMMISSION SCIENTIFIQUE  
DR A. SAUDI

Laboratoire d'accueil : CRCT

#### - Enseignement doctoral et Formation :

- DP02- Lire vite, bien et augmenter sa concentration et sa mémorisation- 21 janvier -28.5 heures
- L'Europe pour les doctorants: de la préparation du diplôme au post-doctorat- 16 mai 2019 - 3.5 heures formation enregistrée
- IP04- Leadership et management- 2 mars 2020 - 25 heures
- EP1 - Stage d'observation tuteuré - 43 heures
- BM05 - Création de souris génétiquement modifiées par le système CRISPR/Cas9- 14 février 2018
- Utilisation et Protection de l'Animal de Laboratoire - Fonction Appicateur - 43 heures
- BM09 - GenomRead- 14 janvier 2019

FILIÈRES

BIOLOGIE STRUCTURALE &  
FONCTIONNELLE  
LAURENT PAQUEREAU

GENES, CELLULES & DEVELOPPEMENT  
GAËLLE LEGUBE

MICROBIOLOGIE  
CLAUDE GUTIEREZ

PHYSIOPATHOLOGIE  
VALERIE PLANAT-BENARD

IMMUNOLOGIE & MALADIES  
INFECTIEUSES  
SYLVIE GUERDER

INNOVATION PHARMACOLOGIQUE  
ANNE BOULOUMIE

CANCEROLOGIE  
CATHERINE MULLER

Je soussignée Florence LARMINAT atteste que :  
**Mme LAURA JAMROG** a pleinement suivi l'enseignement doctoral.

Fait à Toulouse, 3 mars 2021  
La Présidente de la Commission Pédagogique,

Florence LARMINAT



