

École Doctorale des Sciences de la Vie et de la Santé
Institut de génétique et de biologie moléculaire et cellulaire

Thèse présentée par

Marie-Céline DEAU

Soutenue le 26 mars 2021

Pour l'obtention du grade de

Docteur de l'Université de Strasbourg

Sciences du Vivant – Immunologie

**Dissection des spécificités fonctionnelles des
facteurs de transcription de la famille
d'Ikaros au cours du développement précoce
des lymphocytes B**

**Thèse dirigée par
Dr Susan CHAN**

Directeur de recherche, IGBMC

**Rapporteurs externes
Pr Ana CUMANO
Pr Jean-Christophe BORIES**

**Directeur de recherche, Université de Paris
Directeur de recherche, Institut Pasteur**

**Examineur interne
Dr Frédéric COIN**

Directeur de recherche, IGBMC

**Membre invité
Dr Beate HEIZMANN**

Chercheur, IGBMC

*“Victory is not final; failure is not fatal.
It is the courage to continue that counts.”
Winston Churchill*

REMERCIEMENTS

Je remercie les Prs. Ana Cumano et Jean-Christophe Bories ainsi que le Dr. Frédéric Coin d'avoir accepté de juger mon travail.

Je remercie également le Pr. Pauline Soulas et le Dr. Daniel Metzger d'avoir accepté, par trois fois, de faire partie de mon comité de suivi de thèse. Les discussions que nous avons pu avoir m'ont beaucoup aidée au cours de mon travail.

Un immense merci à nos collaborateurs Drs. Simona Sacconi et Dominic Van Essen pour avoir partagé avec moi leurs techniques, leurs conseils et surtout leur anticorps α Venus.

Merci à Philippe et Susan de m'avoir acceptée dans votre laboratoire, de m'avoir confié ce projet passionnément ambitieux et de m'avoir guidée pendant ces 4 années.

Vielen Dank an Beate für deine Unterstützung und deinen Rat in diesen vier Jahren. Nochmals vielen Dank für die Hilfe beim Schreiben dieses Manuskripts!!!

Merci à tous les membres passés et présents du laboratoire. Un merci particulier à Patricia, pas seulement pour toute l'aide quotidienne dans le laboratoire, mais personnellement pour avoir toujours été là, pour m'avoir appris à faire les bredeles et les lamalas, pour le concert Queen Symphonic, pour l'obtention de mon passeport Alsacien en somme, et pour le chat bien sûr !!!!! Un autre immense merci à Céline sans qui cette thèse n'aurait pas été la même. Je venais prendre un peu de sud tous les jours et même si ce n'est pas le même sud, ça fait du bien au cœur et à l'âme!! Merci à Peggy, pour toutes les discussions, scientifiques ou non, les fous rires et pour la découverte de la salle de sport et tout le bien que cela m'a apporté ! Merci à Giovanni, Chiara B et Chiara T, Célestine, Guillaume, Katia, Adina et tous les étudiants qui sont passés par le labo ces dernières années, Léa, Elodie, Vincent, Lauriane, Jessica, pour tous les bons moments partagés, les déjeuners, les cafés, les concours de cuisine, les gâteaux et les rires.

Parce qu'il ne faut pas oublier d'où l'on vient pour savoir où l'on va, je tiens à saluer toutes les personnes qui ont contribué de près ou de loin à mon début de carrière scientifique : Dr. Rozenn Dalbiès-Tran (INRA Nouzilly), M. Eric Palmer et Dr. Alline Paula-Reis (Cryozootech S.A.), Dr. Eric Lippert et Dr. Valérie Prouzet-Mauléon (INSERM U1035) et Dr. Sven Kracker (Institut Imagine). Tout au long de ces années vous m'avez guidée vers ce moment particulier.

Merci encore à mes amis, Marie (le M&M's power jusqu'au bout du monde) & Julien, Marlène, Claire, Louise, Anne-Lise, Aurélie & Claire, Aurélien & Mélody, Sophie, Valentin & Larissa, Jonathan & Marion & Alice & Jules d'être toujours là, tout simplement.

Bien entendu, je voudrais ici prendre le temps de remercier mes parents et toute ma famille pour leur soutien depuis toujours. Je vous dois tout et plus encore. Faudrait qu'j'invente des mots qui existent pas dans le dico pour vous dire à quel point je vous aime.

Merci à Mon Amour pour ton soutien quotidien et ton amour avec lequel la vie est plus douce. Merci d'avoir accepté de quitter ta famille, tes amis et tes collègues pour me suivre dans cette aventure alsacienne. La vie aurait été plus difficile, seule ici. Comme l'a magnifiquement dit Jean d'Ormesson : « Je t'aime dans le temps. Je t'aimerai jusqu'au bout du temps. Et quand le temps sera écoulé, alors, je t'aurai aimé. Et rien de cet amour, comme rien de ce qui a été, ne pourra jamais être effacé ».

Enfin, merci à celui qui me suit comme mon ombre depuis 12 ans. Celui qui s'assoit sur le clavier de l'ordinateur, qui se couche sur mes genoux juste au moment où je dois me lever, à celui qui perd ses poils partout qui court dans tous les sens, qui miaule pendant 30 minutes au moment où je vais me coucher et qui ne supporte pas que je dorme le samedi matin ... Merci, d'apporter tant de mignonnerie dans ma vie.

TABLE DES MATIERES

I.	ABREVIATIONS	12
II.	LISTE DES FIGURES	13
III.	LISTE DES TABLEAUX	14
	INTRODUCTION	16
I.	GENERALITES SUR LES FACTEURS DE TRANSCRIPTION SPECIFIQUES	18
A.	<i>Modèles de liaison à l'ADN</i>	20
B.	<i>Les facteurs de transcription à doigts de zinc C₂H₂</i>	22
1.	Mode de liaison à l'ADN	22
2.	Modèles de prédiction de motifs de liaison.	23
3.	Rôle des linkers dans la liaison à l'ADN	25
II.	LES FACTEURS DE TRANSCRIPTION DE LA FAMILLE D'IKAROS	25
A.	<i>Structure et isoformes</i>	26
1.	La structure générale des facteurs de transcription de la famille d'Ikaros	26
2.	Les différentes isoformes décrites	27
B.	<i>La liaison à l'ADN</i>	30
1.	Rôle du domaine de liaison à l'ADN	30
a)	Implication des doigts de zinc	30
b)	Implication des linkers	31
2.	Implication de l'ADN et motif de liaison	31
C.	<i>Interaction Protéine-Protéine : domaine de dimérisation</i>	32
D.	<i>Mécanismes moléculaires de la régulation de la transcription</i>	34
1.	Association avec NurD	34
a)	Pour l'activation de la transcription	35
b)	Pour la répression de la transcription	35
2.	Association avec le complexe mSin3	36
3.	Association avec SWI/Snf	36
4.	Interaction avec CtBP	37
5.	Association avec PRC2	38
6.	Autres modes de régulations	39
a)	Super-enhancer	39
b)	Hétérochromatine Peri-centromérique (PC-HC)	39
c)	Interaction avec d'autres facteurs de transcription	40
E.	<i>Régulation de l'expression ou de l'activité des facteurs de la famille d'Ikaros</i>	40
1.	Régulation transcriptionnelle	40
2.	Régulation post-traductionnelle	42
a)	Sumoylation	42
b)	Phosphorylation	42
c)	Ubiquitination	43

III.	ROLES DANS LE DEVELOPPEMENT DES LYMPHOCYTES B.....	44
A.	<i>Le développement précoce des lymphocytes B.....</i>	44
1.	De la cellule souche hématopoïétique à l'expression du pré-BCR	44
2.	Du pré-BCR au BCR	47
3.	La voie de signalisation de l'IL7	49
4.	La tolérance des lymphocytes B ou sélection négative	49
5.	La maturation terminale des lymphocytes B.....	50
6.	Les facteurs de transcription de la différenciation des lymphocytes B.....	51
B.	<i>Rôle spécifique d'Ikaros dans le développement des lymphocytes B.....</i>	52
1.	Dans l'engagement vers la lignée lymphoïde.....	53
2.	Dans le développement précoce des lymphocytes B.....	54
3.	Dans la maturation terminale des lymphocytes B.....	54
4.	Le rôle de suppresseur de tumeurs.....	55
a)	Dans les T-ALL	55
b)	Dans les B-ALL	56
5.	Autres pathologies	57
C.	<i>Rôles d'Helios et Aiolos dans le développement B.....</i>	59
1.	Helios.....	59
a)	Dans le développement des lymphocytes B	59
b)	Rôle suppresseur de tumeur.....	59
2.	Aiolos.....	60
a)	Dans le développement des lymphocytes B	60
b)	Dans la maturation des lymphocytes B.....	61
c)	Rôle suppresseur de tumeur.....	62
d)	Hémopathies liées à Aiolos	62
IV.	PROBLEMATIQUE	63
V.	OBJECTIFS	64
	MATERIELS & METHODES	66
I.	CULTURE CELLULAIRE.....	68
A.	<i>J558L-IL7 : Production du surnageant IL-7.....</i>	68
B.	<i>Les cellules BH1.....</i>	68
C.	<i>Les cellules COS-1</i>	68
D.	<i>HEK293T Phoenix Eco</i>	69
II.	CREATION DES LIGNEES CELLULAIRES.....	69
A.	<i>La complémentation bi-moléculaire de fluorescence (BiFC).....</i>	69
B.	<i>Le système tétracycline-ON (tet-ON).....</i>	70
C.	<i>Clonage moléculaire</i>	71
D.	<i>Production rétrovirale.....</i>	72
E.	<i>Transduction des cellules.....</i>	72
F.	<i>Sélection par les antibiotiques.....</i>	72

III.	IMMUNOFLUORESCENCE	73
IV.	IMMUNOPRECIPITATION (IP)	74
V.	IMMOBILISATION D'ANTICORPS SUR BILLES DE SEPHAROSE (ABCAM)	74
VI.	L'IMMUNOPRECIPITATION DE CHROMATINE	76
A.	<i>La préparation des cellules</i>	76
B.	<i>La sonication</i>	78
C.	<i>L'immunoprécipitation</i>	78
D.	<i>La purification de l'ADN</i>	79
E.	<i>Le contrôle par qPCR</i>	80
F.	<i>La préparation de la librairie et séquençage</i>	81
G.	<i>Analyse bioinformatique</i>	81
VII.	L'OMNI-ATAC (ASSAY FOR TRANSPOSASE-ACCESSIBLE CHROMATIN)	81
A.	<i>Le protocole</i>	81
B.	<i>L'analyse bioinformatique</i>	84
VIII.	ANALYSE DE LA TRANSCRIPTION PAR RNA-SEQ	84
A.	<i>Le protocole</i>	84
B.	<i>Analyses bioinformatiques</i>	85
IX.	ISOLATION ET MARQUAGE DES CELLULES PRIMAIRES	85
A.	<i>Isolation des cellules de la moelle osseuse</i>	85
B.	<i>Isolation des cellules de la rate</i>	86
C.	<i>Enrichissement sur colonne</i>	86
1.	<i>Vérification de pureté</i>	86
X.	MARQUAGE INTRACELLULAIRE DES FACTEURS DE TRANSCRIPTION DE LA FAMILLE D'IKAROS DANS LES CELLULES PRIMAIRES ET LES CELLULES BH1	87
A.	<i>Dans les cellules primaires de la moelle osseuse</i>	87
B.	<i>Dans les cellules BH1</i>	87
	RESULTATS	90
I.	LA PREUVE DE CONCEPT	93
A.	<i>Quelle position pour les fragments V1 et V2 : N-terminale ou C-terminale ?</i>	93
B.	<i>Création de la lignée BH1-Ikaros-V1/Ikaros-V2</i>	94
1.	<i>La lignée BH1-rtTA</i>	94
2.	<i>Première stratégie pour l'obtention de la lignée BH1-Ikaros-V1/ikaros-V2</i>	95
3.	<i>Deuxième stratégie pour l'obtention de la lignée BH1-Ikaros-V1/ikaros-V2</i>	95
C.	<i>La présence de V1 et V2 n'empêche pas la localisation normale d'IKAROS dans les cellules</i>	97
D.	<i>V1 et V2 ne perturbent pas la fonctionnalité d'Ikaros dans les cellules BH1</i>	98
E.	<i>L'anticorps αVenus permet l'immunoprécipitation spécifique des dimères</i>	99
F.	<i>V1 et V2 ne force pas la dimérisation des facteurs de transcription</i>	100
II.	ÉTUDE DE LA LIAISON A L'ADN DES FACTEURS DE TRANSCRIPTION DE LA FAMILLE D'IKAROS	102
A.	<i>Aiolos a des spécificités de liaison à l'ADN non partagées avec Ikaros et Helios</i>	102

1.	Le profil de liaison à l'ADN d'Aiolos diffère de ceux d'Ikaros et d'Helios	102
2.	Les gènes spécifiques des lymphocytes B sont des cibles privilégiées d'Aiolos	104
3.	La spécificité de liaison d'Aiolos n'est pas due à sa surexpression.....	105
4.	Les protéines Aiolos et Helios endogènes se colocalisent avec les protéines surexprimées	107
5.	Aiolos se fixe à ses séquences cibles sous forme d'homodimère et d'hétérodimère avec Ikaros	108
B.	<i>La spécificité de liaison à l'ADN d'Aiolos est déterminé par sa séquence protéique et son motif de liaison</i>	<i>112</i>
1.	Le domaine de liaison à l'ADN ne suffit pas à transmettre la spécificité de liaison	112
2.	Le motif de liaison à l'ADN d'Aiolos diffère de celui d'Ikaros	113
III.	ÉTUDE DE LA FONCTIONNALITE DES DIMERES DE LA FAMILLE D'IKAROS.....	117
A.	<i>Comparaison de l'expression d'Ikaros et d'Aiolos dans le système tet-ON avec les cellules primaires</i>	<i>117</i>
1.	Les facteurs de transcription sont fortement surexprimés	117
2.	Adaptation de la dose de doxycycline	117
3.	Cinétique de formation des dimères.....	118
B.	<i>L'expression d'Ikaros et d'Aiolos a un effet d'ouverture sur la chromatine.....</i>	<i>120</i>
C.	<i>Aiolos régule l'expression des gènes spécifiques des lymphocytes B.....</i>	<i>123</i>
D.	<i>Les cellules exprimant Aiolos se différencient plus vite</i>	<i>127</i>
E.	<i>La co-expression d'Ikaros et d'Aiolos est plus efficace dans la régulation de l'expression des gènes spécifiques des lymphocytes B.....</i>	<i>130</i>
	DISCUSSION	134
I.	LES PROPRIETES DE LIAISON A L'ADN DES FACTEURS DE TRANSCRIPTION DE LA FAMILLE D'IKAROS.....	136
II.	LA FONCTIONNALITE DES FACTEURS DE TRANSCRIPTION DE LA FAMILLE D'IKAROS	140
III.	L'HETERODIMERE IKAROS/AIOLOS EST-IL LA CLE DU MYSTERE ?	143
	REFERENCES BIBLIOGRAPHIQUES	146
	ANNEXES	166

I. Abréviations

ADN : Acide DesoxyriboNucléique	PCHC : Hétérochromatine Pericentromérique
ARN : Acie RiboNucléique	PIC : complexe de pré-initiation de la transcription
B-ALL : B-Acute Lymphoblastique Leukemia	RAG : Recombinatin Activating Gene
BCR : B-Cell Receptor	SH : hypermutation somatique
BAFF : B cell Activating Factor	STAT : Signal Transducers and Activators of Transcription
BiFC : Bi-molecular Fluorecent Complementation	SLE : Lupus systémique érythémateux
Cter : carboxy-terminal	SWI/Snf : SWItch/Sucrose Non-Fermentable
CDK9 : Cycline Dependant Kinase 9	TCR : Récepteur des lymphocytes T
CtBP : C-terminal Binding Protein	TSS : Transcription start site
CtIP : CtBP Interacting Protein	V1: partie N-terminale de Venus
CLP : Common Lymphoid Progennitor	V2 : partie C-terminale de Venus
CSR : Class Switch recombination	
DZ : Doigt de Zinc	
FO-B : Follicular B cell	
HDAC : Histone déacétylase	
HSC : Hematopoietic Stem Cell	
Ig : Immunoglobuline	
IL7 : Interleukine 7	
JAK : Janus Kinase	
LMPP : Lymphoid-primed multipotent Progenitor	
MZ-B : Lymphocytes B de la zone Marginal	
MPP : Multipotent Progenitor	
Nter : Amino-terminal	
NurD : Nucleosome Remodeling Complex	
PRC2 : Polycomb Repressiv Complex 2	
P-TEFb: Positive Transcription Elongation Factor	

II. Liste des Figures

FIGURE I.1 : LES MODELES DE DIMERISATION DES FACTEURS DE TRANSCRIPTION (KIM AND LITTLE, 1992)	21
FIGURE I.2 : FIXATION DES DIMERES COMPOSES DE P50 ET P65	21
FIGURE I.3 : STRUCTURE D'UN DOIGT DE ZINC C ₂ H ₂ ET LIAISON A L'ADN	23
FIGURE I.4 : PREDICTION DE LA RECONNAISSANCE DOIGTS DE ZINC C ₂ H ₂ /ADN	24
FIGURE I.5 : PREDICTION DE LA RECONNAISSANCE DOIGTS DE ZINC C ₂ H ₂ /ADN	24
FIGURE I.6 : REPRESENTATION DES FACTEURS DE TRANSCRIPTION DE LA FAMILLE D'IKAROS	27
FIGURE I.7 : LES DIFFERENTES ISOFORMES DECRITES POUR IKAROS, HELIOS ET AIOLOS	29
FIGURE I.8 : MOTIFS DE LIAISON A L'ADN POUR LES FACTEURS DE TRANSCRIPTION DE LA FAMILLE D'IKAROS.	32
FIGURE I.9 : SCHEMA DES DIMERES FORMES PAR IKAROS	34
FIGURE I.9 : ASSOCIATION D'IKAROS AVEC NURD ET P-TEFB	35
FIGURE I.10 : L'ASSOCIATION D'IKAROS AVEC LE COMPLEXE SWI/SNF	37
FIGURE I.11 : ASSOCIATION D'IKAROS AVEC CtBP ET CtIP	37
FIGURE I.12 : LES SITES D'INTERACTION D'IKAROS AVEC SES PARTENAIRES PROTEIQUES	38
FIGURE I.13 : MODE DE REPRESSION DE LA TRANSCRIPTION VIA PRC2	38
FIGURE I.14: IKAROS SE FIXE AUX SUPER-ENHANCERS	39
FIGURE I.15 : PROCESSUS DE POLY-UBIQUITINATION D'IKAROS ET D'AIOLOS PAR CRL4 ^{CRBN} EN PRESENCE DE LENALIDOMIDE OU POMALIDOMIDE	43
FIGURE I.16 : COMPARTIMENTS HEMATOPOÏETIQUES	45
FIGURE I.17 : LE DEVELOPPEMENT PRECOCE DES LYMPHOCYTES B	46
FIGURE I.18 : LES VOIES DE SIGNALISATION INDUITES PAR LE PRE-BCR.....	48
FIGURE I.19 : REGULATION TRANSCRIPTIONNEL DU DEVELOPPEMENT PRECOCE DES LYMPHOCYTES B	51
FIGURE I.20 : EXPRESSION D'IKAROS, HELIOS ET AIOLOS AU COURS DU DEVELOPPEMENT PRECOCE DES LYMPHOCYTES B.	52
FIGURE I.21: RESUME DES PHENOTYPES OBSERVES DANS LES DIFFERENTES LIGNEES DE SOURIS PORTANT UNE MUTATION AU LOCUS D' <i>IKZF1</i>	58
FIGURE I.IV.1 REPRESENTATION THEORIQUE DE LA DISTRIBUTION DES DIMERES DE LA FAMILLE D'IKAROS AU COURS DU DEVELOPPEMENT PRECOCE DES LYMPHOCYTES B.	64
FIGURE MII.1LA COMPLEMENTATION BI-MOLECULAIRE DE FLUORESCENCE	70
FIGURE MII.2 LE SYSTEME TETRACYCLINE-ON (TET-ON)	70
FIGURE M.X.1 ANALYSE DES DIFFERENTES FRACTIONS DE LA MOELLE OSSEUSE	87
FIGURE RI.1 POSITIONNEMENT DES FRAGMENTS V1 ET V2 EN N- OU C-TERMINAL	94
FIGURE RI.2 PREMIERE STRATEGIE POUR L'EXPRESSION DES TRANSGENES LIES AUX FRAGMENTS V1 ET V2	95
FIGURE RI.3 EXPRESSION DE LA FLUORESCENCE DE VENUS PAR COMPLEMENTATION V1+V2	96
FIGURE RI.4 REMPLACEMENT DE L'HYGROMYCINE PAR LA BLASTICIDINE	97
FIGURE RI.5 VISUALISATION DE LA LOCALISATION D'IKAROS PAR IMMUNOFLUORESCENCE.....	98
FIGURE RI.6 DIFFERENCIATION DES CELLULES BH1-IKAROSV1/IKAROSV2	99

FIGURE RI.7 L'ANTICORPS α VENUS PERMET L'IMMUNOPRECIPITATION DES DIMERES V1+V2	100
FIGURE RI.8 V1 ET V2 NE FORCENT PAS LA DIMERISATION	101
FIGURE RII.1 ANALYSE DE LA LIAISON A L'ADN D'IKAROS, HELIOS ET AIOLOS DANS LES CELLULES BH1	103
FIGURE RII.2 LES GENES SPECIFIQUES DES LYMPHOCYTES B SONT DES CIBLES PRIVILEGIEES D'AIOLOS	104
FIGURE RII.3 LES SPECIFICITES DE LIAISON A L'ADN D'AIOLOS NE SONT PAS DUE A LA SUREXPRESSION	106
FIGURE RII.4 LES PROTEINES ENDOGENES HELIOS ET AIOLOS ONT DES SITES DE LIAISON A L'ADN COMMUNS AVEC LES PROTEINES SUREXPRIMEES.	108
FIGURE RII.5 L'HOMODIMERE D'IKAROS NE SE LIE PAS EFFICACEMENT A L'ADN.....	109
FIGURE RII.6 LES GENES SPECIFIQUES DES LYMPHOCYTES B SONT PRINCIPALEMENT CIBLES PAR AIOLOS	111
FIGURE RII.7 APPORT DU DOMAINE DE LIAISON DANS LA SPECIFICITE D'AIOLOS	112
FIGURE II.8 LE MOTIF DE LIAISON D'AIOLOS DIFFERE DE CELUI D'IKAROS	114
FIGURE II.9 ENRICHISSEMENT DES COMBINAISONS DE NUCLEOTIDES ENTOURANT LE MOTIF GGAA DANS LES PICS SPECIFIQUES AIOLOS ET LES PICS COMMUNS IKAROS/AIOLOS	116
FIGURE RIII.1 ADAPTATION DE LA DOSE DE DOXYCYCLINE POUR UNE EXPRESSION PLUS PHYSIOLOGIQUE	119
FIGURE RIII.2 LA FIXATION DES FACTEURS DE TRANSCRIPTION IKAROS ET AIOLOS FAVORISE L'OUVERTURE DE LA CHROMATINE.....	121
FIGURE RIII.3 LES GENES SPECIFIQUES DES LYMPHOCYTES B DONT DIRECTEMENT REGULES PAR AIOLOS.....	124
FIGURE RIII.4 L'EXPRESSION D'AIOLOS PERMET AUX CELLULES DE SE DIFFERENCIER PLUS RAPIDEMENT	128
FIGURE RIII.5 LA CO-EXPRESSION D'IKAROS ET D'AIOLOS PERMET UNE REGULATION PLUS FORTE DE LA TRANSCRIPTION.....	131
FIGURE D.I.1 COMPARAISON DES MOTIFS DE LIAISON A L'ADN DE LA PROTEINE AIOLOS ENDOGENE OU SUREXPRIMEE	137
FIGURE D.I.2 ALIGNEMENT D'IKAROS, HELIOS ET AIOLOS ET ANALYSE DES DOMAINES DESORDONNES	139

III. Liste des tableaux

TABLEAU I.I-1 LES CLASSES DE FACTEURS DE TRANSCRIPTION	19
TABLEAU I.IV-1 COMPARAISON DE LA COMPOSITION DES DOIGTS DE ZINC ET PREDICTION DE MOTIFS DE LIAISON.....	30
TABLEAU I.IV-2 COMPARAISON DES LINKERS ENTRE CHAQUE DOIGT DE ZINC ADJACENT	31
TABLEAU M.II-1 AMORCES POUR LE CLONAGE IKAROS-V1 ET IKAROS-V2.....	71
TABLEAU M.II-2 AMORCES POUR LE CLONAGE DE V1-IKAROS ET V2-IKAROS.....	71
TABLEAU M.V-1 COMPATIBILITE ANTICORPS/PROTEIN A OR G BEADS (ADAPTE DE DYNABEADS)	76
TABLEAU M.VI-1 AMORCES POUR LA PCR Q CISH	80
TABLEAU M.VII-1 AMORCES D'ATAC-SEQ.....	84
TABLEAU M.X-1 LISTE DES ANTICORPS UTILISES	88

Introduction

La régulation de l'expression génique se fait grâce aux facteurs de transcription. Chez les eucaryotes supérieurs il en existe deux classes : les facteurs de transcription généraux qui s'associent à l'ARN polymérase II dans le complexe de pré-initiation de la transcription (PIC) et les facteurs de transcription spécifiques permettant une régulation spatiotemporelle de la transcription de façon fine et précise.

I. Généralités sur les facteurs de transcription spécifiques

Un facteur de transcription est une protéine possédant un domaine de liaison à l'ADN et un domaine d'activation et/ou de répression. En se liant au promoteur de ses gènes cibles ou à des sites proximaux, le facteur de transcription permet le recrutement spécifique du PIC et l'activation de la transcription. Mais il peut également en réprimer l'expression en empêchant le PIC de se lier sur le promoteur. Il a alors un rôle direct ON/OFF sur l'expression de ses gènes cibles. Cependant ces facteurs peuvent également agir comme modulateur de transcription en se liant à des régions plus distales appelés *enhancers* (ou amplificateurs) et *silencers* (ou « inactivateurs »).

Les facteurs de transcription peuvent être ubiquitaires, c'est-à-dire exprimés dans tous les tissus de l'organisme, ou spécifiques, par exemple seulement dans le système hématopoïétique. Là encore ils peuvent être présents dans toutes les cellules de ce tissu ou seulement dans certains types cellulaires. Certains facteurs sont exprimés temporairement, c'est-à-dire seulement à un moment précis du développement cellulaire et d'autres.

Ces facteurs de transcription sont répartis en différentes superclasses suivant la nature de leur domaine de liaison à l'ADN (Tableau 1).

Tableau I.I-1 Les classes de facteurs de transcription

Superclasse 1 : Domaines basiques

- . Classe 1.1 : Leucine zipper (bZIP)
- . Classe 1.2 : Hélice-boucle-hélice (bHLH)
- . Classe 1.3 : Hélice-boucle-hélice/ leucine zipper (bHLH/bZIP)
- . Classe 1.4 : NF-1
- . Classe 1.5 : RF-X
- . Classe 1.6 : bHSH

Superclasse 2 : Domaines de liaison à l'ADN coordonnés par le zinc

- . Classe 2.1 : doigts de zinc Cys4 des récepteurs nucléaires
- . Classe 2.2 : Doigts de zinc Cys4
- . Classe 2.3 : doigts de zinc de type Cys2His2
- . Classe 2.4 : Cys6 cystéine-zinc cluster
- . Classe 2.5 : doigts de zinc de composition alternative

Superclasse 3 : Hélice-coude-Hélice (HTH)

- . Classe 3.1 : Homéodomaine
- . Classe 3.2 : Paired box
- . Classe 3.3 : Fork Head/ winged helix
- . Classe 3.4 : Heatshock factors
- . Classe 3.5 : Tryptophan cluster
- . Classe 3.6 : TEA domain

Superclasse 4 : beta-scaffold factors with minor groove contacts

- . Classe 4.1 : Rel homology region (RHR)
- . Classe 4.2 : STAT
- . Classe 4.3 : p53
- . Classe 4.4 : MADS box
- . Classe 4.5 : beta barrel alpha helix TF
- . Classe 4.6 : TATA-binding protein
- . Classe 4.7 : HMG
- . Classe 4.8 : heteromeric CCAAT factors
- . Classe 4.9 : grainyhead
- . Classe 4.10 : Cold-shock domain factors
- . Classe 4.11 : Runt

Superclasse 0 : les autres facteurs de transcription

- . Classe 0.1 : poing de cuivre
- . Classe 0.2 : HMG(Y)
- . Classe 0.3 : Pocket domain
- . Classe 0.4 : E1A-like factors
- . Classe 0.5 : AP2/EREBP-related factors

La reconnaissance d'un site de liaison particulier sur l'ADN assure la spécificité primaire du facteur de transcription pour son ensemble de gènes cibles mais les interactions avec les bases adjacentes au motif de liaison sont également importantes. La dimérisation des facteurs est fréquente avec des préférences d'orientation et d'espacement entre les motifs (Jolma et al., 2013). Il ressort également de cette étude que les facteurs de transcription à doigts de zinc de type C₂H₂ représentent une famille très hétérogène dans laquelle quasiment chaque facteurs suit un modèle de liaison à l'ADN unique. En effet, les auteurs ont pu décrire 42 modèles différents pour 53 facteurs de transcription dans cette famille alors que pour la famille des facteurs de transcription à homéodomains, 53 modèles sont utilisés par 146 facteurs.

A. Modèles de liaison à l'ADN

Certains facteurs de transcription se lient à l'ADN comme monomères. Ainsi les facteurs de transcription de la famille ETS (Erythroblast Transformation Specific), bien qu'ils possèdent un domaine de dimérisation, ne se lient à l'ADN que sous forme monomérique. De plus, la dimérisation semble être un mécanisme d'inactivation en masquant leur domaine de transactivation (Carrère et al., 1998; Graves and Petersen, 1998; Nye et al., 1992).

D'autres facteurs forment des dimères voire de plus grands oligomères pour assurer leurs fonctions. En 1992, Kim et Little ont décrit deux modèles de dimérisation des facteurs de transcription (Figure I.1). Dans le modèle « dimère », les facteurs de transcription se dimérisent en solution puis se lient à l'ADN. Dans le modèle « monomère », un premier facteur se lie seul à l'ADN puis le deuxième facteur vient soit d'abord se dimériser puis se lier à l'ADN soit l'inverse. En conclusion de cette étude, il y aurait, pour tout facteur de transcription qui se dimérise, un équilibre entre les deux modèles avec une préférence pour l'un des deux (Kim and Little, 1992).

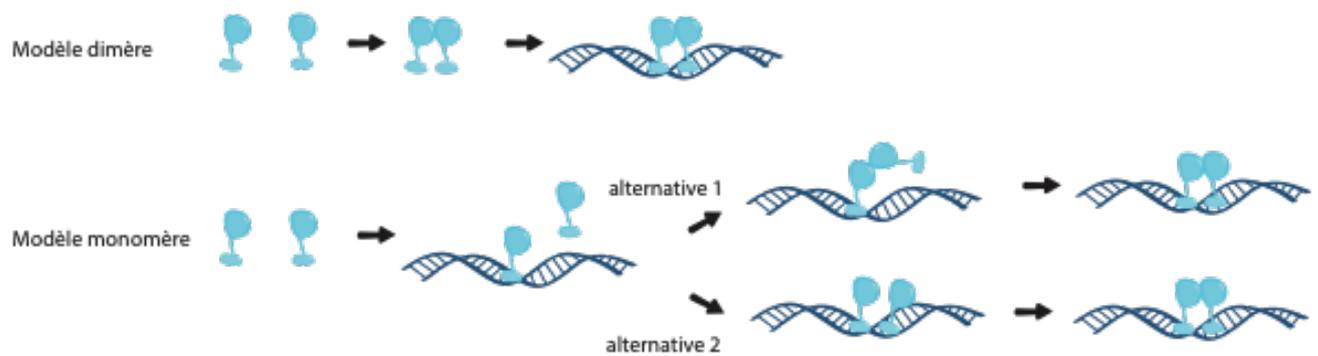


Figure I.1 : Les modèles de dimérisation des facteurs de transcription (Kim and Little, 1992)
 Le modèle dimère propose une dimérisation des facteurs de transcription en solution puis une liaison du dimère à l'ADN. Le modèle monomère propose la liaison à l'ADN d'un monomère puis deux alternatives pour la dimérisation. Soit le second facteur se dimérise au monomère déjà lié puis se lie à l'ADN, soit le second facteur se lie à l'ADN puis dimérise.

Les facteurs de transcription p65 et p50 de la famille Nf- κ B peuvent former des homodimères et des hétérodimères (Figure I.2). Les sites spécifiques des différents dimères sont, en fait, les sites de fixation de chaque monomère, associés en palindrome, pour l'homodimère p50/p50 ou séparés par un nucléotide, pour l'homodimère p65/p65 et l'hétérodimère p50/p65 (Funnell and Crossley, 2012). Chaque dimère est donc dans la capacité de réguler un groupe de gènes cibles qui lui est propre.

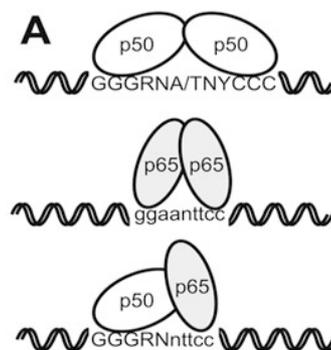


Figure I.2 : Fixation des dimères composés de p50 et p65
 Le motif reconnu par p50 est GGGRNA où R représente une purine (A ou G). Le motif de p65 est GGAA.

Les facteurs cFos et cJun sont des protéines à glissière de leucine permettant leur dimérisation. Lorsqu'ils s'hétérodimérisent ces facteurs forment le complexe AP-1 qui agit comme facteur de transcription. En 2001, Kohler et Schepartz ont montré que le modèle utilisé par l'hétérodimère cFos/cJun passe par une fixation de cJun à l'ADN puis de cFos, la dimérisation se faisant pendant la liaison à l'ADN. De plus, les auteurs ont montré que l'homodimère cJun/cJun existe et est fonctionnel alors que cFos/cFos est trop instable pour être fonctionnel (Kohler and Schepartz, 2001).

Les profils de liaison à l'ADN de deux monomères d'une même famille de facteurs, ici RXR α et PPAR γ , ne sont pas représentatifs de la liaison de l'hétérodimère. De plus, alors que *in vivo*, RXR α est beaucoup plus présent que PPAR γ , il semblerait que l'hétérodimérisation soit favorisée et que la constante d'association de l'hétérodimère à l'ADN soit bien plus forte que celle des monomères (Isakova et al., 2016).

Toutes les études sur la dimérisation des facteurs de transcription montrent que la dimérisation renforce l'affinité des facteurs pour l'ADN et stabilise la liaison. De plus, certaines études mettent en exergue des différences fonctionnelles entre les homodimères et les hétérodimères.

B. Les facteurs de transcription à doigts de zinc C₂H₂

Chez les eucaryotes les facteurs de transcription les plus représentés sont ceux de la classe 2.3, à savoir les facteurs de transcription à doigts de zinc de type Cys₂His₂ ou C₂H₂.

Il existe quatre catégories de facteurs de transcription à doigts de zinc de type C₂H₂ suivant le nombre de doigts de zinc qui composent leur domaine de liaison à l'ADN. Tout d'abord il y a la catégorie des « single » qui ne possèdent qu'un seul DZ nécessitant un domaine spécifique supplémentaire « non doigt de zinc » pour soutenir la liaison à l'ADN. Il y a ensuite la catégorie des « triple » composés de trois doigts de zinc, les « multi-adjacent » qui possèdent quatre ou plus DZ concomitants et enfin les « paires séparées » qui présentent plus de trois DZ couplés deux à deux. Dans le cas de domaine de liaison à l'ADN possédant plusieurs DZ, Les différents DZ adjacents sont séparés par un « Linker ».

1. Mode de liaison à l'ADN

Un doigt de zinc est composé de deux feuillets β et d'une hélice α (Figure I.3). C'est principalement grâce à cette hélice que la liaison à l'ADN peut s'effectuer. En effet, les acides aminés impliqués dans la liaison sont en position -1, +2, +3 et +6 de l'hélice α . Chaque doigt de zinc est ainsi capable de reconnaître un motif de trois nucléotides successifs en *cis* grâce aux acides aminés +6, +3 et -1 et un nucléotide en *trans*, reconnu par l'acide aminé en position +2. (Hossain et al., 2015).

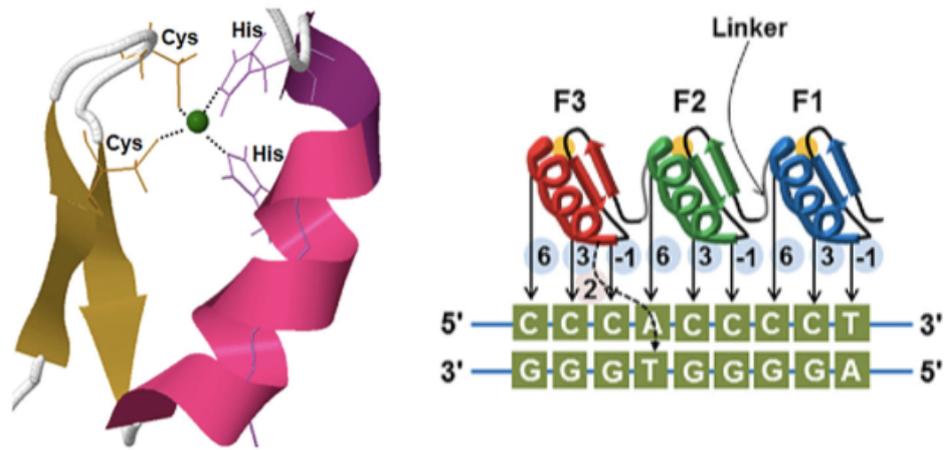


Figure 1.3 : Structure d'un doigt de zinc C_2H_2 et liaison à l'ADN (Hossain et al., 2015)

2. Modèles de prédiction de motifs de liaison.

Alors que la structure des doigts de zinc de type C_2H_2 ainsi que leur mode de liaison à l'ADN sont connus, leurs motifs de liaison ne sont pas toujours bien définis.

Plusieurs équipes se sont heurtées au problème de la spécificité d'interaction entre doigts de zinc et ADN. C'est pourquoi différentes études de prédiction de motifs ou d'analyse à grande échelle des facteurs de transcription à doigts de zinc ont été menées ces dernières années. En utilisant un algorithme d'estimation statistique basé sur la combinaison de données structurales et des exemples de motifs de liaison pour les facteurs de transcription d'une même famille, Kaplan *et al.* ont pu déterminer des préférences de liaison pour chaque acides aminés du doigt de zinc interagissant avec l'ADN (Kaplan et al., 2005). Il semble que l'adénosine soit le nucléotide le moins enrichi dans les domaines de liaison à l'ADN des facteurs de transcription à doigt de zinc C_2H_2 chez la drosophile (Figure RI.4).

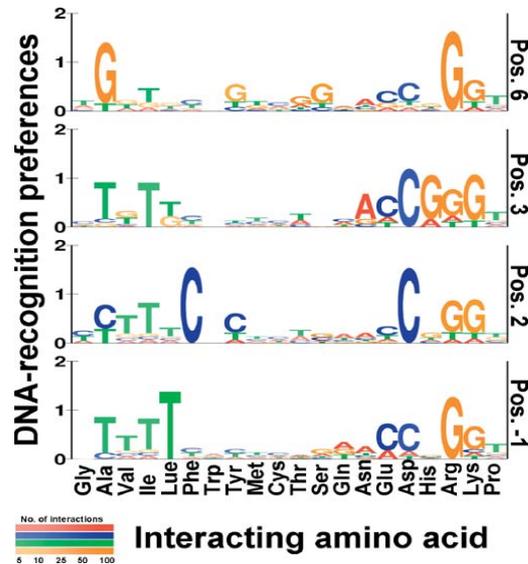


Figure I.4 : Prédiction de la reconnaissance doigts de zinc C₂H₂/ADN (Kaplan et al., 2005)

En utilisant comme modèle d'association des doigts de zinc de la protéine Zif268, les chercheurs ont testé les 20 acides aminés aux positions -1, 1, 2, 3, 5 et 6 de l'hélice α des DZ2 et DZ3. Au regard des 64 possibilités de combinaisons en triplet des 4 bases A,T,C et G, Persikov et collaborateurs ont pu déterminer les préférences de liaisons qui sont résumées dans la figure I.5 (Persikov et al., 2015).

Ce qui ressort de ces différentes études est que la préférence de liaison entre acide aminé et base nucléotidique n'est pas déterminée entièrement par l'acide aminé mais surtout par le contexte d'acides aminés. Il est ainsi très difficile de prédire la préférence de liaison et donc un motif en se basant uniquement sur les bases en position -1,3 et 6.

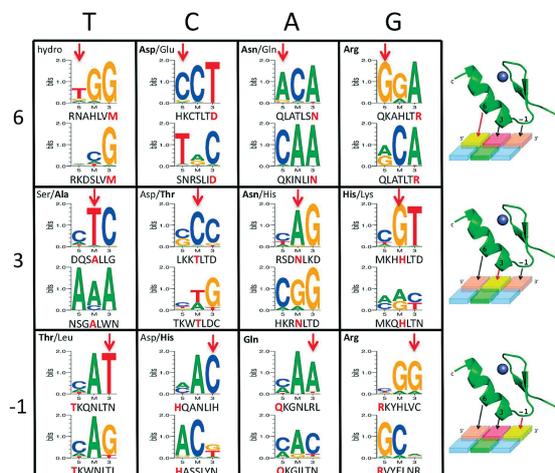


Figure I.5 : Prédiction de la reconnaissance doigts de zinc C₂H₂/ADN

3. Rôle des linkers dans la liaison à l'ADN

Entre chaque doigt de zinc existe une séquence d'acides aminés appelée « linker ». C'est une séquence très conservée parmi les protéines à doigts de zinc de type C_2H_2 dont le consensus est : **HTGEKPY/FXC**. Différentes études ont montré que des mutations ou des modifications, telle que la phosphorylation, dans cette séquence, entraînent une perte d'affinité pour l'ADN. En effet, le remplacement de la glycine par une proline réduit drastiquement la capacité de liaison à l'ADN, un effet plus atténué est observé après le remplacement de la thréonine, de la lysine et de la proline. La lysine permet de stabiliser la liaison à l'ADN en interagissant directement avec les groupements phosphates (Choo and Klug, 1993; Elrod-Erickson et al., 1996). La phosphorylation de la thréonine réduit la capacité de liaison à l'ADN (Jantz and Berg, 2004). Il a été proposé que cette modification fasse partie d'un mécanisme de régulation de la transcription au cours du cycle cellulaire en inactivant de façon coordonnée les facteurs de transcription à doigts de zinc comprenant certaines protéines du complexe de pré-initiation de la transcription également composées de doigts de zinc multi-adjacents de type C_2H_2 tels que TFIID ou TFIIH.

Le rôle crucial de la séquence du linker a été montré par le remplacement d'un linker par un autre, présent dans la même protéine ou dans une protéine différente, ce qui a aboli totalement la liaison à l'ADN. En revanche l'ajout d'acides aminés dans la séquence a pour effet d'augmenter l'affinité de la liaison (Choo and Klug, 1993).

II. Les facteurs de transcription de la famille d'IKAROS

Dans notre laboratoire, nous souhaitons comprendre les mécanismes de la régulation génique impliqués dans l'hématopoïèse. Pour cela, nous étudions les fonctions d'une famille de facteurs de transcription dont les membres sont exprimés dans tous les compartiments hématopoïétiques. En effet, différents laboratoires, dont le nôtre, ont démontré que les facteurs de transcription de la famille d'Ikaros sont impliqués dans le programme de régulation génique assurant le bon déroulement de l'hématopoïèse ainsi que la fonctionnalité des cellules matures. De plus, des mutations perte de fonction ou gain de fonction de ces facteurs de transcription sont retrouvés dans différents cancers hématopoïétiques ou dans les lupus ce qui fait de ces facteurs des suppresseurs de tumeurs.

Les facteurs de transcription de la famille d'Ikaros sont au nombre de cinq dont Ikaros est le membre fondateur mais aussi le plus étudié. Ikaros, Helios, Aiolos, Eos et Pegasus, respectivement codés par les gènes *Ikzf1-5*, sont des facteurs de transcription à doigt de zinc (DZ) de la classe 2.3.

Dans ce travail de thèse, je me suis concentrée sur trois de ces facteurs : Ikaros, Helios et Aiolos. Ces trois protéines ont une structure très conservée entre-elles avec trois domaines principaux et ont, chacune, plusieurs isoformes résultantes d'épissage alternatifs.

A. Structure et isoformes

1. La structure générale des facteurs de transcription de la famille d'Ikaros

La structure de ces facteurs de transcription est très conservée au cours de l'évolution (John and Ward, 2011; Molnár et al., 1996). Leur domaine de liaison à l'ADN, situé en partie amino-terminale (Nter) est composé de quatre doigts de zinc (DZ1 à 4) de type Cys₂His₂ : (F/Y)-X-C-X₂₋₅-C-X₃-(F/Y)-X_{5-ψ}-X₂-H-X₃₋₅-H. Les deux doigts de zinc centraux DZ2 et DZ3 permettent la reconnaissance et la liaison d'un motif central commun : GGAA (Cobb et al., 2000; Molnár and Georgopoulos, 1994). Comme nous l'avons vu précédemment, les facteurs de transcription à doigts de zinc multi-adjacents n'ont besoin que de trois DZ pour la liaison à l'ADN. Dans le cas d'Ikaros, il semblerait que les DZ1 et DZ4 soit alternativement utilisés. Ce concept a été approché par les études de Schjerven et collaborateurs qui ont montré que les DZ1 et DZ4 d'IKAROS lui confèrent des rôles spécifiques, respectivement dans la différenciation des lymphocytes B et sa fonction de suppresseur de tumeur (Schjerven et al., 2013, 2017). Cependant, dans ces expériences de délétion des DZ1 ou DZ4 il reste des gènes cibles communs entre les deux mutants étudiés ce qui ne correspond pas entièrement à une utilisation alternative du troisième DZ en fonction des gènes cibles.

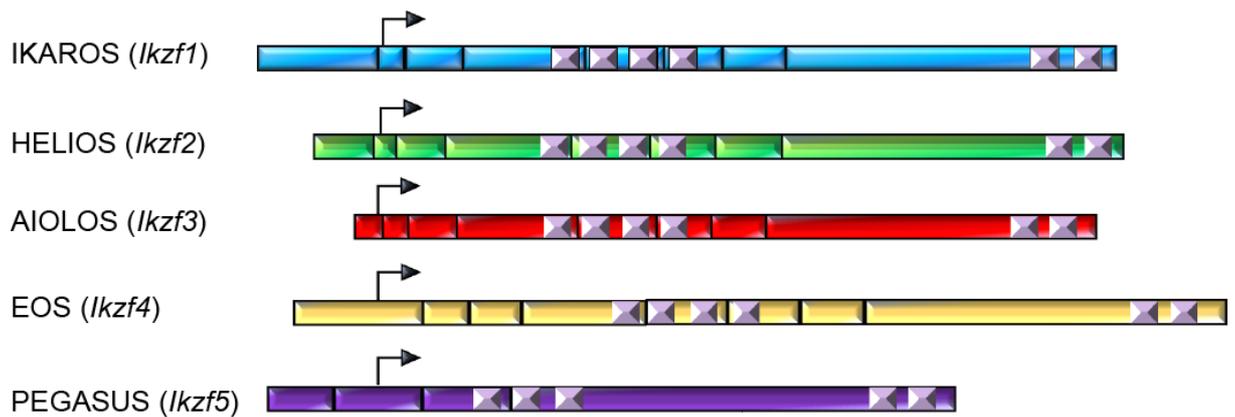


Figure I.6 : Représentation des facteurs de transcription de la famille d'Ikaros
 La flèche représente la position de l'ATG. Les carrés violets représentent la position des doigts de zinc.

La partie carboxy-terminale (Cter) comporte un domaine composé de deux doigts de zinc (DZ5 et DZ6) permettant la dimérisation de la protéine avec elle-même ou avec un autre membre de la famille. Différentes études ont montré la présence des différents membres de la famille dans les mêmes complexes protéiques grâce à des expériences d'immunoprécipitation (Hahm et al., 1998; Kelley et al., 1998; Morgan et al., 1997; Sun et al., 1996).

Enfin, entre les deux domaines à DZ se trouve un domaine bipartite d'activation/répression. Il est composé par la succession d'une série d'acides aminés acides et d'une série d'acides aminés hydrophobes. (Georgopoulos et al., 1997; Sun et al., 1996). Ce domaine est très conservé entre les facteurs de la famille d'Ikaros (Kelley et al., 1998; Koipally et al., 2002).

2. Les différentes isoformes décrites

Chacune de ces protéines existe sous différentes isoformes qui sont les résultats d'épissage alternatifs (Caballero et al., 2007; Hahm et al., 1994; Nakase et al., 2002; Rebollo and Schmitt, 2003). Ils affectent principalement les exons 3 à 5 comportant les séquences codantes du domaine de liaison à l'ADN (Figure I.6). Certaines isoformes sont alors dépourvues de tout ou d'une partie du domaine de liaison à l'ADN. Ainsi les isoformes comportant au moins trois doigts de zinc peuvent efficacement lier l'ADN. En revanche les isoformes dont le domaine de liaison à l'ADN est totalement dégénéré et non fonctionnel peuvent agir comme mutant dominant négatif et sont retrouvés dans certaines cellules leucémiques. Ces protéines possèdent toujours le domaine de dimérisation qui leur permet d'interagir avec la forme WT et ainsi former des dimères non fonctionnels. Par exemple IK-6 est fréquemment

trouvé dans les leucémies aiguë lymphoblastiques (ALL) B (B-ALL) et l'isoforme Hel-5 a été détectée dans les T-ALL (Nakase et al., 2002; Takanashi et al., 2002). Dans les cellules murines saines, les isoformes d'IKAROS les plus abondantes sont Ik-1 et Ik-2, He-1 et He-2 pour HELIOS et Aio-1 pour AIOLOS. Il existe une isoforme d'IKAROS : IK-H ou IK1+ uniquement présente chez l'homme présentant 20 acides aminés supplémentaires entre les exons 2 et 3 (Li et al., 2011).

Dans ce travail je me suis principalement intéressée à la capacité de liaison à l'ADN de ces facteurs de transcription et à leur capacité de dimérisation. C'est pourquoi, les paragraphes suivants présentent plus en détails les connaissances acquises sur les domaines de liaison à l'ADN et de dimérisation.

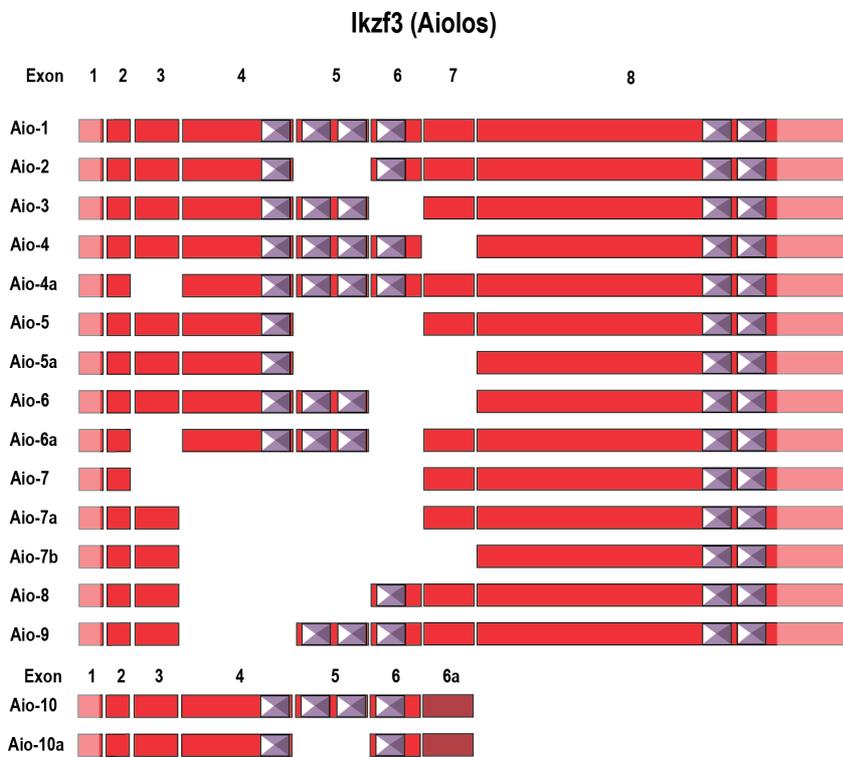
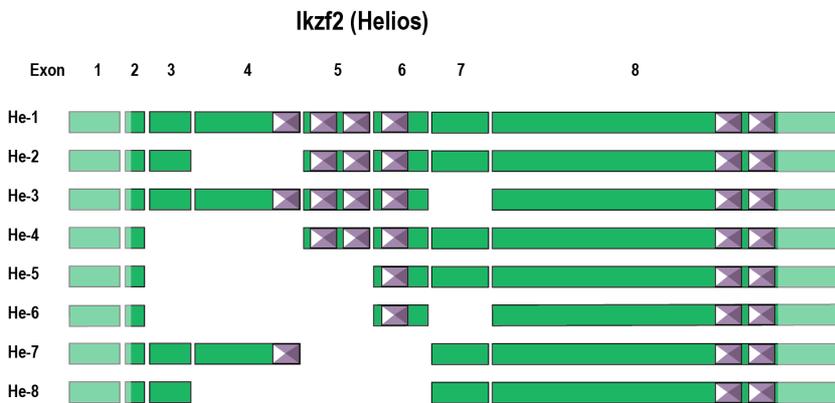
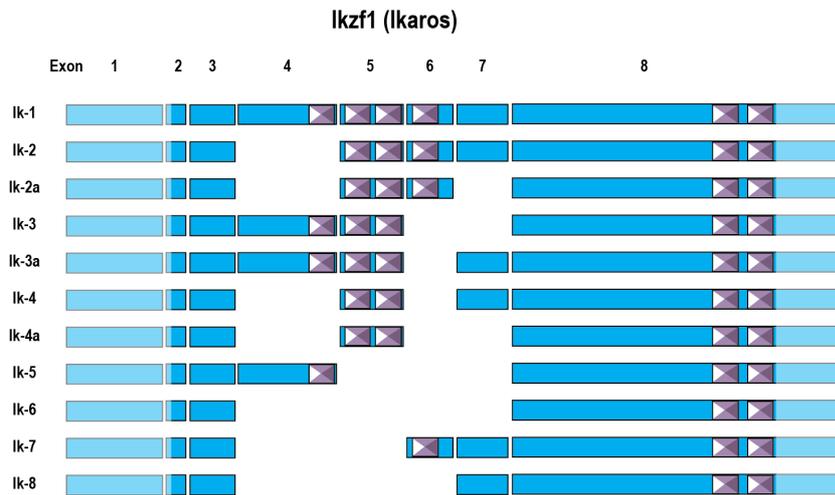


Figure I.7 : Les différentes isoformes décrites pour IKAROS, HELIOS et AIOLOS

Les couleurs atténuées représentent les régions non traduites.

Les isoformes d'Ikaros sont celles décrites par Rebollo & Schmitt en 2003. Les isoformes d'Helios sont celles décrites par Nakase *et al.* en 2002. Les isoformes d'Aiolos sont celles décrites par Caballero *et al.* en 2007. L'exon alternatif 6a est une partie de la séquence intronique entre les exons 5 et 6.

B. La liaison à l'ADN

1. Rôle du domaine de liaison à l'ADN

a) Implication des doigts de zinc

Comme vu précédemment, les facteurs de transcription à doigts de zinc de type C₂H₂ se lient à l'ADN grâce à quatre acides aminés qui font partie de l'hélice du doigts de zinc : les acides aminés en position -1, 2, 3 et 6. La liaison à l'ADN des facteurs de la famille d'IKAROS est surtout médiée par les DZ2 et DZ3. En effet, nous avons vu que les isoformes qui sont dépourvues de ces DZ ne peuvent plus lier l'ADN. Une étude de 2019 a montré que la simple mutation N160A chez AIOLOS avait le même effet sur sa capacité de liaison à l'ADN (Wang et al., 2019). L'acide aminé N160 a son correspondant chez IKAROS (N159) et chez HELIOS (N154) car il s'agit, en fait, du N en position +3 de l'hélice α du DZ2.

Les acides aminés essentiels à la liaison des DZ à l'ADN sont les mêmes pour les trois facteurs (Tableau I.IV.1).

Tableau I.II-1 Comparaison de la composition des doigts de zinc et prédiction de motifs de liaison

	DZ4				DZ3				DZ2				DZ1			
Position	6	3	2	-1	6	3	2	-1	6	3	2	-1	6	3	2	-1
IKAROS	E	S	S	Q	G	A	D	R	R	N	G	Q	V	V	N	G
HELIOS	E	S	S	Q	G	A	D	R	R	N	G	Q	V	V	N	G
AIOLOS	E	S	S	Q	G	A	D	R	R	N	G	Q	V	V	N	G
Prévision de Kaplan	C/G/A			A/G/T	T/A	T		G	G	A/G/C		A/G/T	G/T	G/T		
Prévision de Persikov	C	T		A	T	T		G	G	A		A	T			

Dans leur modèle de prévision de site de liaison à l'ADN combinant des résultats *in vitro* et *in vivo*, Dogan et ses collaborateurs ont déterminé que les DZ1 et DZ4 d'AIOLOS n'ont pas de rôle dans la liaison à l'ADN et que les DZ2 et D3 reconnaissent un motif GGGAA (Dogan et al., 2019).

b) Implication des linkers

Comme nous l'avons noté plus haut, les linkers entre les doigts de zinc du domaine de liaison à l'ADN, dont le consensus est TGEKP, jouent un rôle dans la stabilisation de cette liaison et plus particulièrement la glycine (G) et la lysine (K). Lorsque l'on compare les linkers entre les différents doigts de zinc des protéines de la famille d'IKAROS, la première chose frappante est l'absence de K dans le linker DZ1-DZ2 et du G, dans le cas d'IKAROS (Tableau I.IV.2). De la même façon, le linker DZ3-DZ4, bien qu'il possède le G et le K paraît totalement désorganisé. Ces observations vont dans le sens d'une utilisation stricte des DZ2 et DZ3 pour la liaison à l'ADN et d'une utilisation accessoire des DZ1 et DZ4 qui pourraient lier l'ADN de façon plus transitoire. Il est à noter que la sérine (S) qui remplace le T en première position est également un acide aminé qui peut être phosphorylé permettant ainsi l'inactivation au cours du cycle cellulaire.

Linkers	DZ1-DZ2	DZ2-DZ3	DZ3-DZ4	DZ5-DZ6
IKAROS	TERPFQ	SGEKPFK	SVGKPHK	GCHGFRDPFE
HELIOS	TGERPFH	SGEKPFK	SVGKPHK	GYRDPLE
AIOLOS	TGERPFQ	TGEKPFK	SVEKPYK	GFRDPFE

Tableau I.II-2 Comparaison des linkers entre chaque doigt de zinc adjacent

2. Implication de l'ADN et motif de liaison

Sur la base de données JASPAR recensant les motifs des facteurs de transcription dans différents groupes taxonomiques, il n'y a de motif que pour IKAROS chez Homo sapiens (Figure I.7.A.) (Fornes et al., 2020). Lorsque l'on compare les motifs décrits pour les trois facteurs dans la littérature, il apparaît bien un motif central GGAA mais aussi quelques variations. Certaines publications décrivent des motifs tels que CTGGGAA ou TGGGAA ou encore GGGAA pour IKAROS (Ma et al., 2010; Molnár and Georgopoulos, 1994). A l'exception du second motif décrit pour Aiolos, GGAA semble être le cœur de chaque motif (Figure I.7).

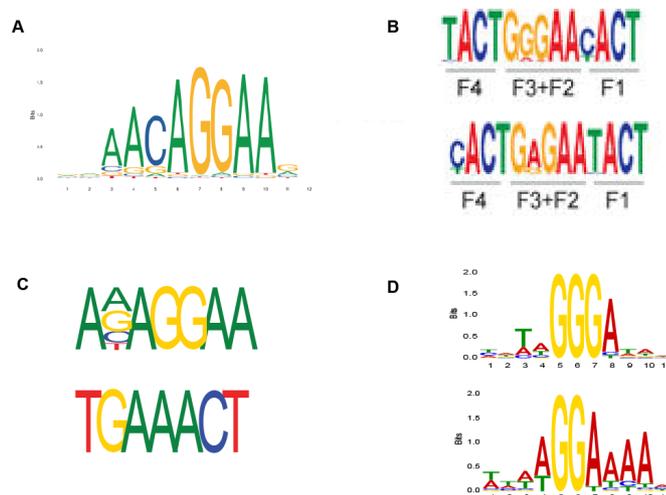


Figure 1.8 : Motifs de liaison à l'ADN pour les facteurs de transcription de la famille d'Ikaros.
 A. Motif de la base de données JASPAR pour IKAROS chez l'Homme
 B. Motifs décrits par pour IKAROS chez la souris (Schjerven et al., 2013)
 C. Motifs décrits pour AIOLOS chez la souris (Hung et al., 2016)
 D. Motifs décrits pour HELIOS, mis en forme avec ggseqLogo (Hahm et al., 1998; Wagih, 2017)

C. Interaction Protéine-Protéine : domaine de dimérisation

Les protéines de la famille d'Ikaros possèdent deux doigts de zinc de type C₂H₂ en position C-terminale qui ne participent pas à la liaison à l'ADN mais sont responsables de la formation des homodimères et des hétérodimères (Hahm et al., 1994). Le linker entre les DZ5 et DZ6 des protéines de la familles d'Ikaros ne ressemble pas du tout au consensus TGEKP retrouvé dans les domaines de liaison à l'ADN, renforçant l'idée que ces doigts de zinc ne sont pas impliqués dans une interaction protéine/ADN (Tableau 2).

Ainsi, en 2003, McCarty et collaborateurs ont entrepris d'étudier les propriétés de ce domaine de dimérisation et de déterminer les acides aminés impliqués dans l'interaction protéine-protéine. Le remplacement des DZ5 et DZ6 du domaine de dimérisation par les DZ2 et DZ3 du domaine de liaison à l'ADN abolit la dimérisation ce qui montre que ces deux DZ ont des propriétés spécifiques à l'interaction protéine-protéine (McCarty et al., 2003). Il semble également que l'acide aspartique en position 18 du domaine de dimérisation, dans le DZ5, soit en position 469 de la protéine Ikaros, ait un rôle déterminant dans la propriété de dimérisation. Cependant l'abolition de la dimérisation par la mutation D18Q ou D469Q a été réalisée sur une protéine tronquée ne possédant qu'un domaine N-terminal réduit et le domaine de dimérisation. Sur le plan de l'évolution ce domaine de dimérisation, pourtant structurellement conservé chez les homologues d'Ikaros comme Hunchback (drosophile), a néanmoins changé, empêchant la dimérisation d'Ikaros avec Hunchback (McCarty et al., 2003). Le

remplacement de non moins de 11 résidus est nécessaire pour rétablir une interaction entre ces deux protéines. Cette évolution montre l'importance et la haute sélectivité de ce domaine. En effet, malgré une certaine homologie entre le domaine de dimérisation d'Ikaros et le domaine C-terminale de la protéine humaine Trps1 (Tricho-Rhino-Phalangeal Syndrome Proteine), les deux protéines n'interagissent pas. Il n'est donc pas suffisant de posséder un domaine de dimérisation avec deux doigts de zinc de type C_2H_2 pour interagir avec Ikaros.

Ceci a de nouveau été démontré en 2004 par l'étude de Westman *et al.* qui ont réussi à résoudre la structure du domaine de dimérisation d'Eos, le quatrième membre de la famille d'IKAROS. Les chercheurs ont mis en évidence des acides aminés impliqués dans l'homodimérisation et/ou l'hétérodimérisation d'Eos avec lui-même ou IKAROS, Pegasus ou encore Trps1. Nous y retrouvons D469 qui est impliqué dans toutes les formes de dimérisation (Westman *et al.*, 2004). Il est à noter ici qu'Eos, qui fait partie de la famille d'Ikaros et qui est capable de se dimériser avec lui, est aussi capable d'interagir avec Trsp1 alors que nous avons vu précédemment d'Ikaros ne le peut pas. Il y a donc, au sein de la même famille, des affinités de dimérisation différentes favorisant certains partenaires. Ce travail a été confirmé et approfondi en 2008 par Brayer et Segal qui ont déterminé quels acides aminés sont nécessaires à l'interaction d'Eos avec Ikaros, Pegasus et Trsp1 (Brayer and Segal, 2008). Alors que les 42 substitutions testées n'ont que peu d'effet sur l'interaction en Eos et Pegasus, 66% d'entre elles abolissent l'interaction Eos-Trsp1 et 4, au sein du DZ5 d'Ikaros, l'interaction Eos-Ikaros. Cependant, bien qu'elle ait été testée la substitution D469A ne semble pas ici abolir la dimérisation.

L'ensemble de ces études montre que l'interaction protéine-protéine médiée par les doigts de zinc C_2H_2 n'est pas encore bien définie. Il semble, cependant, que ce domaine de dimérisation soit hautement sélectif ce qui pourrait signifier que les différents dimères puissent avoir des rôles particuliers.

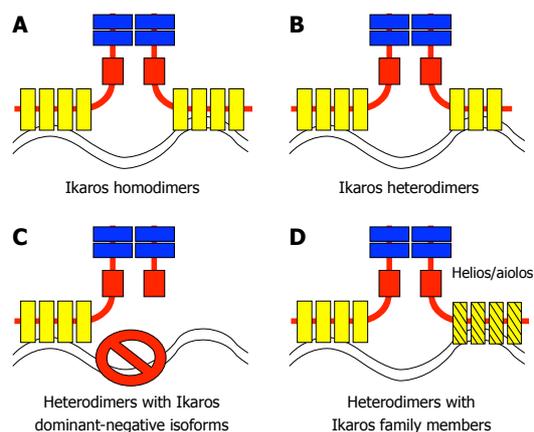


Figure I.9 : Schéma des dimères formés par Ikaros
Li et al., 2011

D. Mécanismes moléculaires de la régulation de la transcription

Les facteurs de transcription de la famille d'Ikaros ne sont pas des facteurs de transcription classique. Ils sont bien sûr capables de recruter le complexe de pré-initiation de la transcription aux promoteurs de leurs gènes cibles mais ils fonctionnent principalement par le recrutement de complexes de remodelage de la chromatine tels que NuRD (Nucleosome remodeling Deacetylase), SWI/Snf (Switch/Sucrose non fermentable) ou PRC2 (Polycomb Repressor Complex 2). Les différents mécanismes connus vont être détaillé dans les sections suivantes.

1. Association avec NurD

Le complexe NuRD est un des complexes majeurs du remodelage de la chromatine. Des altérations de son activité sont impliqués dans des défauts du développement embryonnaire et dans diverses pathologies humaines tels que les cancers (Basta and Rauchman, 2015). Le complexe est composé de différentes protéines permettant d'assurer des fonctions enzymatiques (HDAC1 et 2, Mi2 α,β) et la liaison à l'ADN (MTA1/2/3, MBD2/3, Rbbp4/7, Gata2a/2b). Certaines de ces protéines sont dépendantes du type cellulaire.

L'association des facteurs de transcription de la famille d'Ikaros avec le complexe NuRD peut être impliquée soit dans l'activation, soit dans la répression de la transcription.

a) Pour l'activation de la transcription

Il a été montré dans la lignée de lymphocytes T CD4+ humains, Jurkatt, qu'IKAROS interagit avec les complexes NuRD, via la sous-unité Mi2 β , et p-TEFb (positive Transcription-Elongation factor), via la sous-unité CDK9 (cyclin-dépendant kinase 9) (Bottardi et al., 2014). Cette même étude montre qu'Ikaros est capable de recruter ces deux complexes aux sites de démarrage de la transcription (TSS – transcription Start Site) de ses gènes cibles pour en favoriser la transcription et l'élongation dans les progéniteurs hématopoïétiques murins. De plus, ce recrutement est sensible aux niveaux d'Ikaros présents dans la cellule ce qui contribue aux phénotypes d'haplo-insuffisances d'Ikaros décrit dans certaines pathologies hématopoïétiques (Figure I.9).

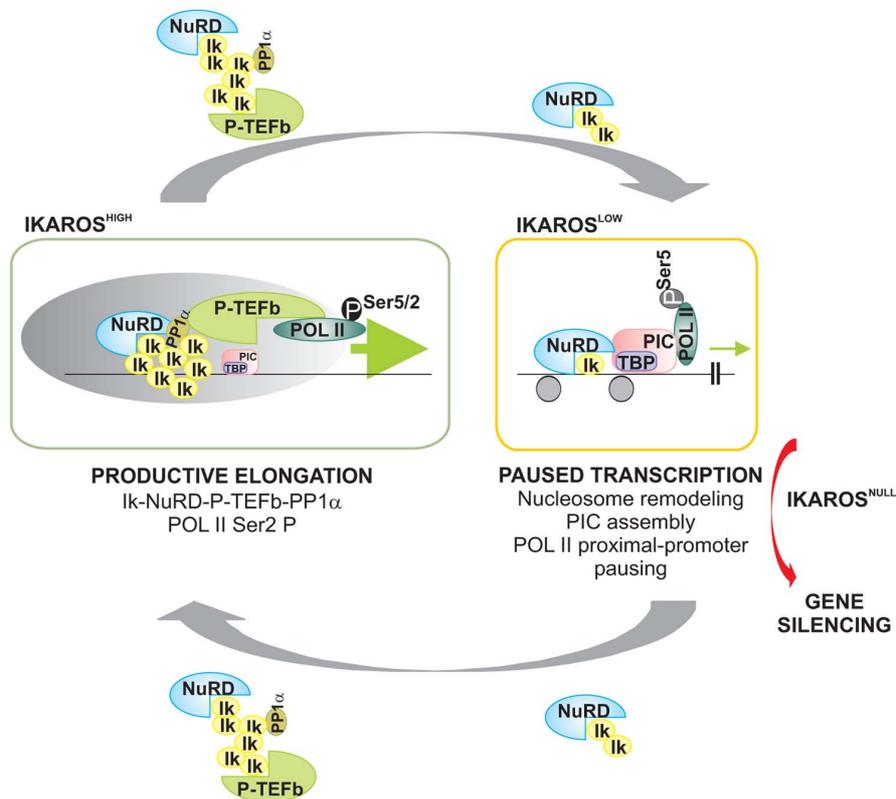


Figure I.9 : Association d'Ikaros avec NuRD et p-TEFb

b) Pour la répression de la transcription

Une autre étude a cependant mis en évidence le recrutement du complexe NuRD par Ikaros, grâce à sa liaison avec la sous-unité Mi-2 β , qui entraîne l'éviction de l'ARN polymérase II, la compaction de la chromatine et sa relocalisation au niveau de

l'hétérochromatine péri-centromérique. Ce processus étant indépendant de l'activité des HDAC1/2 (Histone déacétylases) (Liang et al., 2017).

Il est également possible de retrouver une répression dépendante des HDAC lors de l'activation des lymphocytes T où la localisation de Mi-2 β et HDAC-1 au niveau de l'hétérochromatine est dépendante d'Ikaros. Ainsi, Ikaros recrute le complexe NuRD au niveau de la chromatine accessible, les histones sont désacétylées et la région compactée est dirigée vers l'hétérochromatine. Mais IKAROS peut également recruter NuRD au niveau de sites déjà inaccessibles pour maintenir le statut de la chromatine (Kim et al., 1999). Aiolos est également retrouvé dans ces complexes et interagit avec Mi-2 β mais ne semble pas pouvoir le recruter en absence d'Ikaros. Helios s'associe également avec NuRD, en interagissant avec Mi-2 β , et Ikaros dans les thymocytes (Sridharan and Smale, 2007). De la même façon, Helios est localisé au niveau de l'hétérochromatine.

2. Association avec le complexe mSin3

Le complexe mSin3 est associé à une répression de la transcription de façon dépendante des HDAC. Ikaros est capable d'interagir avec différents composants de ce complexe tels que mSin3A, mSin3B, SAP18 (mSin Associated Protein), HDAC1 and HDAC2. Une fois encore, Aiolos est présent dans le complexe dans les lymphocytes T (Koipally et al., 1999). Cette association avec mSin3 ne se fait pas au niveau des *foci* de PC-HC, offrant à Ikaros un autre mode de répression de la transcription. Cependant, cette association avec mSin3 est également dépendante du type cellulaire et du stade de différenciation.

3. Association avec SWI/Snf

Le complexe de remodelage de la chromatine SWI/Snf a originellement été identifié grâce à des études sur la levure dans laquelle il est impliqué dans la régulation de l'expression de 5% des gènes. Chez les mammifères, deux variants ont été identifiés : BAF et PBAF. Comme NuRD il est composé de plusieurs sous-unités dont Brg1, une ATPase, qui en est la sous-unité catalytique (Roberts and Orkin, 2004). Des mutations de Brg1 prédisposent au développement de tumeurs.

Ikaros et Aiolos co-localisent avec des protéines du complexes SWI/Snf dans les lymphocytes T. Il a été montré par immunoprécipitation qu'Ikaros interagit

effectivement avec Brg-1. Cette association pourrait être responsable de l'ouverture de la chromatine et donc la rendre accessible à d'autres facteurs pour permettre la transcription (Figure I.10) (Kim et al., 1999).

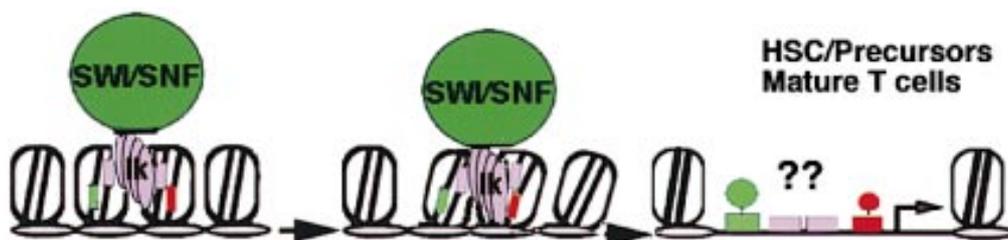


Figure I.10 : L'association d'IKAROS avec le complexe SWI/Snf

4. Interaction avec CtBP

CtBP (C-terminal Binding Protein) agit avec CtIP (CtBP Interacting Protein) pour réprimer la transcription en empêchant le recrutement du complexe de pré-initiation de la transcription. Ikaros mais pas Aiolos est capable d'interagir avec CtBP grâce au motif PEDLS présent dans son domaine N-terminal (Figure I.11). En revanche, Aiolos peut se lier à CtIP directement dans un mécanisme de répression indépendant de CtBP (Koipally & Georgopoulos, 2000, 2002b).

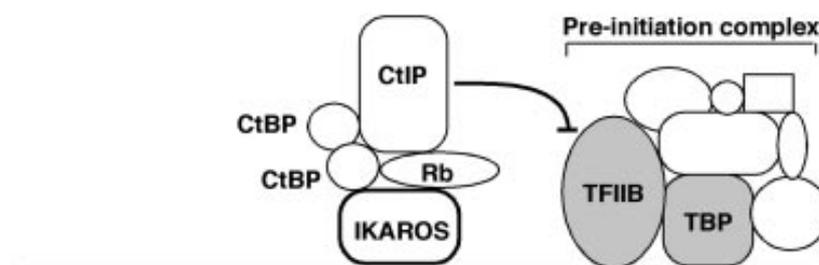


Figure I.11 : Association d'IKAROS avec CtBP et CtIP

Tous les mécanismes précédemment expliqués sont possible grâce à l'interaction d'Ikaros avec diverses partenaires protéiques. Les capacités de liaison d'Ikaros à ces protéines sont résumés dans le travail de Koipally et Georgopoulos en 2002 (Koipally and Georgopoulos, 2002b) (Figure I.12)

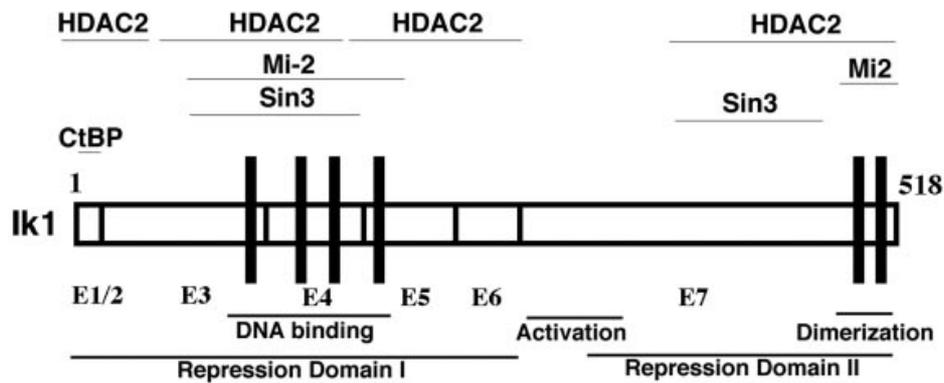


Figure I.12 : Les sites d'interaction d'IKAROS avec ses partenaires protéiques

5. Association avec PRC2

Ikaros agit également au niveau d'enhancers en voie d'inactivation avec EBF1, un facteur de transcription spécifique des lymphocytes B, pour réprimer la transcription de gènes qui pourraient interférer avec le processus de différenciation des cellules B. Les promoteurs de ces gènes sont également ciblés par le complexe PRC2 (Polycomb Repressiv Complex 2) qui entraîne les modifications des marques des histones vers des marques répressives (Hu et al., 2016). Même si l'activité de PRC2 est perdue lorsque Ikaros WT est remplacé par un dominant négatif, il ne semble pas y avoir d'interaction directe entre Ikaros et PRC2 (Figure I.13). En revanche, il a été montré qu'au cours de la différenciation des lymphocytes T, Ikaros recrute PRC2 au niveau des promoteurs de ses gènes cibles pour en réprimer la transcription (Oravec et al., 2015).

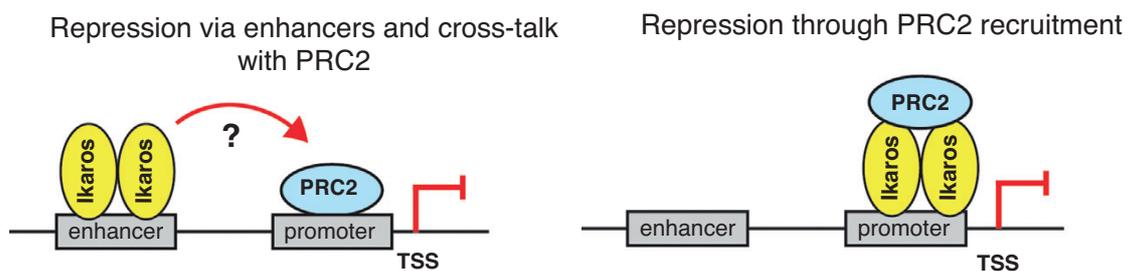


Figure I.13 : Mode de répression de la transcription via PRC2 (Heizmann et al., 2018)

6. Autres modes de régulations

a) Super-enhancer

Ikaros se lie au niveau de super-enhancers qui constitué de différents sites de liaisons pour des facteurs de transcription majeurs. Par exemple, dans les lymphocytes pré-B, il a été montré qu'IKAROS interagit avec d'autres facteurs de transcription spécifiques de la différenciation des lymphocytes B : Pax5, E2A, IRF4 au niveau de super-enhancers actifs contribuant à l'activation de la transcription des gènes cibles (Figure I.14). Il semble que la présence d'IKAROS soit indispensable pour la fonctionnalité de ces super-enhancers non pas pour le recrutement des autres facteurs de transcription, sauf peut-être AIOLOS et E2A, mais pour permettre le maintien des marques permissives des histones (par exemple : H3K4me2 et H3K27Ac) (Hu et al., 2016).

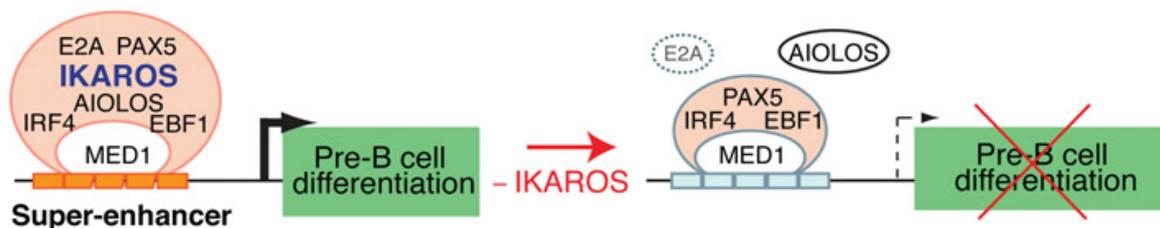


Figure I.14: IKAROS se fixe aux super-enhancers

Par comparaison entre le nombre et la localisation de super-enhancers avant et après introduction d'IKAROS dans des thymocytes au stade de différenciation DN3, il a été démontré que la liaison d'IKAROS sur l'ADN peut également entraîner la formation de nouveaux super-enhancers (Ding et al., 2019).

b) Hétérochromatine Peri-centromérique (PC-HC)

Un autre mode de répression de la transcription par Ikaros est le recrutement de ses gènes cibles au niveau l'hétérochromatine pericentromérique (PC-HC). Une étude menée en 1997 met en évidence la localisation d'Ikaros au niveau de l'hétérochromatine pericentromérique dans les lymphocytes pré-B en phase G1 et G2 du cycle cellulaire. De plus, ils semble qu'Ikaros puisse recruter ses gènes cibles dont l'expression doit être réprimée au cours de la différenciation des lymphocytes B vers

les foci d'hétérochromatine (Brown et al., 1997). En 2000, Cobb et collaborateur ont montré que le domaine de liaison à l'ADN d'Ikaros est nécessaire à sa localisation dans l'hétérochromatine peri-centromérique et qu'Ikaros est aussi localisé au niveau de séquences répétées riches en GC situé au sein des centromères des chromosomes et qui s'appellent les γ -satellites (Cobb et al., 2000). Helios et Aiolos sont également localisés au niveau des *foci* de PC-HC (Hahm et al., 1998). De la même façon les isoformes d'Helios et Aiolos dépourvues des DZ2 et DZ3 ne sont plus capable de rejoindre les foci de PC-HC (Caballero et al., 2007).

c) Interaction avec d'autres facteurs de transcription

Pour contrôler l'expression de ses gènes cibles, il a été montré qu'Ikaros est capable de recruter les facteurs de la famille GATA grâce à son domaine C-terminal (Bottardi et al., 2013).

L'interaction entre Aiolos et Blimp-1 (B lymphocyte-induced maturation protein-1) joue un rôle important dans la survie des cellules plasmatisques impliquées dans le myélome multiple. Les deux premiers doigts de zinc de Blimp-1 assurent l'interaction avec le domaine N-terminal d'Aiolos, mais pas avec Ikaros. Les deux protéines lient alors l'ADN pour réguler l'expression de leurs gènes cibles (Hung et al., 2016).

E. Régulation de l'expression ou de l'activité des facteurs de la famille d'Ikaros

1. Régulation transcriptionnelle

Peu d'études ont été menées sur la régulation transcriptionnelle des facteurs de la famille d'Ikaros. Il semblerait, néanmoins, que l'expression d'Ikaros et d'Aiolos dans les lymphocytes pré-B soit liée aux facteurs de transcription IRF4 et IRF8. En effet chez les souris double mutantes pour ces deux facteurs, Ikaros et Aiolos sont peu exprimés par rapport au contrôle et les cellules sont bloquées dans leur différenciation (Lu, 2003). La réexpression d'IRF4 ou d'IRF8 permet l'augmentation de l'expression d'Ikaros et d'Aiolos et IRF4 le fait plus efficacement qu'IRF8. Ceci permet de réenclencher le processus de différenciation montrant que le rôle d'IRF4 et IRF8 dans la différenciation précoce des lymphocytes B est médiée par Ikaros et Aiolos (Ma et al., 2008). Au niveau supérieur, Fra2 (FOS related antigen 2) a été montré comme

régulateur de l'expression d'IRF4 dans les lymphocytes pro-B (Ubieta et al., 2017). De plus, l'expression d'IRF4 semble être contrôlée par BLNK (B-cell Linker protein) aussi connue sous le nom de SLP-65 (Thompson et al., 2007). Cette protéine fait partie de la voie de signalisation du récepteur des cellules pre-B, le pre-BCR. Cependant dans le cas d'une absence de BLNK dans les lymphocytes pro-B, l'expression d'Ikaros reste inchangée alors qu'Aiolos est absent ce qui suggère que l'expression d'Aiolos dans les lymphocytes pré-B est contrôlée par la voie de signalisation du pré-BCR.

Malgré le fait que l'on puisse faire un parallèle entre différenciation B et T pour l'expression d'AIOLOS, aucune étude n'a clairement montré que l'expression d'Aiolos puisse être médiée par la voie de signalisation du pré-TCR (récepteur des lymphocytes pre-T).

La régulation de la transcription d'Helios n'a pas été très étudiée. Cependant au cours du développement fœtal il semble que la transcription d'Helios soit régulée par Gsx2 et Dlx1&2 dans certaines cellules neurales (Martín-Ibáñez et al., 2017).

Bien qu'il ne soit pas traduit, l'exon 1 d'*Ikzf1* a un rôle particulier dans la régulation de son expression, il en est le promoteur. Il y a, en fait, deux exons 1a et 1b qui sont individuellement inactivés en fonction de la lignée cellulaire (myéloïde ou lymphoïde). De plus, plusieurs enhancers ont été identifiés dans les différents introns du gène (Yoshida et al., 2013). Ces enhancers peuvent être spécifiques du stade de différenciation. L'un d'eux est, en fait, en amont du gène dans la région 5' à moins de 10 kilobases (kb) de l'exon 1. Dans notre laboratoire nous sommes intéressés aux cellules de patients atteints de leucémie aiguë lymphoblastique B (B-ALL) ayant une délétion de l'exon 1 d'*IKZF1*. En général l'absence de l'exon 1 est due à une délétion plus large de la région 5' du gène *IKZF1*, parfois de plusieurs dizaines voire centaines de kb (Morel, Deau et al., 2019). En effet, les cellules présentant une délétion de la région 5' de façon mono- ou bi-allélique ont une réduction de l'expression d'IKAROS, respectivement de moitié ou totale. Cela confirme l'importance de cette région dans la régulation de la transcription d'*Ikzf1* et la présence de régulateurs de la transcription. Récemment, le super-enhancer E120 a été caractérisé. Il est à 120kb en amont du gène *Ikzf1* et sa délétion entraîne une réduction significative de la transcription d'*Ikzf1* (Alomairi et al., 2020).

2. Régulation post-traductionnelle

La localisation et l'activité d'Ikaros sont régulées par des modifications post-traductionnelles.

a) Sumoylation

La sumoylation d'Ikaros sur les lysines K58, K240 et K425 ne change pas la localisation d'Ikaros au niveau de l'hétérochromatine mais empêche sa liaison avec les complexes de remodelage de la chromatine et son rôle dans la répression de la transcription de ses gènes cibles (Gómez-del Arco et al., 2005). Cela a pour effet de réduire son efficacité dans l'inhibition de l'expansion des cellules leucémiques chez la souris. La sumoylation apparaît alors comme une modification entraînant la réduction voire la perte de fonction de suppresseur de tumeur d'Ikaros (Apostolov et al., 2016).

Les motifs de sumoylation (ψ KxE, où ψ représente un résidu hydrophobique) aux positions K58 et K240 semblent conservés chez Aiolos aux positions K61 et K244. Seule la position K58 est conservée chez Helios en position K51. Cependant aucune étude n'a mis en évidence la sumoylation d'Aiolos et d'Helios à ces positions.

b) Phosphorylation

Grâce à des expériences de phospho-mimétisme, il a été montré que la phosphorylation d'Ikaros par CK2 (Casein Kinase 2) au sein du domaine C-terminale joue un rôle dans la capacité d'IKAROS à réguler la progression du cycle cellulaire entre les phases G1 et S (Gómez-del Arco et al., 2004). De plus, la phosphorylation des sérines 13 et/ou 294 par CK2, entraîne la délocalisation d'Ikaros des *foci* de PC-HC (Gurel et al., 2008). La tyrosine kinase BTK est également capable de phosphoryler les sérines 214 et 215 qui se trouvent à dans la séquence du DZ4 d'Ikaros et cela permettrait de contrôler sa localisation nucléaire et d'augmenter sa capacité de liaison à l'ADN (Ma et al., 2013). La protéine SYK (Spleen Tyrosine Kinase) est également capable de phosphoryler Ikaros aux positions S358 et S361, au sein du domaine de dimérisation, favorisant sa localisation nucléaire et sa liaison à l'ADN (Uckun et al., 2012). La phosphorylation apparaît comme mécanisme de régulation de la fonction d'Ikaros au cours de la différenciation des thymocytes. En effet, il est possible

d'associer un profil de phosphorylation d'Ikaros à différents stades de différenciation des lymphocytes T (Sridharan and Smale, 2007).

Comme expliqué précédemment, la phosphorylation des thréonines ou sérines des linkers entre les doigts de zinc du domaine de liaison à l'ADN abolit la liaison pendant la phase G2/M du cycle cellulaire (Dovat et al., 2002). Les substitutions phospho-mimétiques abolissent la localisation dans les foci de PC-HC montrant que cette localisation est dépendante de la liaison à l'ADN. La kinase responsable de la phosphorylation des linkers au cours du cycle cellulaire n'est toujours pas identifiée.

c) Ubiquitination

L'ubiquitination est un mode de régulation négative par dégradation de la protéine. Certaines molécules immuno-modulatrices (ImiD) tels que le Lénalidomide ou le Pomalidomide ciblent le complexe humain CRL4^{CRBN} (Cullin Ring E3 ubiquitin ligase complex ; Cereblon) permettant alors le recrutement d'IKAROS et d'AIOLOS et leur ubiquitination les menant à la dégradation via le protéasome (Gandhi et al., 2014).

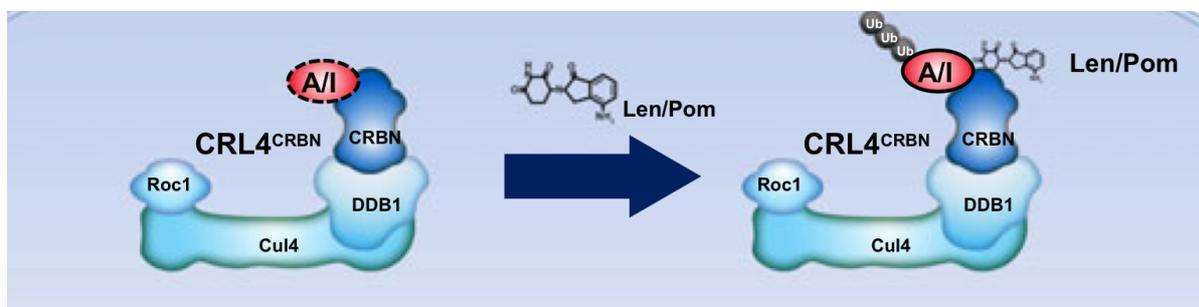


Figure I.15 : Processus de poly-ubiquitination d'IKAROS et d'AIOLOS par CRL4^{CRBN} en présence de Lenalidomide ou Pomalidomide

III. Rôles dans le développement des lymphocytes B

Les lymphocytes B ont été nommés ainsi pour la « Bourse de Fabrice », organe présent chez les oiseaux, qui a été décrit comme le lieu de production des cellules sécrétant des anticorps (Glick et al., 1956). Dans le sang, ils représentent 15 à 20% des lymphocytes circulants. Ils sont caractérisés par l'expression, à leur surface, du récepteur spécifique des lymphocytes B, le BCR (B Cell Receptor) et de l'antigène spécifique des lymphocytes B, CD19. Les lymphocytes B sont impliqués dans la réponse humorale en sécrétant des anticorps spécifiques à l'antigène et en présentant cet antigène aux lymphocytes T. Ces deux spécificités en font des cellules incontournables du système immunitaire.

Les lymphocytes B se développent en deux temps. Le premier est indépendant de l'antigène et se déroule dans la moelle osseuse, c'est le développement précoce. Dans un second temps, la maturation terminale est dépendante de l'antigène et se déroule dans les organes lymphoïdes secondaires tels que la rate et les ganglions lymphatiques. Cette maturation permet la diversification du répertoire des anticorps et de différencier les lymphocytes B en lymphocytes B mémoires ou en plasmocytes.

Pour ce travail de thèse, mon modèle d'étude est le développement précoce des lymphocytes B et plus particulièrement les dernières étapes de ce processus.

A. Le développement précoce des lymphocytes B

1. De la cellule souche hématopoïétique à l'expression du pré-BCR

La production des cellules sanguines, ou hématopoïèse, commence dans le sac vitellin au tout début du développement embryonnaire. Les cellules produites alors sont principalement des érythrocytes qui permettront l'oxygénation des tissus de l'embryon. Puis, dans fœtus, le foie fœtal devient le siège de l'hématopoïèse. Une dernière migration a lieu vers la moelle osseuse où les cellules souches hématopoïétiques resteront tout au long de la vie (Galloway and Zon, 2003). La cellule souche hématopoïétique a été nommée ainsi grâce à sa capacité de reconstitution de l'hématopoïèse chez des sujets irradiés. De récentes études ont montré que cette population cellulaire est en fait divisée en groupe de cellules avec différents potentiels de reconstitution (Ali et al., 2017; Wilson et al., 2015). Ainsi les cellules ayant une capacité de reconstitution de l'hématopoïèse définitive sont appelées Long-Term (LT)-

HSC. Puis, dans le compartiment suivant nous retrouvons des cellules qui ont une capacité de reconstitution réduite à quelques mois et qui sont appelées Short-Term (ST)-HSC. Enfin, la dernière catégorie regroupe les progéniteurs multipotents (MPP) qui ont une capacité de reconstitution de quelques semaines. Ces cellules peuvent encore être divisées en trois catégories chacune regroupant des MPP qui sont biaisés vers l'une des trois lignées hématopoïétiques : myéloïde, lymphoïde et érythrocytaire (Pietras et al., 2015). En effet, les MPP4 ont un potentiel lymphoïde plus important que les MPP2 et 3.

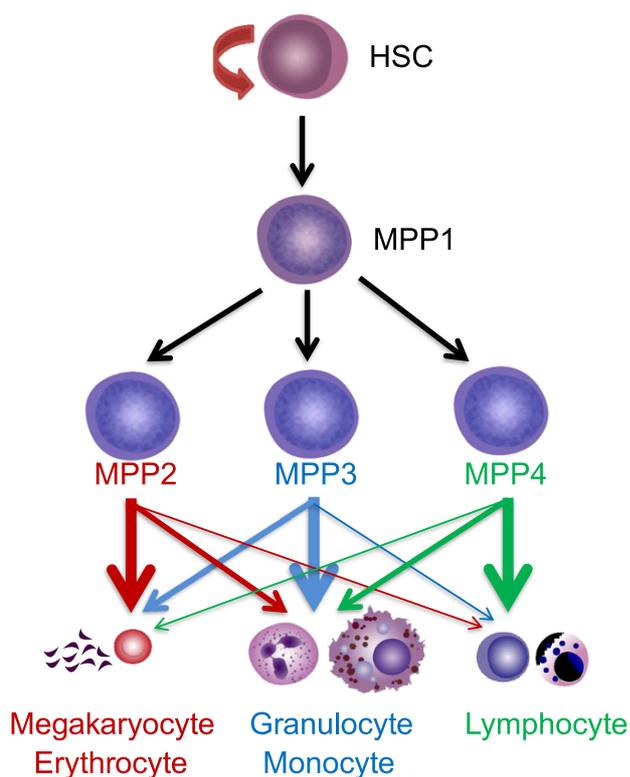


Figure I.16 : Compartiments hématopoïétiques (Cheng et al., 2020)

Les marqueurs phénotypiques des MPP4 (LSK+, Flk2/Flt3+, CD150+, CD48+) recourent ceux des progéniteurs multipotents lymphoïdes (LMPP) qui avaient précédemment été décrits (Adolfsson et al., 2005; Pietras et al., 2015). Sous l'influence, et plus particulièrement la diminution de l'expression du facteur de transcription PU.1, codé par le gène *Spi1*, les MPP4/LMPP s'engagent dans la voie lymphocytaire (DeKoter and Singh, 2000). Les progéniteurs lymphoïdes communs (CLP), ont été identifiés en 1997 par des expériences de transplantation chez des souris irradiées, ces cellules ne pouvant générer que des lymphocytes B, T et Natural Killer (NK) (Kondo et al., 1997).

Les cellules entrent dans le développement précoce des lymphocytes B comprenant plusieurs stades de différenciation. Deux nomenclatures se font face pour nommer ces différents stades, la nomenclature de Rolink (ou nomenclature de Bâle) et la nomenclature de Hardy (ou nomenclature de Philadelphie) (Hardy et al., 1991; Rolink et al., 1993) (Figure I.17).

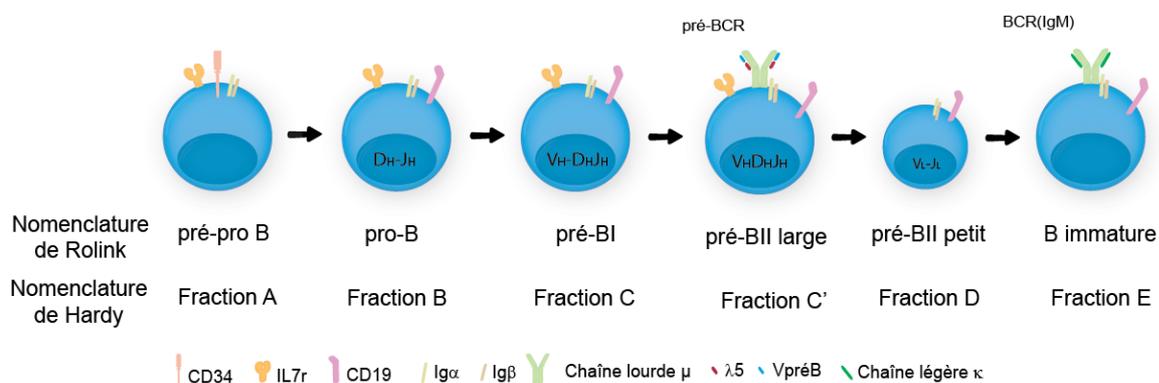


Figure I.17 : Le développement précoce des lymphocytes B

Chaque stade est caractérisé phénotypiquement par la combinaison de différentes protéines membranaires et par les étapes du réarrangement des loci des chaînes lourdes et légères de immunoglobulines (Ig). En effet, le récepteur des lymphocytes B (BCR) passe par plusieurs étapes de maturation avec d'être exprimé à la surface des lymphocytes B. Le BCR est, en fait, un anticorps, ou immunoglobuline (Ig), composé de deux chaînes lourdes et de deux chaînes légères identiques entre-elles. La chaîne lourde est composée d'une partie constante, ancrée à la surface des cellules, qui est vectrice de la fonctionnalité de l'anticorps en solution et une partie variable qui permet la reconnaissance de l'antigène. A l'état germinatif, la partie variable est constituée de différents segments divisés en fragments et regroupés par famille. Par exemple, le segment V_H (variable) est composé d'une centaine de fragments regroupés en 13 familles. Il y a ainsi le segment D_H (diversité) et J_H (jonction). La première étape du réarrangement du *locus* des chaînes lourdes des Ig se déroule dans la fraction B et consiste à relier des fragments D_H (Diversity heavy chain) et J_H (Joining Heavy chain). La liaison entre V_H et D_HJ_H (Variety Heavy chain) intervient dans la fraction C (Hardy et al., 1991).

Ce processus est orchestré par les protéines RAG1 et RAG2 (Recombination Activating Gene) et assure la pérennité du développement précoce des lymphocytes B (Mombaerts et al., 1992; Shinkai et al., 1992). L'enzyme TdT (Terminal

deoxynucleotid Transferase) codée par le gène *Dntt* permet d'augmenter la diversité du répertoire des immunoglobulines en ajoutant des nucléotides dans la région variable de la chaîne lourde au moment du réarrangement. Elle n'est, cependant, pas indispensable au développement des lymphocytes B (Gilfillan et al., 1993, 1995). Les cellules n'ayant pas réussi à réarranger le locus des chaînes lourdes sont dirigées vers l'apoptose.

Le réarrangement du locus des chaînes lourdes permet l'expression de la chaîne lourde μ , qui, associée aux chaînes légères de substitution $V_{\text{préB}}$ et $\lambda 5$, compose le pré-BCR (Kudo and Melchers, 1987; Sakaguchi and Melchers, 1986). La chaîne lourde μ ne possédant pas de domaine intracellulaire, indispensable à la transduction du signal, c'est l'hétérodimère $Ig\alpha/Ig\beta$ qui assure cette fonction pour le pré-BCR et, plus tard, le BCR. Ces protéines, respectivement codées par les gènes *Cd79a* et *Cd79b* dont les ARNm sont détectés dès la fraction A, sont exprimées dès la fraction B. Les cellules de la fraction C' expriment donc le pré-BCR.

2. Du pré-BCR au BCR

La voie de signalisation du pré-BCR induit la prolifération cellulaire, sa propre répression et la différenciation. Le mode d'activation du pré-BCR n'est pas encore claire, il pourrait s'agir d'un ligand non encore identifié ou de l'agrégation naturelle des pré-BCR à la surface des cellules. La prolifération cellulaire, ou sélection positive, est potentiellement due aux chaînes légères $V_{\text{préB}}$ et $\lambda 5$ (Melchers et al., 1995). La voie de signalisation du pré-BCR assure, en collaboration avec la voie IL7/STAT5 la prolifération et la survie cellulaire (Fleming and Paige, 2002).

De plus, il a été montré que la voie du pré-BCR est impliquée dans la répression de la transcription des gènes des chaînes de substitution $V_{\text{préB}}$ et $\lambda 5$. Cet effet passe par une répression mono-allélique de la transcription des gènes codant pour ces protéines. Les gènes seront ensuite totalement réprimés et délocalisés de l'euchromatine vers l'hétérochromatine (Parker et al., 2005).

C'est la voie de signalisation du pré-BCR via SLP-65 (ou BLNK) qui initie d'une part l'arrêt du cycle cellulaire par la répression de la voie PI3K/Akt de manière dépendante de Syk, une tyrosine kinase associée au motifs ITAM des co-recepteurs $Ig\alpha$ et $Ig\beta$ (Herzog et al., 2008). Enfin, le pré-BCR est impliqué dans la réexpression de RAG1 et RAG2 pour le réarrangement des *loci* des chaînes légères via Foxo1 (Amin and Schlissel, 2008; Dengler et al., 2008; Fleming and Paige, 2002). Il existe

deux *loci* codant pour les chaînes légères des immunoglobulines : κ et λ . Le *locus* de la chaîne κ est le premier à être réarrangé. Soit c'est un succès et la chaîne légère κ sera transcrite et traduite, soit le *locus* est détruit au profit du réarrangement du *locus* de la chaîne λ (Hieter et al., 1981). Chez la souris, 95% des chaînes légères sont des chaînes κ et seulement 5% sont des chaînes λ (McGuire and Vitetta, 1981). Cette étape du développement se déroule dans la fraction D. Le lymphocyte pré-B devient alors un lymphocyte B immature exprimant un BCR composé de la chaîne lourde μ et de la chaîne légère κ , chez la souris. De la même façon que pour la chaîne lourde, les réarrangements infructueux mènent à l'apoptose

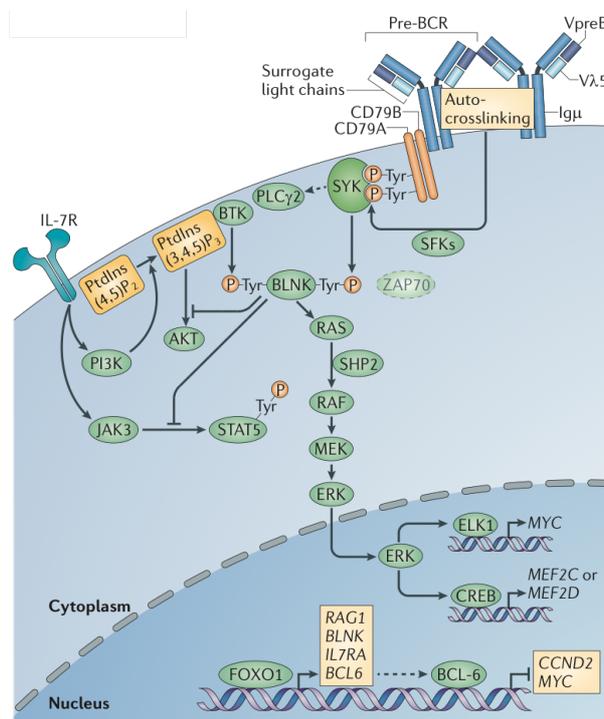


Figure I.18 : Les voies de signalisation induites par le pré-BCR (Rickert, 2013)

L'entrée dans cette dernière étape du développement précoce va de pair avec l'arrêt de la dépendance à l'IL7. La signalisation du pré-BCR permet aussi la répression de FAK (Focal Adhesion Kinase) et des intégrines $\alpha4\beta1$ et favorise l'expression de CXCR4 (Fistonich et al., 2018). Ceci permet la migration des cellules, au sein de la moelle osseuse, d'une niche riche en IL7 vers une niche pauvre en IL7.

3. La voie de signalisation de l'IL7

La voie de signalisation de l'IL7 est nécessaire au développement des lymphocytes B. Elle assure jusqu'au stade C' la survie, la différenciation et la prolifération en régulant différentes voies de signalisation et d'activation de la transcription (Clark et al., 2014).

L'interleukine 7 (IL7) est produite par certaines cellules stromales de la niche hématopoïétique mais également dans le foie fœtal, la rate et le thymus. Elle est impliquée dans la survie et la prolifération des lymphocytes pro-B. Son récepteur (IL7R) est composé de deux sous-unités α (IL7 α) et γ qui est commune avec le récepteur des interleukines 2, 4, 9, 15 et 21. Le récepteur IL7- α , dont l'expression est régulée par PU.1, est exprimé à la surface des CLP (DeKoter et al., 2002).

Les Janus kinases (JAK) s'associent au récepteur de l'IL7. Particulièrement JAK3 s'associe avec la chaîne γ et JAK1 avec la chaîne α . La phosphorylation du récepteur conduit au recrutement des protéines STAT. Dans les lymphocytes B, seuls les protéines STAT1, STAT3 et STAT5 sont activées par l'activation de la voie de signalisation de l'IL7. Bien qu'il ait été montré que STAT3 soit absolument nécessaire à la lymphopoïèse B (Chou et al., 2006), c'est STAT5 qui est préférentiellement activée et qui permet l'activation de la transcription des gènes nécessaire au développement des lymphocytes B (Heltemes-Harris et al., 2011)

La voie de signalisation de l'IL7 est indispensable au réarrangement distal V_H-D_HJ_H (Corcoran et al., 1998). En effet l'activation de STAT5 permet l'ouverture de la chromatine à proximité du locus des chaînes lourdes des Ig (Bertolino et al., 2005). En revanche, la voie de signalisation de l'IL7/JAK/STAT, bien qu'elle permette leurs expressions, empêche l'accès des protéines RAG au *loci* des chaînes légères Ig κ grâce à la liaison de STAT5 au niveau de l'enhancer E κ (Mandal et al., 2011). De façon intracellulaire, la protéine LNK, codée par le gène *Sh2b3*, interagit directement avec la protéine JAK3 (Janus Kinase 3) empêchant la phosphorylation des protéines STAT5a et STAT5b et ainsi leurs activations (Cheng et al., 2016).

4. La tolérance des lymphocytes B ou sélection négative

La tolérance du soi est caractérisée par une non-réponse immunitaire face à un antigène spécifique. Ici, les antigènes du soi, c'est-à-dire de l'individu, ne doivent pas être reconnus par les récepteurs des lymphocytes B ce qui déclencherait une réponse

humorale et une maladie auto-immune. Les réarrangements successifs des chaînes lourdes et légères entraînent statistiquement la possibilité d'obtenir un BCR auto-réactif. Pour éviter que ces cellules ne sortent de la moelle et atteignent les organes lymphoïdes secondaires, les lymphocytes B immatures sont exposés à des antigènes du soi dans le microenvironnement médullaire. Les lymphocytes B auto-réactifs peuvent alors être éliminés par apoptose ou anergie ou réenclencher leurs programmes de réarrangement des chaînes légères jusqu'à l'obtention d'un BCR non auto-réactif (Gay et al., 1993; Goodnow et al., 1988; Nemazee and Bürki, 1989).

5. La maturation terminale des lymphocytes B

Le lymphocyte B immature sort alors de la moelle osseuse pour rejoindre les organes lymphoïdes secondaires tels que la rate ou les ganglions lymphatiques. Il est alors appelé lymphocytes B transitionnel. Un marqueur important est l'expression du récepteur du B-cell activating factor, BAFF (BAFF-R ou BR3 : BLyS receptor 3). Il a été montré qu'en absence du BAFF-R, les lymphocytes B restaient bloqués au stade immature et que, chez les souris mutantes, il y avait une absence de lymphocytes B périphériques. De plus, bloquer ce récepteur avec un antigène compétitif de BAFF inhibe la production d'anticorps ce qui en fait un des régulateur majeur de la maturation des lymphocytes B (Yan et al., 2001).

Suivant la force du signal via le BCR et l'activation de la voie Notch, les lymphocytes B transitionnels deviennent soit des lymphocytes B de la zone marginale (MZ) de la rate soit des lymphocytes B folliculaires (FO). Les MZ-B restent donc en périphérie de la rate et sont responsable de la réponse thymo-indépendante, c'est-à-dire de façon indépendante des lymphocytes T. A la rencontre d'un antigène exogène ils se différencient en plasmocytes, cellule sécrétrice d'anticorps, et produisent des anticorps IgM, possédant la chaîne lourde μ des Ig. Les FO-B, eux, sont responsables de la réponse thymo-dépendante. Les FO-B activés par la rencontre avec un antigène exogène, présente cet antigène aux lymphocytes T auxiliaires présents dans la zone T de la rate. Sous l'influence des cytokines sécrétées par ces lymphocytes T, les cellules peuvent se différencier en plasmocytes à IgM de haute affinité à courte durée de vie lorsque leur BCR est suffisamment spécifique de l'antigène ou former les centres germinatifs où ils passent par deux étapes concomitantes permettant l'augmentation de l'affinité de l'anticorps et la diversification des fonctions effectrices : l'introduction de mutations dans les régions variables des Ig, c'est l'hypermutation

somatique (SH) et le réarrangement de la région constante des Ig permettant le changement de classe des anticorps, c'est la commutation isotypique (CSR : Class Switch Recombination). Les cellules rencontrent alors des lymphocytes T auxiliaires folliculaires et des cellules présentatrices de l'antigène, les cellules dendritiques, pour leurs sélections. Si SH et CSR ont effectivement permis d'augmenter l'affinité de l'anticorps pour l'antigène, les lymphocytes B se différencient soit en plasmocytes à longue durée de vie soit en lymphocytes B mémoires. Si SH et CSR ont été désavantageux, les cellules sont dirigées vers l'apoptose (Weller and Descatoire, 2015).

Les plasmocytes sont généralement une durée de vie de quelques jours mais certains d'entre eux vont constituer une mémoire humorale et seront localisés dans des niches de la moelle osseuse où ils perdureront plusieurs années.

6. Les facteurs de transcription de la différenciation des lymphocytes B

La différenciation des lymphocytes B est orchestrée par différents facteurs de transcription (Figure I.19). Ainsi, EBF1 est un facteur de transcription essentiel de l'engagement des CLP vers la voie lymphocytaire B (Zandi et al., 2008). Son expression est induite par la voie de signalisation de l'IL7 (Dias et al., 2005). En collaboration avec E2A et FOXO1, il confirme l'engagement dans la voie de différenciation B (Mansson et al., 2012). Ces facteurs agissent ensemble pour activer la transcription de PAX5 (Decker et al., 2009; Lin et al., 2010). PAX5 permet l'inhibition des gènes spécifiques des autres lignées et ainsi le maintien de l'identité cellulaire (Cobaleda et al., 2007; Delogu et al., 2006). Il a un rôle prépondérant dans le développement précoce des lymphocytes B car il régule l'expression de SLP-65, protéine clé de la signalisation du pré-BCR (Schebesta et al., 2002).

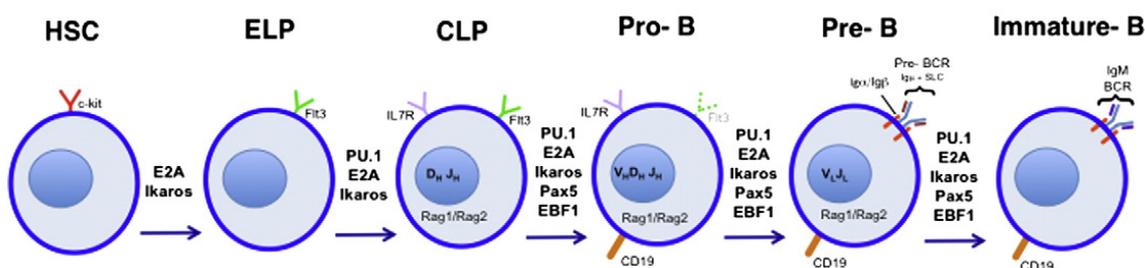


Figure I.19 : Régulation transcriptionnelle du développement précoce des lymphocytes B (De and Barnes, 2014)

B. Rôle spécifique d'Ikaros dans le développement des lymphocytes B

Ikaros, Helios et Aiolos sont exprimés au cours du développement précoce des lymphocytes B. Ils favorisent l'engagement des cellules dans la lignée B, assurent son bon déroulement et sont nécessaires aux fonctionnalités des lymphocytes B matures. Cependant, ces trois facteurs ne sont pas exprimés de la même façon au cours du développement B et malgré leur grande homologie de structure et leur motif commun de liaison à l'ADN, ils n'ont pas les mêmes fonctions.

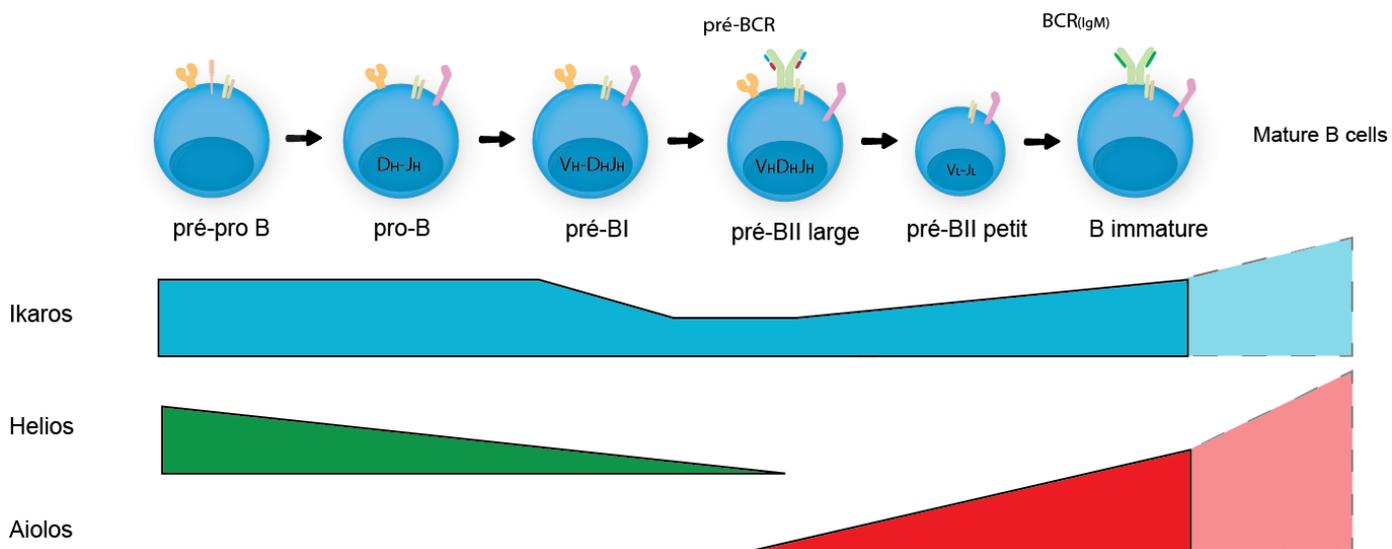


Figure I.20 : Expression d'Ikaros, Helios et Aiolos au cours du développement précoce des lymphocytes B. (adapté d'Immgen & (Heizmann et al., 2013))

Ikaros est toujours exprimé malgré quelques variations dans son niveau d'expression. En revanche Helios et Aiolos sont presque exclusifs, ils ne peuvent être retrouvés exprimés ensemble qu'au stade pré-BII large. Avant ce stade les cellules expriment Ikaros et Helios, après ce stade les cellules expriment Ikaros et Aiolos. Le stade pré-BII large apparaît alors comme transitionnel entre deux programmes de régulation au cours duquel Ikaros est le membre le plus exprimé et Helios et Aiolos sont beaucoup moins représentés.

Notre laboratoire, ainsi que d'autres, a montré qu'Ikaros joue un rôle prépondérant dans l'engagement et la différenciation des cellules B. Aiolos semble, lui, assurer la bonne fonction des cellules matures. Différents modèles murins ont permis d'appréhender les rôles cruciaux d'Ikaros dans l'engagement vers le lignage

lymphoïde, la différenciation des pro-B et l'expression du BCR. Dans la figure I.21 sont résumées les différentes études menées au cours de ces vingt dernières années sur le rôle d'Ikaros dans la lymphopoïèse B. Ces diverses mutations atteignent la lymphopoïèse B de trois façons. Les mutations entraînant l'absence de la protéine ou de sa capacité à lier l'ADN (*Ikzf1*^{null/null}, *Ikzf1*^{DN/DN}, *Ikzf1*^{plstc/plstic}) mènent à une absence d'engagement vers la lignée B et l'absence totale de lymphocytes B dans la moelle osseuse et la périphérie. D'autres mutations induisent des altérations de la cinétique de différenciation avec de légers blocages à certains stades et de la maturation des lymphocytes (*Ikzf1*^{L/L}, *Ikzf1*^{ΔF4/ΔF4}). Ces deux premiers types de mutations mènent au développement de lymphomes T. Enfin, la mutation conditionnelle *Ikzf1*^{Flox/Flox} et la délétion de du doigt de zinc 1 (*Ikzf1*^{ΔF1/ΔF1}) mènent à un blocage total de la développement précoce des lymphocytes B aux stades C' et B respectivement et une absence de lymphocytes B en périphérie. Dans ce dernier cas, il n'y a pas de développement d'hémopathie T. Ceci montre qu'Ikaros est donc nécessaire à l'engagement vers la lignée lymphoïde B et au développement précoce des lymphocytes B.

1. Dans l'engagement vers la lignée lymphoïde

Ikaros est indispensable à la mise en place du programme transcriptionnel permettant le maintien des MPP4/LMPP et leur engagement vers les CLP. En effet, Ikaros agit en régulant l'expression de gènes indispensables au développement lymphocytaire tels qu'IL7r, Flt3, Notch1 et Ccr9 dès le stade MPP4/LMPP (Ng et al., 2009). En contrôlant l'expression de Flt3, IKAROS restreint le destin des cellules vers le lignage lymphoïde (Yoshida et al., 2006). En se liant sur un enhancer particulier de PU.1, Ikaros permet l'arrêt de la boucle d'autorégulation de PU.1, menant à la diminution de son expression ce qui est nécessaire à l'engagement vers la lignée B (Zarnegar and Rothenberg, 2012). Ainsi l'absence d'IKAROS au stade MPP4/LMPP, empêche les cellules de s'engager vers les CLP et favorise leur différenciation vers la voie myéloïde (Wang et al., 1996; Yoshida et al., 2006). Il en va de même si Ikaros est présent mais muté dans son domaine de liaison à l'ADN (Georgopoulos et al., 1994). Pour marquer l'entrée dans la lignée B, les expressions de Ly6D et de RAG1 sont nécessaires (Inlay et al., 2009). Si l'un des deux ou les deux sont absents alors il n'y a pas de différenciation B. L'étude menée par Reynaud et al. a montré que les gènes *Rag1*, *Rag2* et *Dnnt* ne sont plus exprimés dans les cellules

Ikzf1^{-/-} (Reynaud et al., 2008). Ces résultats montrent qu'Ikaros joue un rôle essentiel dans l'engagement des cellules dans la lignée lymphoïde et le maintien de l'identité cellulaire.

2. Dans le développement précoce des lymphocytes B

Dans notre laboratoire, nous avons montré qu'une mutation hypomorphique d'Ikaros abolit la lymphopoïèse B fœtale et réduit drastiquement le compartiment des lymphocytes B dans la moelle osseuse après la naissance (Kirstetter et al., 2002).

Ikaros est nécessaire pour le réarrangement des *loci* des chaînes légères des Ig et l'expression du BCR. Puisque l'absence totale d'Ikaros conduit à une absence d'engagement dans la lignée B, il a fallu utiliser une délétion conditionnelle pour étudier son rôle dans la différenciation des lymphocytes B. Ainsi, dans le laboratoire, nous avons étudié le développement précoce des lymphocytes B lorsqu'Ikaros est absent à partir de la fraction A. Bien qu'Ikaros soit nécessaire pour l'expression des protéines RAG, une fois ces protéines exprimées, Ikaros ne semble pas nécessaire pour le réarrangement et l'expression de la chaîne lourde μ , les cellules atteignent la fraction C' et expriment le pré-BCR (Heizmann et al., 2013). En revanche, les cellules déficientes pour Ikaros restent bloquées à la fraction C'. Ikaros a plusieurs rôles prépondérants à cette étape du développement. Il antagonise la voie IL7/STAT5 par une liaison compétitive à l'ADN (Heizmann et al., 2020). Il en résulte une répression de l'expression des chaînes légères de substitution $V_{\text{préB}}$ et $\lambda 5$, l'expression des protéines RAG1 et RAG2 pour l'initiation du réarrangement des chaînes légères, une perte d'adhésion aux cellules stromales par répression des intégrines, et un arrêt du cycle cellulaire (Heizmann et al., 2013; Joshi et al., 2014; Sabbattini et al., 2001). Le passage de C' à D nécessite la perte de l'adhésion cellulaire et l'utilisation d'inhibiteur des intégrines dans ces cellules mènent à la mort cellulaire. Les cellules bloquées en fraction C' expriment de façon aberrante les intégrines permettant l'adhésion aux cellules stromales de la moelle osseuse qui sont le support de la différenciation jusqu'à ce stade et permettent l'expansion de cette fraction (Joshi et al., 2014).

3. Dans la maturation terminale des lymphocytes B

Les souris *Ikzf1*^{L/L} présentent un nombre réduit de lymphocytes B périphériques et ces cellules ont un seuil d'activation bas mais forment peu de centres germinatifs

lors une stimulation antigénique. Cependant les niveaux des différentes immunoglobulines dans le sérum sont normaux, à l'exception des IgG₃ qui sont très réduits voire inexistantes (Kirstetter et al., 2002). Les FO-B *Ikzf1*^{L/L} présentent une activation constitutive de la voie de signalisation du BCR et principalement une augmentation de la phosphorylation des protéines ERK1/2 et p38. Ikaros serait donc un régulateur de l'activation des FO-B en régulant la transcription des gènes codants pour les protéines ERK1/2 et p38 (Heizmann et al., 2016). Ikaros prévient également l'hyperactivation des FO-B en contrôlant l'expression de TLR9 (Toll-like receptor 9) dont la voie de signalisation restreint la survie des lymphocytes B auto-réactifs activés (Schwickert et al., 2019).

Ikaros apparaît également comme un régulateur majeur de la commutation isotypique en régulant l'accessibilité de la chromatine pour la protéine AID (Activation-induced cytidine deaminase) et en se liant directement aux régions régulatrices des gènes des immunoglobulines (Sellars et al., 2009, 2011).

4. Le rôle de suppresseur de tumeurs

a) Dans les T-ALL

L'absence d'Ikaros mène au développement quasiment systématique d'hémopathies suggérant un rôle suppresseur de tumeur. Bien que son rôle dans la lymphopoïèse B ne soit plus à démontrer, chez la souris, l'absence d'Ikaros aux différents stades de la différenciation mène au développement de leucémies aiguës lymphoblastiques T (T-ALL). Dans notre laboratoire, nous avons montré que ces pathologies sont liées à une absence de régulation de la voie Notch1 qui apparaît comme constitutivement active (Dumortier et al., 2006). En effet, dans les cellules *Ikzf1*^{L/L}, la plupart des cellules leucémiques ont une délétion du promoteur de Notch1 favorisant l'utilisation d'un promoteur caché, ou cryptique, menant à l'expression de la partie intracytoplasmique et active de la protéine Notch1, N_{ICD} (Jeannet et al., 2010). L'activation constitutive de cette voie de signalisation promeut la transformation leucémique, la survie et la prolifération des cellules (Ferrando, 2009).

b) Dans les B-ALL

Chez l'homme, l'absence d'Ikaros au cours de l'hématopoïèse ne mène que rarement au développement de T-ALL (Marçais et al., 2010). En revanche, différents types de mutation d'Ikaros ont été retrouvés dans des leucémies aiguës lymphoblastiques B (B-ALL) pédiatriques (15%) et de l'adulte (40-50%). Ces mutations sont de différentes natures. Il peut s'agir d'une délétion totale du chromosome 7 portant le gène *Ikzf1* chez l'homme, c'est ce que l'on appelle une monosomie 7 (Heerema et al., 2004). Des délétions de certains exons d'*Ikzf1* sont également décrites : $\Delta 4-7$, est une délétion des exons 4 à 7 menant à l'expression de l'isoforme IK6 (Figure I.5) connue pour agir comme dominant négatif ; $\Delta 2-7$ qui mène à la délétion de l'exon 2 porteur de l'ATG et conduisant très probablement à un allèle nul. Dans notre laboratoire nous avons décrit la délétion de la partie 5' de l'exon 1, contenant le promoteur du gène et menant à la perte de son expression (Morel, Deau et al., 2019). Le dernier type de mutation est la mutation ponctuelle conduisant à l'insertion d'un codon STOP prématuré et l'expression d'une protéine tronquée, instable. Ces mutations somatiques d'Ikaros dans les B-ALL sont des événements secondaires de la transformation leucémique mais sont un marqueur de pronostic défavorable pour le patient.

Chez la souris, l'haplo-insuffisance d'*Ikzf1* combinée à la présence de la protéine de fusion BCR-Abl mènent au développement de B-ALL de manière accélérée par rapport aux souris n'ayant pas de mutation d'Ikaros (Virely et al., 2010).

L'absence d'une protéine Ikaros fonctionnelle empêche la répression de la voie de signalisation IL7/STAT5 ce qui contribue à favoriser la survie et la prolifération. Dans le cas d'une mutation entraînant l'activation constitutive de STAT5, les souris développent des T-ALL et B-ALL (Katerndahl et al., 2017). Les protéines aberrantes telles que BCR-Abl, résultant de l'expression de gènes de fusion et qui sont responsables de la transformation leucémique, activent la voie de STAT5. La réexpression de la forme WT d'Ikaros dans ces cellules mène à la réduction de la croissance leucémique (Schjerven et al., 2017)

5. Autres pathologies

Lorsque les mutations d'Ikaros sont germinales, elles conduisent peu fréquemment au développement de leucémies mais favorisent le développement d'immunodéficiences communes variables (CVID) (Eskandarian et al., 2019; Kuehn et al., 2016). Cette pathologie est caractérisée par une hypogammaglobulinémie c'est-à-dire une forte réduction des anticorps sériques, l'absence de plasmocytes et un nombre normal de lymphocytes B. Cela peut résulter d'un dysfonctionnement au niveau de l'activation des lymphocytes B naïfs par les lymphocytes T auxiliaires, de façon intrinsèque ou extrinsèque aux lymphocytes B, ou au niveau de la commutation isotypique dans les centres germinatifs. Plusieurs défauts génétiques peuvent mener au développement de CVID. Dans le cas d'une absence de protéine Ikaros fonctionnelle, le CVID est associé à une réduction drastique du nombre de lymphocytes B.

Une mutation ponctuelle dans le DZ3 d'Ikaros et le développement d'un Lupus systémique érythémateux (SLE) a récemment été mise en évidence (Van Nieuwenhove et al., 2018). Les lymphocytes B périphériques montrent une augmentation de l'expression des molécules impliquées dans l'activation et une diminution de celles impliquées dans la répression de l'activation. Dans un modèle murin exprimant, après l'engagement des cellules vers le lignage B, une forme dominant-négatif d'Ikaros, les chercheurs ont montré la production d'anticorps auto-réactifs menant au développement de pathologies auto-immunes (Wojcik et al., 2007).

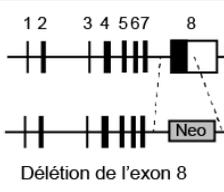
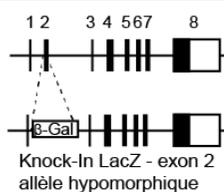
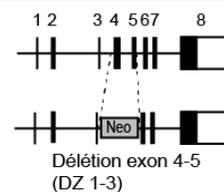
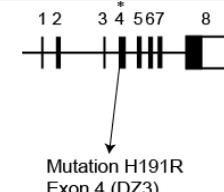
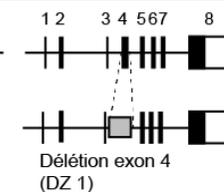
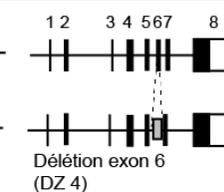
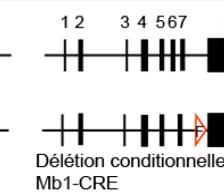
Lignée murine	<i>Ikzf1</i> ^{null/null}	<i>Ikzf1</i> ^{L/L}	<i>Ikzf1</i> ^{DN/DN}	<i>Ikzf1</i> ^{Plstc/Plstc}	<i>Ikzf1</i> ^{ΔF1/ΔF1}	<i>Ikzf1</i> ^{ΔF4/ΔF4}	<i>Ikzf1</i> ^{Flox/Flox}
Mutation	 Délétion de l'exon 8	 Knock-In LacZ - exon 2 allèle hypomorphique	 Délétion exon 4-5 (DZ 1-3)	 Mutation H191R Exon 4 (DZ3)	 Délétion exon 4 (DZ 1)	 Délétion exon 6 (DZ 4)	 Délétion conditionnelle exon 8 Mb1-CRE
Phénotype B	Lymphocytes B absents	Léger blocage de la différenciation B B activés CSR perturbé	Lymphocytes B absents	Lymphocytes B absents	Blocage à la fraction B Réduction de la fraction C'	Réduction des progéniteurs B Augmentation de la fraction C'	Blocage de la différenciation en fraction C'
Phénotype T	Lymphocytes T foetaux absents T activés post-natal Biais vers CD4+	Composition cellulaire normale du thymus Thymocytes et lymphocytes T activés	Lymphocytes T absents	Lymphocytes T absents	T foetaux présents Légère réduction de la cellularité thymique	T foetaux absents Réduction de la cellularité thymique	Normal
Phénotype hématopoï- tique	Légère réduction de l'activité des HSC NK et pDCs absents	NK et pDCs absents Augmentation des précurseurs neutro- philes dans le foie foetal	Forte réduction de l'activité des HSC Réduction des progéniteurs eryth- roïdes NK et pDCs absents	Perte des LT-HSC Augmentation des progéniteurs GM Anémie foetale fatale	NK et pDCs présents	NK et pDCs absents	
Hémopathie T	Une fraction des <i>Ikzf1</i> ^{null/+} développent une hémopathie T	Lymphomes thymiques à 10 mois	Lymphomes thymiques à 4 mois	Hémopathie T à 4 mois	absence	Lymphomes thymiques à 10 mois	absence
Références	(Avitahl et al., 1999; Wang et al., 1996; Winandy et al., 1999; Yoshida et al., 2006)	(Dumortier et al., 2006; Heizmann et al., 2016; Kirstetter et al., 2002; Sellars et al., 2009)	(Georgopoulos et al., 1994; Winandy et al., 1995)	(Mantha et al., 2007; Papathanasiou et al., 2003)	(Schjerven et al., 2013)	(Schjerven et al., 2013)	(Heizmann et al., 2013; Schwickert et al., 2014)

Figure I.21: Résumé des phénotypes observés dans les différentes lignées de souris portant une mutation au locus d'*Ikzf1*
(Adapté de Marke et al., 2018)

C. Rôles d'Helios et Aiolos dans le développement B

1. Helios

Helios est présent principalement dans le compartiment des cellules souches hématopoïétiques et des progéniteurs multipotents (Kelley et al., 1998). Dans l'ensemble des cellules matures, seuls les lymphocytes T régulateurs présentent une expression d'Helios significative (Hahm et al., 1998; Kim et al., 2015; Takatori et al., 2015).

a) Dans le développement des lymphocytes B

Au cours de la différenciation des lymphocytes B, l'expression d'Helios diminue jusqu'à ne plus être détectable avant la fin du développement précoce. Cela laisse penser que son rôle est négligeable cependant il semblerait que cette régulation de son expression soit nécessaire au développement normal et à la fonction des lymphocytes B. L'absence d'Helios n'empêche pas l'engagement dans le lignage B ni la différenciation des lymphocytes. En revanche, elle contribue à diminuer le nombre de CLP et donc le nombre absolu de lymphocytes B périphérique (données non publiées du laboratoire). Une surexpression d'Helios au cours du développement précoce des lymphocytes B entraîne la production d'auto-anticorps, la diminution des MZ-B et le développement de lymphome (Dovat et al., 2005). Comme nous le verrons plus tard, ce phénotype est similaire à celui décrit dans les mutants Aiolos null.

b) Rôle suppresseur de tumeur

Différentes isoformes d'Helios ont été retrouvées dans les leucémies T adultes (ALT). La plupart de ces isoformes ne possèdent pas le domaine de liaison à l'ADN et ont un rôle dominant négatif prouvant le rôle de suppresseur de tumeur d'Hélios dans les ALT (Asanuma et al., 2013). La surexpression d'Helios dans la lignée leucémique T, Jurkat, montre un arrêt de la prolifération et une reprise du processus de

différenciation, confirmant le rôle du suppresseur de tumeur d'HELIOS dans les T-ALL (Li et al., 2018).

Dans les cellules souches leucémiques myéloïdes la délétion d'HELIOS favorise la différenciation et l'apoptose (Park et al., 2018).

Enfin, des altérations d'HELIOS ont été retrouvées dans des sous-types « Low Hypodiploid » de B-ALL (Iacobucci and Mullighan, 2017). Les cellules comportent entre 32 et 39 chromosomes et ont fréquemment des mutations de TP53.

2. Aiolos

En 1997, Morgan et collaborateur ont montré qu'Aiolos était un partenaire privilégié d'Ikaros, ces deux protéines pouvant former des hétérodimères, mais également des homodimères, grâce au domaine C-terminal (Morgan et al., 1997). Dans cette étude, les chercheurs ont également montré qu'Aiolos est un meilleur activateur de la transcription qu'Ikaros et que la co-expression des deux facteurs avait un effet intermédiaire sur la transcription.

a) Dans le développement des lymphocytes B

L'expression d'Aiolos dans le développement précoce des lymphocytes B est induite par la voie de signalisation du pré-BCR (Schwickert et al., 2014; Thompson et al., 2007). Ce qui signifie qu'Aiolos est exprimé au cours du développement précoce des lymphocytes B à partir de la fraction C'. Dès le début, Aiolos participe activement à la répression des gènes des chaînes légères de substitution *Igll1* ($\lambda 5$) et *Vpreb* avec Ikaros (Thompson et al., 2007). L'implication d'un hétérodimère Ikaros/Aiolos dans la relocalisation de *locus* de $\lambda 5$ vers l'hétérochromatine a été suggérée dans une étude récente mettant en évidence le rôle de la protéine OBF1 dans ce processus (Karnowski et al., 2008). OBF1 (Oct Binding Factor 1) est un co-activateur transcriptionnel spécifique des lymphocytes B. Les souris déficientes pour cette protéine présentent une altération de la maturation et de l'activation des lymphocytes B (Nielsen et al., 1996). Cependant il est exprimé, comme Aiolos à partir du stade pré-B et les souris doubles mutantes pour Aiolos et OBF1 présentent une forte réduction du nombre de cellules au stade Immature ainsi que dans la périphérie. Cela est dû à

un léger blocage au stade pré-BII en raison de l'absence de répression de l'expression de $\lambda 5$.

Aiolos tient également un rôle important dans l'arrêt du cycle cellulaire en réprimant, avec Ikaros, l'expression de c-Myc, ce qui lève la répression de l'expression de l'inhibiteur de cycle cellulaire p27 et qui inhibe l'expression de la cycline D3, cycline spécifique du passage G1/S (Ma et al., 2010). A partir de là, l'expression d'Aiolos est croissante et atteint son maximum dans les cellules B matures.

Dans le cas d'une absence d'Aiolos au cours de la différenciation des lymphocytes B, le développement précoce n'est pas bloqué comme dans le cas de l'absence d'Ikaros, mais sa cinétique est altérée. Il semble qu'il y ait une accélération de la progression de la fraction A vers la fraction B, et de la fraction C' à la fraction D. En revanche, les stades de maturation ultérieurs semblent compromis (Wang et al., 1998)

b) Dans la maturation des lymphocytes B

De plus, Aiolos a été décrit comme régulateur majeur de la maturation des lymphocytes B vers un stade effecteur. Les souris mutantes pour Aiolos présentent une augmentation de la quantité d'auto-anticorps dans leurs *sera*, ainsi qu'une augmentation des isotypes IgE, IgG1 et IgG2a sans immunisation. Les populations recirculantes, de la zone marginale ou de la cavité péritonéale sont réduites. De plus, les sujets âgés développent des lymphomes B et des syndromes de lymphoprolifération T. Les souris malades ne présentant pas de syndrome lymphoprolifératif pourraient, en fait, avoir développé une maladie auto-immune, sachant que des auto-anticorps dirigés contre le rein sont détectables à partir de 5 semaines (Wang et al., 1998).

Aiolos est aussi régulateur du destin des cellules matures en favorisant leur engagement vers la voie folliculaire et donc la réponse à l'antigène, en collaboration avec la tyrosine kinase de Bruton, Btk, et le récepteur du complément, CD21 (Cariappa et al., 2001). Il semble également qu'Aiolos soit indispensable à la génération des plasmocytes à longue durée de vie (Cortés and Georgopoulos, 2004).

c) Rôle suppresseur de tumeur

Les souris âgées déficientes pour Aiolos développent des lymphomes B et leurs *sera* contiennent des taux anormaux d'auto-anticorps permettant probablement le développement de maladie auto-immune (Wang et al., 1998).

L'inactivation d'*Ikzf3* est retrouvée dans la forme « presque haploïde » des B-ALL (Iacobucci and Mullighan, 2017). Les cellules comportent entre 24 et 31 chromosomes et ont souvent des mutations activatrices de Ras.

Au contraire, la surexpression d'Aiolos dans les leucémies lymphocytiques chroniques B (B-CLL) semble associée à un meilleur pronostic (Duhamel et al., 2008). Cette surexpression est due à l'augmentation de la transcription d'*IKZF3*, cible de la voie NF- κ B qui est constitutivement active dans les B-CLL. AIOLOS protège les cellules de l'apoptose en régulant les protéines de la famille de Bcl2 (Billot et al., 2011).

d) Hémopathies liées à Aiolos

Le myélome multiple (MM) est caractérisé par la prolifération excessive des plasmocytes dans la moelle osseuse. Bien qu'il n'en soit pas la cause, Aiolos agit comme agent anti-apoptotique en collaboration avec Blimp-1 pour promouvoir la survie des plasmocytes dans les Myélomes Multiples (Hung et al., 2016). Le traitement avec le Lénalidomide menant à la dégradation d'Aiolos permet ainsi d'interrompre la voie Aiolos/Blimp-1 dans cette pathologie.

La surexpression d'Aiolos prédispose les lymphocytes B circulant à une différenciation en plasmablastes qui se concrétise par la sur-activation de la voie de signalisation de BAFF, conduisant à la sécrétion d'anticorps anti-dsDNA et l'inflammation des tissus dans le lupus systémique érythémateux (SLE) (Nakayama et al., 2017). L'absence d'Aiolos mène également au développement de cette pathologie. En revanche, les souris doubles mutantes pour Aiolos et OBF1 ne développent plus de réaction auto-immune et de SLE (Sun et al., 2003). Ceci montre l'importance de la présence d'Aiolos et sa juste régulation pendant la lymphopoïèse B.

IV. PROBLEMATIQUE

Ikaros, Helios et Aiolos ont une forte homologie structurelle et un domaine de liaison à l'ADN composé de doigts de zinc dont les propriétés de liaison sont tout à fait identiques. Ces trois protéines sont retrouvées dans les *foci* de PC-HC et partagent leurs partenaires protéiques ainsi que leurs modes de régulations de la transcription de leurs gènes cibles. Cependant, au cours du développement des lymphocytes B, Ikaros semble avoir des propriétés non redondantes pour le passage pré-BII/Immature B, alors que les deux autres facteurs sont exprimés au stade pré-BII et qu'Aiolos semble être également important pour certains aspects de ce processus. L'absence d'Ikaros ou d'Aiolos au cours du développement précoce des lymphocytes B mènent à des phénotypes très différents ce qui renforce l'idée que malgré les fortes similitudes, ces facteurs ont des spécificités fonctionnelles qui leurs sont propres. Grâce à leur domaine C-terminal, ces protéines peuvent former des homodimères et des hétérodimères. Il est donc envisageable que les différents dimères aient des fonctions particulières. Or, les dimères formés par les facteurs de transcription de la famille d'Ikaros n'ont jamais été étudiés. De plus, l'absence d'Ikaros au cours du développement B se traduit par l'absence de l'homodimère d'Ikaros mais également celle des hétérodimères Ikaros/Helios et Ikaros/Aiolos (Figure P.1). Il est donc impossible de déterminer quel dimère est réellement nécessaire au processus de différenciation des lymphocytes pré-B. C'est également le cas dans le phénotype Aiolos null.

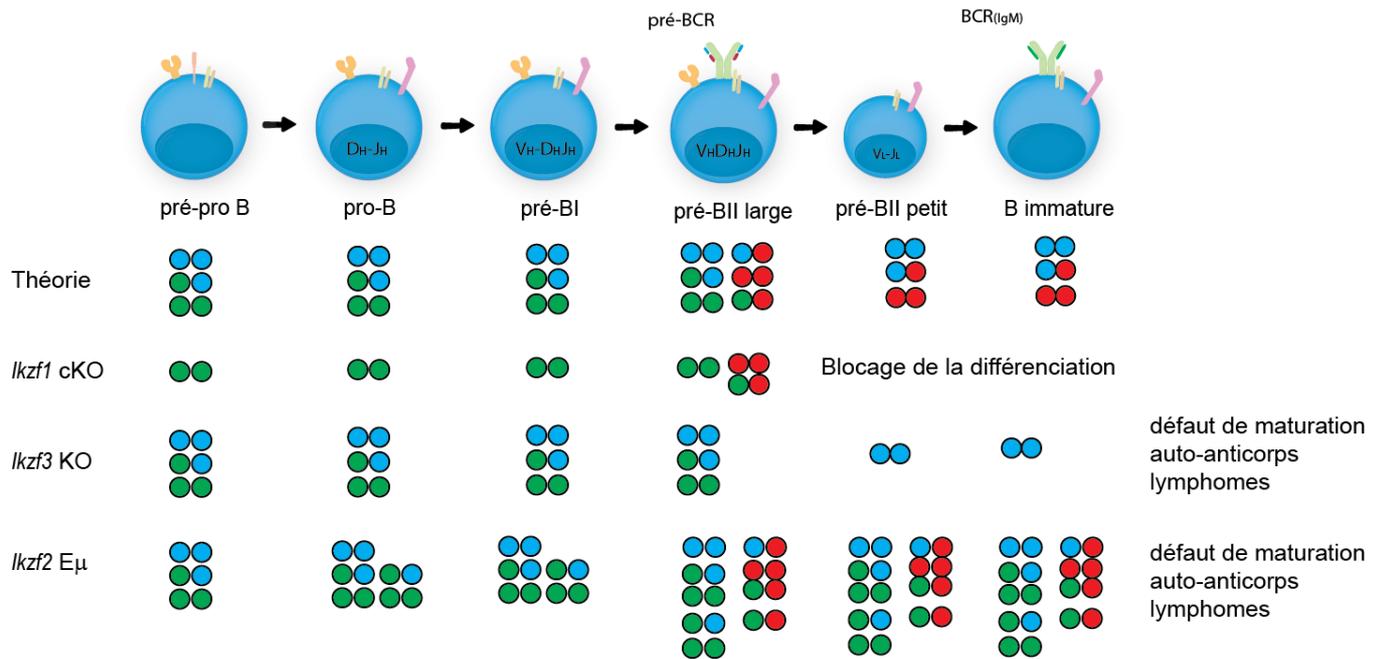


Figure I.IV.1 Représentation théorique de la distribution des dimères de la famille d'Ikaros au cours du développement précoce des lymphocytes B. Ikaros est représenté en bleu, Helios en vert et Aiolos en rouge

V. OBJECTIFS

Ce travail est centré sur la caractérisation des spécificités fonctionnelles des facteurs de transcription de la famille d'Ikaros en répondant à ces différentes questions :

- Les différents facteurs ont-ils des motifs de liaison à l'ADN spécifiques ?
- Les différents dimères ont-ils des gènes cibles spécifiques ?
- Comment les régulent-ils ?

Matériels & Méthodes

I. Culture cellulaire

A. J558L-IL7 : Production du surnageant IL-7

Don de M. Reth – (Winkler et al., 1995)

Ces cellules de myélomes dérivées de souris BALB/c ont été transfectées avec l'ADN complémentaire de l'interleukine 7 murine grâce à un vecteur BMG néo (*Bacillus subtilis* genome portant un gène de résistance à la néomycine). Ces cellules sont cultivées dans du DMEM (Dulbecco's Modified Eagles's Medium) supplémenté de 5% de sérum de veau fœtal.

Le surnageant contenant l'interleukine 7 (IL-7) est récupéré après 10 jours de culture des cellules sans changement de milieu. Ce surnageant est filtré à l'aide d'une unité de filtration Stericup 500mL avec un filtre 0,22µm (Millipore). Il est ensuite soit utilisé directement comme supplément au milieu de culture des cellules BH1, soit congelé.

B. Les cellules BH1

Les cellules BH1 sont issues de la moelle osseuse d'une souris portant un knock-out conditionnel pour *Ikzf1* dans la lignée B (Heizmann et al., 2013). Elles sont cultivées dans du milieu IMDM (Iscove Modified Dulbecco Media) supplémenté avec 10% de sérum de veau fœtal tet-free (dutscher), 10% de surnageant IL-7, 1% de Pénicilline/streptomycine et 50µg/ml de gentamycine. La concentration cellulaire initiale est de 300 000 cellules /ml et les cellules sont passées au 1/6 tous les deux jours.

C. Les cellules COS-1

Les cellules COS-1 sont des fibroblastes dérivés de rein de singe. Elles sont cultivées dans un milieu DMEM (1g/L glucose) supplémenté de 5% FCS et de gentamycine (20µg/ml). Ces cellules ont été utilisées pour la production de protéines destinées aux expériences de retard sur gel.

D. HEK293T Phoenix Eco

Ces cellules sont dérivées de la lignée HEK 293T, une lignée dérivée de cellules humaines embryonnaires de rein. Elles sont utilisées pour la production simplifiée de rétrovirus car elles produisent déjà les protéines d'enveloppe et les protéines Gag (Group Antigens) et POL (reverse transcriptase). Pour la production rétrovirale il faut transfecter ces cellules avec un plasmide contenant l'information génétique désirée encadrée de LTR. Ces cellules sont maintenues en culture dans un milieu DMEM (4,5g/L glucose) +10% SVF inactivé + 1mM Pyruvate de sodium + Peniciline/Streptomycine

II. Création des lignées cellulaires

Dans ce travail j'utilise une technique de complémentation bi-moléculaire de fluorescence couplée au système tétracycline-ON dans les cellules BH1.

A. La complémentation bi-moléculaire de fluorescence (BiFC)

En collaboration avec le laboratoire de Simona Sacconi (IRCAN, Nice)

La complémentation bi-moléculaire de fluorescence (BiFC : Bi-molecular Fluorescence Complentation) repose sur le principe que deux moitiés non fluorescentes d'une protéine fluorescente sont capables de reconstituer la fluorescence de la protéine initiale. Dans ce travail, j'utilise la protéine fluorescente Venus qui est divisée en deux moitiés non fluorescentes V1 et V2. Les fragments V1 et V2 sont liés aux facteurs de transcription de la famille d'Ikaros grâce à un linker composé d'un enchaînement de sérines et de glycines (SGSGGGGSGGGGSSG) permettant le mouvement. Ainsi, lors de la dimérisation des facteurs de transcription les deux parties V1 et V2 pourront se lier et restituer la fluorescence de Venus, c'est la complémentation.

De plus, nos collaborateurs ont développé un anticorps monoclonal capable de reconnaître un épitope présent sur la protéine complète Venus et sur la protéine Venus reconstituée par V1 et V2 mais qui est absent de V1 et de V2 lorsque les fragments

sont isolés. Cet anticorps permet alors l'immunoprécipitation spécifique des dimères V1/V2 comme schématisé dans la figure MII.1.

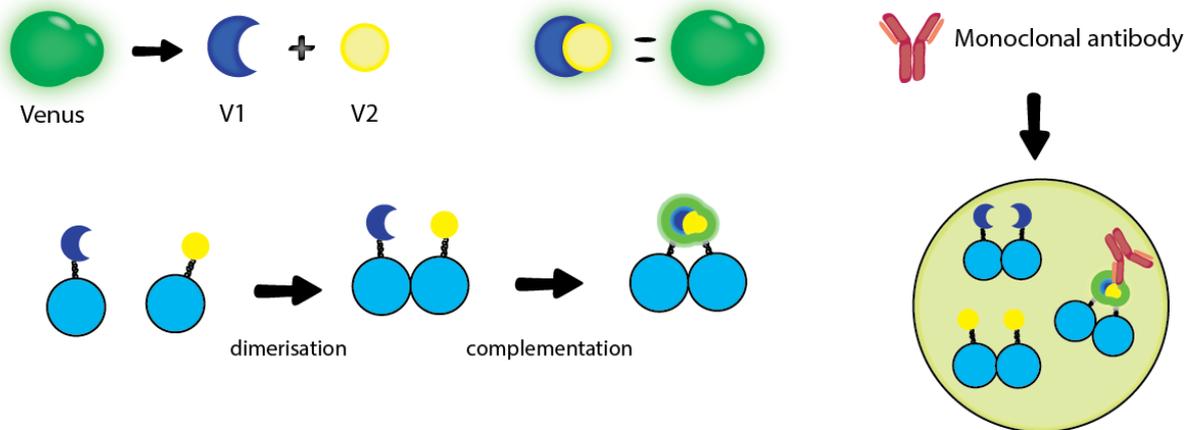


Figure MII.1 La complémentation bi-moléculaire de fluorescence

La protéine Venus, fluorescente, est divisée en deux moitiés V1 et V2, non fluorescente. Lorsque V1 et V2 se trouvent à proximité l'un de l'autre, ils sont capables de restituer la fluorescence de Venus. Ainsi, V1 et V2 sont liés aux facteurs de transcription grâce à un Linker puis, lorsque les facteurs se dimérisent, cela permet à V1 et V2 de se compléter et restituer la fluorescence de Venus. De plus, l'anticorps monoclonal α Venus permet l'immunoprécipitation spécifique des dimères formés par V1 et V2.

B. Le système tétracycline-ON (tet-ON)

Le système tet-On permet un contrôle précis et efficace de l'activation de la transcription des gènes situés en aval d'un promoteur TRE (Tetracycline Responsive element). Ainsi en présence de doxycycline, un analogue de la tétracycline, les facteurs de transcription de la famille d'Ikaros seront exprimés dans les cellules BH1.

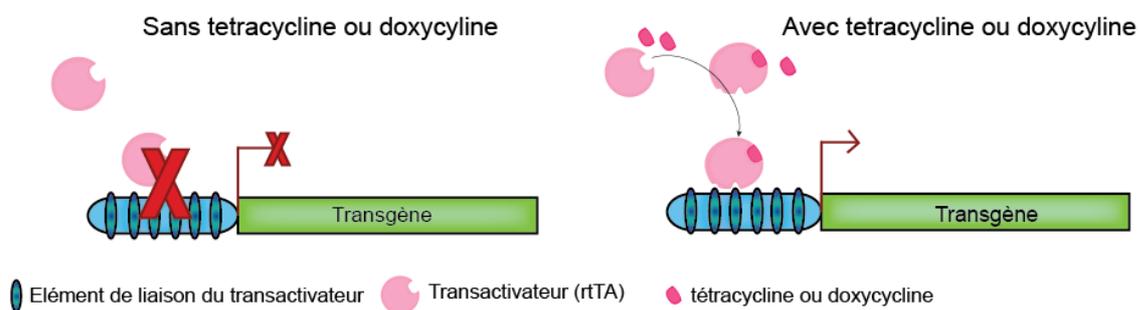


Figure MII.2 Le système tétracycline-ON (tet-ON)

An absence de tétracycline ou de doxycycline, le transactivateur rtTA n'est pas capable de se lier à ses éléments de liaison au niveau du promoteur. Il n'y a pas d'activation de la transcription. En présence de tétracycline ou de doxycycline, le rtTA change de conformation et est capable de se lier au promoteur et la transcription est alors possible.

La transposition de la BiFC couplée au système tet-ON dans les cellules BH1 s'est faite par transduction rétrovirale et antibiosélection. La première étape a été la décision

de positionner les fragment V1 et V2 en partie N-terminale ou en partie C-terminale des protéines.

C. Clonage moléculaire

La première étape a été de lier les parties N-ter et C-ter, respectivement V1 et V2, de la protéine Venus aux parties N-ter ou C-ter d'Ikaros. J'ai réalisé ces constructions par PCR en amplifiant d'un côté l'ADNc d'Ikaros et de l'autre le linker SGSGGGSGGGGSSG avec soit V1 soit V2 de sorte que les deux PCR aient un domaine d'homologie. Ceci permet de réaliser une troisième PCR d'alignement en utilisant les amorces externes des deux PCR précédentes. Les quatre constructions ont été clonées par restriction enzymatique/ligation dans un vecteur pMSCV.

Tableau M.II-1 Amorces pour le clonage Ikaros-V1 et Ikaros-V2

Nom de l'amorce	Séquence 5' → 3' domaine d'homologie	Fragment amplifié
IK-V1/V2 Frag1 FOR	CCGGAATTAGATCTCTCGAGCCACCATGGATG	Ikaros
IK-V1/V2 Frag1 REV	GAGCCACCGCCACCGGATCCGCT CAGGTGGTAACGATGCTCCC	
IK-V1 Frag2 REV	GGAATTTACGTAGCGGCCGCTTACTGCTTGTC	Linker-V1
IK-V1/V2 Frag2 FOR	AGCATCGTTACCACCTG AGCGGATCCGGTGGCGGT	Linker-V2
IK-V2 Frag2 REV	GGAATTTACGTAGCGGCCGCTTACTTGTACAG	

Tableau M.II-2 Amorces pour le clonage de V1-Ikaros et V2-Ikaros

Nom de l'amorce	Séquence 5' → 3' domaine d'homologie	Fragment amplifié
V1-IK Frag1 FOR	CCGGAATTAGATCTCTCGAGATGGTGAGCAAGGGCGAG	V1-Linker
V1/V2-IK Frag1 REV	TGACCCTCATCGACATCCATT TCCGGAGGACCCACCACCTCCAG	
V2-IK Frag1 FOR	CCGGAATTAGATCTCTCGAGATGAAGAACGGCATCAAGGC	V2-Linker
V1/V2-IK Frag2 FOR	GAGGTGGTGGGTCTCCGGA ATGGATGTCGATGAGGGTCAAG	Ikaros
V1/V2-IK Frag2 REV	GGAATTTACGTAGCGGCCGCTCAGCTCAGGTGGTAACGATGCTCC	

Après avoir testé les différentes combinaisons, j'ai, sur le même principe, créer les constructions Aiolos-V1, Aiolos-V2, Helios-V1 et Helios-V2. J'ai ensuite cloné les six constructions possédant les fragments V1 et V2 en partie C-ter dans les plasmides p259 comportant un promoteur du système tétracycline et un gène de résistance à un antibiotique sous le contrôle d'un promoteur constitutif. Ainsi, le gène de résistance sera toujours exprimé permettant la sélection des cellules transduites.

D. Production rétrovirale

Les cellules Phoenix Eco sont mises en cultures dans une boîte de culture cellulaire de 60x15mm à 200 000 cellules/ml dans 10mL. Le lendemain, le milieu de culture est totalement changé par 5mL de milieu frais. À côté, deux tubes de 15mL sont préparés avec 1,5 mL d'Opti-MEM sans antibiotiques. Dans le premier, 50µl de Lipofectamine 2000 (Invitrogen) sont pipetés et, dans le second, 10µg d'ADN plasmidique. Après 5 minutes d'incubation à température ambiante, l'opti-MEM contenant l'ADN est ajouté au goutte-à-goutte dans celui contenant la lipofectamine. Après 20 minutes d'incubation à température ambiante, les complexes lipofectamine/ADN sont ajoutés au goutte-à-goutte sur les cellules. Les boîtes de cultures sont alors replacées dans l'incubateur à 37°C 5%CO₂. Après 24 heures, le milieu de culture est totalement remplacé par 5mL du milieu des cellules qui seront transduites. Le surnageant viral est alors récolté 24 heures et 48 heures après. Il est soit immédiatement utilisé pour une transduction soit congelé dans par aliquots de 500µl à -80°C.

E. Transduction des cellules

500 000 cellules BH1 sont pipetées dans un eppendorf 1,5mL et culotées par centrifugation 5 minutes à 1200 rpm. Le culot est alors resuspendu dans 500µl de suspension virale supplémentée de polybrène (8µg/ml) puis déposé dans un puit d'une plaque de culture 24 puits. La plaque est centrifugée pendant 10 minutes à 2000rpm à température ambiante puis placée dans un incubateur à 37°C 5% CO₂. Après une heure d'incubation, 1mL de milieu de culture est ajouté aux cellules. Le lendemain, le surnageant est jeté et les cellules resuspendue dans 3mL de milieu neuf.

F. Sélection par les antibiotiques

Les cellules transduites avec le rétrovirus permettant l'insertion de l'ADNc du transactivateur sont sélectionnées avec la généticine (G418) à une concentration finale de 1mg/mL. Les cellules exprimant les facteurs des transcription liés au fragments V2 ont été sélectionnées avec de la puromycine (Invitrogen) à une

concentration finale de 3µg/ml. Les facteurs de transcription liés aux fragments V1 sont co-exprimés d'abord avec le gène de résistance à l'hygromycine (Sigma, 400µg/ml) ou avec le gène de résistance à la blasticidine (Invitrogen, 20µg/ml).

III. Immunofluorescence

La complémentation de fluorescence par V1 et V2 peut se visualiser directement au microscope à fluorescence.

Les cellules traitées avec la doxycycline sont récupérées et lavées deux fois dans du milieu de culture sans sérum et comptées. 70 000 cellules sont déposées dans une goutte de 100µl sur une lame puis laissées 20 minutes à température ambiante dans le noir. Les cellules sont alors déposées sur la lame et le surnageant est ôté avec une pipette puis lavées deux fois avec 200µl de tampon de lavage (PBS 1X, BSA1%, Tween 0.05%). Les cellules sont alors fixées avec 100µl de paraformaldéhyde (PFA) 4% pendant 10 minutes dans le noir. La PFA est ensuite enlevée et les cellules lavées comme précédemment. Puis les cellules sont perméabilisées avec 100µl de tampon de perméabilisation (PBS1X, tritonX-100 0.5%) pendant 10 minutes dans le noir puis lavées comme précédemment. Préalablement à un marquage intracellulaire, les cellules sont incubées pendant 10 minutes à un tampon de blocage (tampon de lavage, 5% de sérum dépendant de l'origine de l'anticorps pour le lapin- α Ikaro, on utilise un sérum de lapin). Les cellules sont lavées comme précédemment puis incubées pendant 1 heure avec l'anticorps primaire (α Ikaro) dilué au 1/100^{ème} dans du tampon de lavage. Les cellules sont lavées trois fois puis incubées avec 100µl d'anticorps secondaire (anti-lapin AF568) au 1/400^{ème} dans du tampon de lavage pendant 45 minutes. Après trois lavages les lames sont légèrement séchées puis 7µl de DAPI (50-100ng/ml) dilué dans de la solution de montage (Glycérol 75%, PBS1X 20%, Propylgalate 5%) sont appliqués sur une lamelle qui sera déposée sur les cellules. La lamelle est ensuite scellée grâce à du vernis à ongle. Les lames sont conservées dans une boîte à 4°C jusqu'à l'observation au microscope à fluorescence.

IV. Immunoprécipitation (IP)

Pour vérifier la capacité de l'anticorps α Venus à précipiter spécifiquement des dimères j'ai réalisé des expériences d'immunoprécipitation.

Les cellules traitées 24 heures à la doxycycline sont récupérées et lavées avec du PBS1X deux fois puis comptées au bleu de trypan. Le culot est alors resuspendu dans un tampon de lyse (TrisHCl pH7.5 20mM, NaCl 150mM, EDTA pH8 5mM, NaPIP 5mM, NaVO₄ pH10 1mM, NaPO₄ pH7.6 20mM, β -glycerophosphate 3mM, NaF 10mM, TritonX-100 1%, Inhibiteur de protéase (Roche) 1X, Inhibiteur de phosphatase 1X). Le lysat est alors passé au sonicateur (Bioruptor, Diagenode) pendant 15 minutes en position haute selon des cycles alternant 20 secondes ON et 15 secondes OFF. Le lysat est ensuite centrifugé pendant 10 minutes à 14000rpm à 4°C pour éliminer les débris puis les protéines sont dosées. Après une étape de clarence avec les billes de sépharose protéines G pendant 1 heures, le surnageant est récupéré, 100 μ g de protéines sont mis de côté pour l'INPUT et le reste est mis en incubation sur la nuit avec de nouvelles billes et l'anticorps. Le lendemain, les échantillons sont centrifugés, le surnageant ôté, un aliquot de même volume que l'input est mis de côté et les billes sont lavées 3 fois avec du tampon de lyse. Du tampon de charge (4X : TrisHCl pH6.6 200mM, SDS 8%, Glycerol 40%, EDTA 50mM, bleu de bromophenol 0.08%, β -mercaptoéthanol 578mM) est alors ajouté aux billes, à l'INPUT et au surnageant d'IP. Les échantillons sont incubés 5 minutes à 95°C puis chargés sur un gel de polyacrylamide 7,5%.

Le choix des billes de sépharose, protéine G ou protéine A se fait en fonction de l'anticorps que l'on va utiliser par la suite pour l'immunoprécipitation.

V. Immobilisation d'anticorps sur billes de sépharose (Abcam)

Pour améliorer l'efficacité de l'immunoprécipitation avec l'anticorps α Venus, j'ai immobilisé l'anticorps sur des billes de sépharose

Les billes de sépharose couplées aux protéines A ou G sont à choisir suivant le tableau de compatibilité.

Les billes sont lavées deux fois avec du PBS 1X et remises en suspension dans du PBS 1X à une concentration de 50% puis laisser en rotation à 4°C pendant la nuit. Le lendemain, les billes sont centrifugées à 14000 rpm pendant 1 minute et le surnageant est aspiré. Le tampon de dilution (PBS 1X, 1mg/ml BSA) est ajouté à un ratio 1:1 et les billes mises en rotation pendant 10 minutes à 4°C. Après une centrifugation, le surnageant est aspiré et la solution d'anticorps (20µg d'anticorps pour 30ml de billes dans du tampon de dilution) est ajouté à un ratio de 1 :1. Le mélange est fait par pipetage puis rotation à 4°C pendant une heure. Après centrifugation, le surnageant est ôté et remplacé par du tampon de dilution à un ratio 1 :1. Après une rotation de 5 minutes à 4°C et une centrifugation le surnageant est remplacé par du PBS 1X à un ratio 1 :1

Le diméthyle pimélimidate (DMP) est préparé extemporanément à une concentration de 13mg/ml dans du tampon de lavage (0,2M de triéthanolamine dans du PBS). Le DMP est alors ajouté aux billes à un ratio 1 :1 et le mélange est mis en rotation pendant 30 minutes à températures ambiante. Les billes sont alors lavées avec 1ml de tampon de lavage et après une rotation de 5 minutes à température ambiante, centrifugées et le surnageant enlevé. Cette opération DMP + lavage est renouvelée deux fois. Après la dernière centrifugation, le surnageant est enlevé et 50mM d'éthanolamine sont ajoutés à un ratio volumique de 1 :1 pour « quencher » le DMP, en rotation pendant 5 minutes à température ambiante puis les billes sont lavées avec du PBS. Cette opération est répétée une fois. Pour retirer l'excès d'anticorps non fixés, les billes sont lavées avec 1ml de glycine 1M, pH3, en rotation pendant 10 minutes à température ambiante. Cette étape est répétée une fois. Les billes sont alors lavées trois fois dans du PBS Tween-20 0,01% puis elles peuvent être stockées à 4°C pour quelques jours. Pour éviter une contamination bactérienne, de l'azide de sodium peut être ajouté.

Tableau M.V-1 Compatibilité anticorps/protein A or G beads (adapté de Dynabeads)

+++ Liaison forte, ++ liaison moyenne, + liaison faible, - pas de liaison			
Espèce	Sous classe d'immunoglobuline	Protéine A	Protéine G
Humain	IgG1, IgG2, IgG4	+++	+++
	IgG3	+	+++
	IgD	-	-
	Fab	+	+
	ScF _v	+	-
Souris	IgG1	+	+++
	IgG2a, IgG2b, IgG3	+++	+++
	IgM	-	-
Rat	IgG1	+	++
	IgG2a	-	+++
	IgG2b	-	+
	IgG2c	+++	+++
Chèvre	IgG1	+	+++
	IgG2	+++	+++
Mouton	IgG1	+	+++
	IgG2	+++	+++
Bovin	IgG1	+	+++
	IgG2	+++	+++
Cheval	IgG(ab)	+	-
	IgG(c)	+	-
	IgG(T)	-	+++
Lapin	Tous les Ig	+++	++
Chien	Tous les Ig	+++	+
Chat	Tous les Ig	+++	+
Cochon	Tous les Ig	+++	+
Cochon d'inde	Tous les Ig	+++	+
Poulet	Tous les Ig	-	-

VI. L'Immunoprécipitation de Chromatine

A. La préparation des cellules

Cinquante millions de cellules sont lavées deux fois dans du PBS 1X à température ambiante avec une centrifugation de 10 minutes à 1500 rpm. Puis les cellules sont comptées à l'aide d'une cellule de comptage Neubauer et du bleu trypan. Les cellules sont resuspendues dans du PBS 1X à une concentration de $5 \cdot 10^6$ cellules/ml dans un

tube de 50mL. Pour la première étape de crosslink, une solution stock de Disuccinimidyle glutarate (DSG) est préparée extemporanément à une concentration de 163mg/ml dans du Diméthylsulfoxyde (DMSO). Le DSG est alors ajouté à la suspension cellulaire pour obtenir une concentration finale de 652µg/ml. Le tube de 50mL est alors placé sur un agitateur pendant 45 minutes à température ambiante. Cette étape permet de fixer les protéines d'un même complexe en créant un lien de 7,7Å entre deux terminaisons NH₂. Puis les cellules sont culotées par centrifugation 10 minutes à 1500 rpm à température ambiante et lavées à deux reprises dans 20 ml de PBS 1X. Le culot cellulaire est alors resuspendu dans une solution de paraformaldéhyde (PFA) à 1% dans du PBS 1X. Cette étape de 10 minutes à température ambiante permet de fixer les complexes protéiques à l'ADN. Puis de la glycine est ajouté à une concentration finale de 0,25M pendant 5 minutes à température ambiante pour emprisonner la PFA libre. Les cellules sont alors culotées par centrifugation 10 minutes à 1500 rpm à 4°C. Puis les cellules sont resuspendues dans 20ml de PBS1X froid et placées sur un agitateur à 4°C pendant 15 minutes puis culotées par centrifugation comme précédemment. Cette étape de lavage est répétée une fois.

Puis les cellules sont resuspendue dans 1ml du premier tampon de lyse LB1 (Tris-HCl pH8 50mM, EDTA pH8 2mM, NP-40 0,1%, Glycérol 10%) complémenté d'inhibiteurs de protéase (cOmplete™, EDTA-free Protease Inhibitor Cocktail – Merck milipore) dans un tube eppendorf de 1,5ml. Les cellules sont incubées dans la glace pendant 10 minutes puis centrifugées à 1500rpm pendant 10 minutes à 4°C. Le culot est resuspendu dans 1mL du tampon de lyse LB2 (Tris-HCl pH8 10mM, NaCl 200mM, EDTA pH8 1mM, EGTA pH8 0,5mM) supplémenté avec des inhibiteurs de protéase, incubé 10 minutes dans la glace et les noyaux sont culotés par centrifugation 10 minutes à 1500 rpm à 4°C. Les noyaux sont alors lavés deux fois dans 800µl de tampon LB2.3 (2/3 LB2, 1/3 LB3 (Tris-HCl pH8 50mM, NaCl 200mM, EDTApH8 1mM, EGTA pH8 0,5mM, Deoxycholate de sodium 0,1%, N-laurylsarcosine 0,5%) supplémenté avec des inhibiteurs de protéases puis resuspendus dans ce même tampon à une concentration de 100.10⁶ cellules /ml. À cette étape, les échantillons peuvent être conservés à -80°C.

B. La sonication

La sonication des échantillons se fait dans des tubes 1,5ml TPX (#C30010010 diagenode) à $25 \cdot 10^6$ cellules par tube, soit 250 μ l d'échantillon dans un sonicateur Bioruptor® Plus (Diagenode). Un cycle de sonication est : 5 minutes 15 secondes ON / 15 secondes OFF puissance HIGH. Pour les cellules BH1, 6 cycles de sonication sont nécessaires puis les débris sont culotés par centrifugation à 14000 rpm pendant 10 minutes à 4°C. Pour obtenir des fragments de chromatine entre 300pb et 150pb. Pour contrôler la sonication, 50 μ l d'échantillons sont prélevés et additionné de NaCl à une concentration finale de 200mM puis placés à 65°C pendant au moins 4 heures. L'ADN est alors purifié sur colonne avec le kit PCR&Gel Clean up (Machery Nagel) puis migré sur gel d'agarose 2%.

C. L'immunoprécipitation

Les échantillons sont divisés dans trois eppendorf 1,5mL à raison de 150 μ l par tube puis dilué au millième dans du tampon de dilution (Tris-HCl pH8 50mM, NaCl 200mM, EDTA pH8 1mM, NP-40 0,5%). Pour l'étape de « pre-clear », les billes de sépharose couplées aux protéines A ou G sont lavées deux fois dans du tampon de dilution et resuspendues à une concentration de 50%. Dans chaque tube contenant l'échantillon, 80 μ l de billes sont ajoutées. Les tubes sont alors placés en rotation pendant une heure à 4°C. Après une centrifugation de 5 minutes à 1500rpm à 4°C, le surnageant est récupéré et placé dans un nouveau tube eppendorf 1,5mL et 20 μ l par tubes sont prélevés (60 μ l au total) pour l'Input et conservés à -20°C. L'anticorps est ajouté dans les échantillons, par exemple pour Ikaros, 5 μ g d'anticorps sont ajoutés dans chaque tube. Puis les tubes sont placés en rotation à 4°C sur la nuit. Le lendemain, 40 μ l de billes à 33% sont ajoutés dans les tubes et du Sodium Dodecyl Sulfate (SDS) à une concentration finale de 0,1% et les tubes sont remis en rotation à 4°C pendant une heure. Les tubes sont alors centrifugés pendant 5 minutes à 1500 rpm à 4°C. Le surnageant est récolté et conservé dans un tube de 15mL à -20°C. Les billes sont lavées avec 800 μ l de tampon de lavage (Tris-HCl pH8, NaCl 200mM, EDTA pH8 2mM, NP-40 1%, SDS, 0,1%), une fois rapidement puis cinq fois en alternant 5 minutes de rotation à 4°C et la centrifugation. Puis les billes sont lavées trois fois en alternant 5

minutes en rotation à 4°C et la centrifugation avec du TE 50-10 (Tris-HCl 50mM – EDTA 10mM). L'ADN est alors élué en resuspendant les billes dans 35µl de TE 50-10 supplémenté avec 2% de SDS. Après une incubation de 5 minutes à température ambiante, les billes sont centrifugées pendant 5 minutes à 1500rpm. Le surnageant est récupéré à l'aide d'une seringue d'1ml et une aiguille 29G. Cette étape est répétée avec 20µl de TE 50-10 2%SDS. Au total, 55µl sont récupérés de chaque tube faisant un total de 165µl par échantillons.

D. La purification de l'ADN

L'Input est récupéré et décongelé. Du NaCl est ajouté à une concentration finale de 200mM. Les tubes Input et échantillons sont alors placés dans un thermoblock à 65°C pendant une nuit.

L'ADN de l'Input est purifié sur colonne à l'aide du kit PCR&Gel Clean up (Machery Nagel) en suivant les recommandations du fabricant et élué dans 120µl de tampon d'éluion. L'ADN de l'échantillon est purifié à l'aide du kit de Qiagen MinElute Reaction Clean up en suivant les recommandations du fabricant et élué dans 40µl de tampon d'éluion.

L'ADN peut aussi être purifié avec des tubes Heavy Phase-Lock gel 1,5mL (Eppendorf). Dans ce cas, les tubes contenant le gel sont centrifugés 20 secondes à 14000g. Un volume de phénol-chloroforme 1 :24 alcool isoamyle est ajouté à un volume d'échantillon et le mélange se fait par pipetage. Le mélange est alors déposé sur le gel et le tube est centrifugé 5 minutes à 14000g. Un volume de chloroforme est ajouté à la phase aqueuse et le mélange est fait par pipetage et le tube est centrifugé pendant 5 minutes à 14000g. La phase aqueuse est alors récupérée et placée dans un nouveau tube Eppendorf 1,5mL. Du glycogène est alors ajouté à une concentration finale de 66µg/mL et le mélange est vortexé. Le glycogène est utilisé pour améliorer l'efficacité de la précipitation de l'ADN. La précipitation de l'ADN se fait en ajoutant 2,5 volumes d'éthanol 100% et après avoir vortexé le tube pendant 10 secondes, celui-ci est placé une heure à -80°C. L'ADN est culoté par centrifugation pendant 15 minutes à 14000g à 4°C. Le surnageant est jeté et l'ADN est lavé avec 1mL d'éthanol à 100% et centrifugé 5 minutes à 14000g à 4°C. Le surnageant est jeté et le tube est laissé

ouvert pendant 10 minutes jusqu'à évaporation totale de l'éthanol. Le culot d'ADN est resuspendu dans de l'eau sans nucléase avec les volumes précédemment cités.

E. Le contrôle par qPCR

Pour vérifier que l'expérience d'immunoprécipitation de chromatine a fonctionné, une PCR quantitative est réalisée pour deux sites présents dans la séquence génomique du gène Cish. L'un est connu pour être un site de liaison des protéines de la famille d'Ikaros (Cish184) et l'autre est un contrôle négatif de fixation (Cish4200).

Tableau M.VI-1 Amorces pour la PCRq Cish

Nom des amorces	Séquence
Cish 184 F	GTCCAAAGCACTAGACGCCTG
Cish 184 R	TTCCCGGAAGCCTCATCTT
Cish 4200 F	TACCCCTTCCAACCTCTGACTGAGC
Cish 4200 R	TTCCCTCCAGGATGTGACTGTG

Pour la préparation de la gamme, l'input est utilisé à différentes concentration : 1X ; 0,5X ; 0,2X ; 0,04X ; 0,008X ; 0,0016X ; 0,00032X. Le mastermix est composé pour 1 puit de 6µl SYBRGreen, 1,6µl H₂O et 2,4µl Primer Mix 5µM. La qPCR est alors réalisée sur un Lightcycler ®480 (software release 1.5.0 SP3) en suivant le programme suivant :

	Temperature (°C)	Temps
Dénaturation	95	15 min
Cycles d'amplification	95	10 sec
	57	30 sec
Mesure pour quantification	72	15 sec
Fusion	95	5 sec
	60	5 sec
	95	Continue
Refroidissement	37	-2.2°C/s

Un graphique représentant le $\log_{(10)}$ de la concentration en fonction du nombre de cycle pour les points de la gamme. La droite de tendance linéaire donne une équation avec laquelle on peut calculer la concentration relative des échantillons. Ainsi, on peut calculer la concentration relative de l'ADN immunoprécipité en fonction de l'Input.

F. La préparation de la librairie et séquençage

Après une purification de l'ADN grâce aux billes SPRIselect avec un ratio 1.8, la concentration de l'ADN est évaluée avec un Qubit (Invitrogen) et le kit HS dsDNA (High Sensitivity double strand DNA) La préparation de la librairie se fait avec le kit Multiplex library (Diagenode) en suivant les instructions du fabricant. Les librairies sont alors purifiées pour enlever les primers et les séquences de plus de 600bp. La qualité de la librairie est évaluée grâce au 2100 Bioanalyzer (Agilent Technologies) avec une puce High-Sensitivity DNA (#5067-4627). Le séquençage est réalisé sur un séquenceur HiSeq4000 Illumina 1x50b.

G. Analyse bioinformatique

Tous les outils d'analyse sont disponibles sur : <http://www.usegalaxy.org>.

L'alignement des séquences sur le génome *Mus musculus* 10 (mm10) est fait avec Bowtie 2 (version 2.3.4.3). Puis les résultats sont filtrés sur la qualité de l'alignement (MAPQ=40) en utilisant Samtoolsview (version 1.9). L'identification des pics est faite avec MACS2callpeak (2.1.1.20160309.6) en regroupant les fichiers contrôles (Input) d'une même analyse. L'annotation des pics est faite avec Homer (version 0.0.5, disponible sur <http://www.usegalaxy.org>).

VII. L'Omni-ATAC (Assay for Transposase-Accessible Chromatin)

A. Le protocole

Protocole d'Amanda Ackerman (28/01/2019), Laboratoire Kaestner

Ce protocole est basé sur les publications de Buenrostro *et al.* (Buenrostro *et al.*, 2015, 2013)

Après avoir lavé 50 000 cellules avec du PBS 1X (phosphate Buffer Saline) dans un eppendorf 1,5mL, le culot cellulaire a été remis en suspension dans 50µl de tampon de lyse (48,5µl de tampon de resuspension (49,25mL d'H₂O sans nucléase, 500µl Tris-

HCl pH7,5 1M, 150µl MgCl₂ 1M, 100µl NaCl 5M), 0,5µl de NP-40 10%, 0,5µl the Tween-20 (10%) et 0,5µl Digitonine 1%). Après 3 minutes d'incubation dans la glace, 1mL de tampon de lavage (990µl de tampon de resuspension, 10µl de Tween-20 10%) a été ajouté et le tube a été inversé trois fois pour l'homogénéisation. L'échantillon a été ensuite centrifugé 10 minutes à 4°C à 500g. Le culot est alors composé des noyaux et le surnageant est composé du cytoplasme.

Les noyaux ont été remis en suspension dans le tampon de transposition (25µl de 2X Tagment DNA Buffer, 16,5µl de PBS 1X, 5µl d'H₂O nuclease-free, 2,5µl de transposase Tn5 (Illumina), 0,5µl de Tween-20 10%, 0,5µl de digitonine 1%) et placé dans un thermoblock à 37°C pendant 30 minutes avec une agitation à 1000rpm.

L'ADN transposé a ensuite été isolé et purifié en utilisant le kit de Qiagen MinElute Reaction CleanUp en suivant les instructions du fabricant. L'élution de l'ADN s'est fait dans 10µl de tampon d'élution. À ce stade, l'ADN peut être conservé à -20°C.

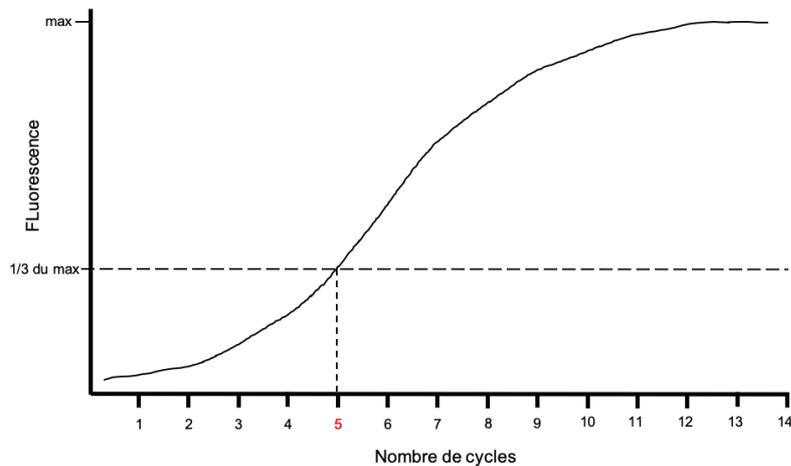
Pour préparer la librairie il faut réaliser une PCR dont les amorces sont répertoriées dans le tableau. Dans un tube de PCR, les 10µl d'ADN échantillon ont été ajouté à un mix contenant 10µl d'H₂O sans nucléase, 25µl de NEBnext High-Fidelity 2X PCR Master Mix (Illumina), 2,5µl de l'amorce Ad1_noMX (20µM) et 2,5µl d'une amorce Ad2.* possédant un index (20µM). Le programme de PCR est le suivant :

- 72°C - 5 minutes
 - 98°C - 30 secondes
 - 98°C - 10 secondes
 - 63°C - 30 secondes
 - 72°C - 1 minute
- } 5

Après ces 5 premiers cycles de PCR, une qPCR peut être réalisée pour évaluer le nombre de cycles de PCR que l'on doit ajouter. Pour cela, 5µl de librairie partiellement amplifiée sont prélevés et ajoutés à un mix de qPCR contenant 5µl de NEBNext High Fidelity 2X PCR Master Mix (Illumina), 3,85µl d'H₂O sans nucléase, 0,5µl d'amorce Ad1, 0,5µl d'amorce Ad2.*, 0,15µl de SYBR Green I 100X. Le programme de qPCR est le suivant :

- 98°C – 30 secondes
 - 98°C – 10 secondes
 - 63°C – 30 secondes
 - 72°C – 1 minute
- } 20 cycles

Il faut ensuite réaliser le graphique de la fluorescence en fonction du nombre de cycle. Le nombre de cycles de PCR à ajouter correspond au nombre de cycle qui permettent d'atteindre 1/3 du maximum de fluorescence, comme décrit dans le graphique.



Pour rajouter les cycles nécessaires, il suffit de replacer le tube de PCR dans le thermocycleur et de rajouter les cycles en suivant ce programme :

- 98°C – 10 secondes
 - 98°C – 10 secondes
 - 63°C – 30 secondes
 - 72°C – 1 minute
- } N

Pour évaluer la qualité de la librairie, 1µl de la PCR est analysé sur le 2100 Bioanalyzer (Agilent Technologies) avec une puce High-Sensitivity DNA (#5067-4627). Deux cas de figures se présentent alors, soit la librairie nécessite de nouveaux cycles d'amplification, soit la librairie est satisfaisante, c'est-à-dire que l'ADN libre, les fragments comportant un et deux nucléosomes sont bien détectés. La purification de la librairie se en utilisant des billes SPRIselect (Beckman Coulter) avec un ratio (billes/ADN) de 1,8. Afin d'être certain d'enlever l'intégralité des amorces, la purification est faite deux fois.

Le séquençage de 2x100 bases est fait sur un séquenceur Hiseq 4000.

Tableau M.VII-1 Amorces d'ATAC-seq

Nom du primer	Séquence du primer
Ad1_noMX	AATGATACGGCGACCACCGAGATCTACACTCGTCGGCAGCGTCAGATGTG
Ad2.2_CGTA TAG	CAAGCAGAAGACGGCATA CGAGATCTAGTACGGTCTCGTGGGCTCGGAGATGT
Ad2.3_AGGCAGAA	CAAGCAGAAGACGGCATA CGAGATGCTCAGGAGTCTCGTGGGCTCGGAGATGT
Ad2.5_GGACTCCT	CAAGCAGAAGACGGCATA CGAGATCATGCCTAGTCTCGTGGGCTCGGAGATGT
Ad2.6_TAGGCATG	CAAGCAGAAGACGGCATA CGAGATGTAGAGAGGTCTCGTGGGCTCGGAGATGT
Ad2.7_CTCTCTAC	CAAGCAGAAGACGGCATA CGAGATAGCGTAGCGTCTCGTGGGCTCGGAGATGT
Ad2.8_CAGAGAGG	CAAGCAGAAGACGGCATA CGAGATCAGCCTCGGTCTCGTGGGCTCGGAGATGT
Ad2.9_GCTACGCT	CAAGCAGAAGACGGCATA CGAGATTGCCTCTTGTCTCGTGGGCTCGGAGATGT

B. L'analyse bioinformatique

Les données ont été analysées en utilisant le pipeline ENCODE ATAC-seq (V.1.5.1). L'alignement des reads est réalisé avec Bowtie2 (v.2.3.4.3) (Langmead and Salzberg, 2012). L'identification des pics a été faite en utilisant MACS2 (v.2.2.4). Les comparaisons d'intérêts (Ikaros vs CTRL et Aiolos vs CTRL) ont été faites en utilisant la méthode proposée en 2014 par Love, Huber et Anders (Love et al., 2014) : DeSeq2 (v.1.24.0).

VIII. Analyse de la transcription par RNA-seq

A. Le protocole

Les cellules BH1-rtTA, BH1-Ikaros-V1/Ikaros-V2 et BH1-Aiolos-V1/Aiolos-V2 ont été traitées pendant 24 heures avec des doses adaptées de doxycycline : 2µg/ml, 50ng/ml et 35 ng/ml respectivement. Les ARNs ont été extraits suivant les recommandations du kit RNeasy Mini (# 74104) de Qiagen.

Les bibliothèques ont été réalisées en suivant les recommandations du kit TruSeq Stranded mRNA (Illumina #20020595).

Le séquençage de 1x50 bases a été réalisé sur le séquenceur HiSeq4000 (Illumina).

B. Analyses bioinformatiques

Les reads ont été préprocessées pour retirer les adaptateurs de la librairie, les séquences polyA et les séquences de faibles qualités (score <20). Puis, les reads de moins de 50 bases ont été retirées des analyses. Les reads conservées ont été alignées sur le génome mm10 (*Mus musculus*) en utilisant le logiciel STAR (v.2.5.3a). La quantification de l'expression génique a été faite en utilisant les reads alignées une seule fois en utilisant htseq-count (v.0.6.1.p1) avec les annotations d'Ensembl (v.90). Seules les reads allouées de façon non-ambiguë ont été gardées pour la normalisation. Les reads ont été normalisées avec la méthode proposée par Anders et Huber (Anders and Huber, 2010). Les comparaisons d'intérêt ont été faites en utilisant DeSeq2 (v.1.16.1).

IX. Isolation et marquage des cellules primaires

Les cellules primaires sont issues de souris sauvage (WT) ou mutantes ($\Delta Ikzf3$). Les souris sont sacrifiées par exposition au CO₂.

A. Isolation des cellules de la moelle osseuse

Pour l'isolation des cellules de la moelle osseuse, les fémurs, tibias et péronés sont récupérés et broyés dans un mortier dans une solution de PBE (PBS 1X, BSA 0,5%, EDTA 2mM). Pour un enrichissement total, la moelle épinière et le sternum sont également récupérés et broyés. Le broyat est filtré sur des membranes de 50µm. Après une centrifugation des 10 minutes à 1500rpm à 4°C, le culot est resuspendu dans une solution de NH₄Cl 0,83% pour l'érythrolyse et incubé à température ambiante pendant 5 minutes. Après deux lavages en PBE, les cellules sont prêtes pour l'immunomarquage.

B. Isolation des cellules de la rate

La rate est récupérée dans un tampon PBE puis écrasée à l'aide d'un piston de seringue dans une boîte de pétri de 60x15mm. La suspension cellulaire est passée sur un tamis de 50µm puis centrifugée pendant 10 minutes à 1500rpm, 4°C. Le culot est resuspendu dans une solution de NH₄Cl 0,83% pour l'érythrolyse et incubé à température ambiante pendant 5 minutes. Après deux lavages au PBE, les cellules sont prêtes pour un immunomarquage.

C. Enrichissement sur colonne

Pour récupérer les lymphocytes B spléniques, les cellules de la rate sont comptées au bleu de trypan puis centrifugées 10 minutes à 1500rpm à 4°C. Le culot cellulaire est resuspendu à une concentration de 20.10⁶ cellules/ml dans du PBE + CD19-biotine (1/100^{ème}). La suspension est laissée sur glace pendant 30 minutes puis les cellules sont lavées deux fois avec du PBE et finalement resuspendues à une concentration de 10⁷/90µl. Un aliquot de 20µl est prélevé pour la vérification de pureté. Les billes magnétiques streptavidine (MACS 130.048.101) sont ajoutées à une concentration de 10µl/10⁷cellules. Le mélange est laissé sur glace pendant 15 minutes puis lavé avec 2ml de PBE et centrifugé. Les cellules sont resuspendues dans du PBE à une concentration de 10⁸cellules/500µl. La colonne MACS LS est placée sur le séparateur MACS et rincée avec 3ml de PBE. La suspension cellulaire est chargée sur la colonne et les cellules non marquées sont récupérées, un aliquot est prélevé pour la vérification de pureté. La colonne est lavée trois fois avec du PBE puis ôtée du séparateur et placée sur un tube de récolte. Pour récupérer les cellules retenues sur la colonne il faut alors appliquer 5ml de PBE sur la colonne et utiliser le piston pour forcer le passage du tampon. Les cellules sont resuspendues et un aliquot est prélevé pour la vérification de pureté.

1. Vérification de pureté

Les trois aliquots « avant », « pendant », « après » sont récupérés et centrifugés puis le culot est resuspendu dans 100µl de PBE + mix d'anticorps (CD11b-FITC 1/100^{ème}, CD3-Cy5 1/100^{ème}, Streptavidine-PE 1/400^{ème}). Le mélange est gardé

sur glace pendant 30 minutes puis lavé avec du PBE. Les cellules marquées sont alors analysées par cytométrie de flux.

X. Marquage intracellulaire des facteurs de transcription de la famille d'Ikaros dans les cellules primaires et les cellules BH1

A. Dans les cellules primaires de la moelle osseuse

Les cellules primaires sont mises à une concentration de $20 \cdot 10^6/\text{ml}$ dans du PBE + mix d'anticorps (Tableau M.X.1) à la concentration optimale. Le marquage se fait pendant 30 minutes sur glace puis les cellules sont lavées dans du PBE. Ensuite les cellules sont fixées puis perméabilisées et marquées en suivant les recommandations du kit Foxp3 / Transcription Factor Staining Buffer Set (eBiosciences).

Les différentes fractions des cellules de la moelle osseuse sont alors analysées grâce suivant la stratégie présentée figure M.X.1.

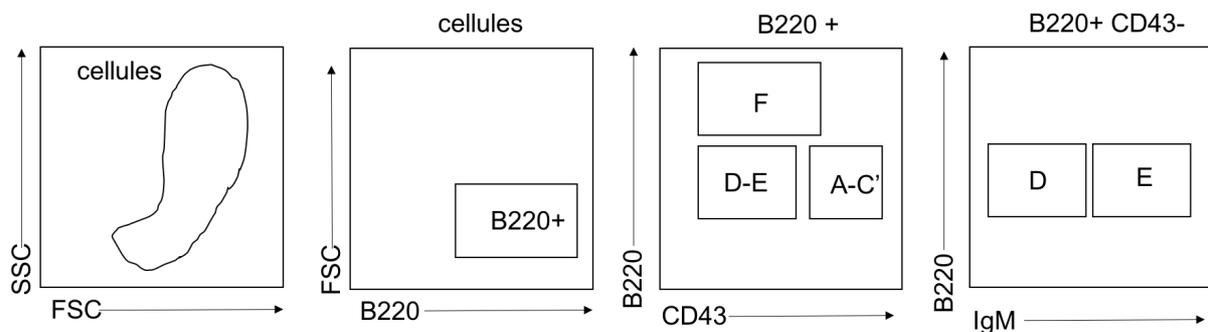


Figure M.X.1 Analyse des différentes fractions de la moelle osseuse

B. Dans les cellules BH1

Les cellules BH1 sont récupérées en culture puis lavées au PBS. Le marquage avec les anticorps primaires α Ikaros, α Aiolos ou α Helios (Tableau M.X.1) puis l'anticorps secondaire α Lapin-PE est fait en suivant les recommandations du kit Transcription Factor Staining Buffer Set (eBiosciences).

Tableau M.X-1 Liste des anticorps utilisés

Spécificité	Conjugaison	Concentration	Fournisseur	Référence
Ikaros C-ter	Purifié	1/800	IGBMC	
Aiolos	Purifié	1/100	Cell signaling	D1C1E
Aiolos	PE	1/100	eBioscience	12-5789-80
Helios	Purifié	1/100	Cell signaling	D8W4X
IgM	Cy5	1/300	Southern biotech	SBA 1020-15
Kappa LC	BV421	1/100	BD Biosciences	562888
CD179b(λ 5)	Biotine	1/100	BD Biosciences	551865
B220	FITC	1/100	Invitrogen	RM2601-3
CD19	Biotin	1/100	eBioscience	13-0193
CD43	BV421	1/100	BD biosciences	562958
Anti-lapin	PE	1/400	Jackson	711-116-152
Streptavidine	APC-Cy7	1/400	BD Biosciences	554063
Streptavidine	AF700	1/400	Invitrogen	S-21383
Streptavidine	PE	1/400	Jackson	016-110-084

Résultats

Dans le laboratoire nous avons établi une lignée de lymphocytes pré-B murins (BH1) à partir des cellules de la moelle osseuse d'une souris *Ikaros*^{Flox/Flox} (Heizmann et al., 2013). Ces cellules sont déficientes pour Ikaros et sont bloquées dans leur différenciation au stade pré-B. La réexpression d'Ikaros dans ces cellules permet de reprendre le processus de différenciation et d'obtenir, *in vitro*, des cellules exprimant le BCR. Ces cellules expriment de faibles niveaux d'Aiolos et Helios qui ne sont donc pas capable de soutenir la différenciation des lymphocytes pré-B. Ainsi, Ikaros semble, ici, avoir une fonction unique. Cependant, l'absence d'Ikaros ne conduit pas seulement à l'absence des homodimères d'Ikaros mais également à l'absence des différents hétérodimères qu'il pourrait former avec Helios et Aiolos. Il est donc impossible de savoir quel dimère est nécessaire à la différenciation des lymphocytes pré-B.

Pour étudier spécifiquement les dimères, j'ai utilisé un système de complémentarité bi-moléculaire de fluorescence (BiFC), plus précisément décrit dans le matériel et méthode, qui repose sur l'association de chaque monomère avec une moitié (V1 ou V2) non fluorescente de la protéine fluorescente Venus. Lorsque les deux monomères forment le dimère, les deux moitiés de Venus s'associent et reconstituent la fluorescence de Venus. Cela permet de directement visualiser la dimérisation par cytométrie de flux ou immunofluorescence. Grâce à notre collaboration avec l'équipe de Simona Sacconi (IRCAN, Nice), j'ai pu également utiliser un anticorps (α Venus) reconnaissant un épitope présent uniquement sur la protéine Venus et sur la protéine Venus reconstituée par V1 et V2. Ainsi, la réalisation d'expérience d'immunoprécipitation de chromatine avec cet anticorps m'a permis d'étudier la liaison à l'ADN des différents dimères de la famille de facteurs de transcription Ikaros. Cela donne l'avantage d'utiliser le même anticorps pour les différents dimères et ainsi de ne pas induire de biais dû à une efficacité différente des anticorps spécifiques des facteurs de transcription.

Dans ce travail de thèse, j'ai étudié les différents dimères formés par les facteurs de transcription de la famille d'Ikaros suivant deux grands axes de travail : leurs spécificités de liaison à l'ADN et leurs spécificités fonctionnelles au cours du développement précoce des lymphocytes B. Ces deux grands axes seront présentés, dans cet exposé, à la suite de la preuve de concept de la méthode.

I. La preuve de concept

Dans notre laboratoire nous avons acquis une expertise sur le rôle d'Ikaros dans le développement précoce des lymphocytes B. C'est pourquoi j'ai décidé de valider le concept de mon projet avec Ikaros. J'ai détaillé dans cette section les étapes nécessaires à l'obtention d'une lignée BH1 exprimant l'homodimère d'Ikaros : Ikaros-V1/Ikaros-V2. J'ai ensuite vérifié que la présence des fragments V1 et V2 ne perturbent ni la localisation d'Ikaros ni sa fonctionnalité spécifique au cours de la différenciation des lymphocytes pré-B.

A. Quelle position pour les fragments V1 et V2 : N-terminale ou C-terminale ?

Les facteurs de la famille d'Ikaros possèdent leur domaine de liaison à l'ADN en position N-terminale (Nter) et leur domaine de dimérisation en position C-terminale (Cter). Pour vérifier à quelle position, Nter ou Cter, les fragments V1 et V2 doivent être positionnés, j'ai testé les combinaisons suivantes : Ikaros-V1/Ikaros-V2 ; Ikaros-V1/V2-Ikaros ; V1-Ikaros/Ikaros-V2 et enfin V1-Ikaros/V2-Ikaros où la position de V1 et de V2 par rapport à Ikaros représente la position en Nter ou Cter.

Pour cette expérience, l'expression des transgènes est accompagnée par celle d'un gène rapporteur dont la séquence se situe après une séquence IRES (Internal Ribosome Entry Site) dans la construction plasmidique. Ainsi la protéine Ikaros liée en Nter ou Cter au fragment V1 sera accompagnée de l'expression de hCD4 (CD4 humain) et Ikaros liée au fragment V2 par TdTomato (Figure RI.1)

Toutes les combinaisons testées sont capables d'exprimer la fluorescence de Venus. Les combinaisons N-ter/N-ter et C-ter/C-ter sont les deux qui présentent une intensité de fluorescence plus importante. Cela pourrait refléter une association plus stable de V1 avec V2. Cependant, seule la combinaison C-ter/C-ter présente une population cellulaire qui semble avoir diminué en taille, ce qui est la première marque de la reprise de la différenciation des cellules BH1 ré-exprimant Ikaros. Au regard de cette expérience, j'ai donc choisi de positionner les fragments V1 et V2 en partie Cter des facteurs de transcription de la famille d'Ikaros.

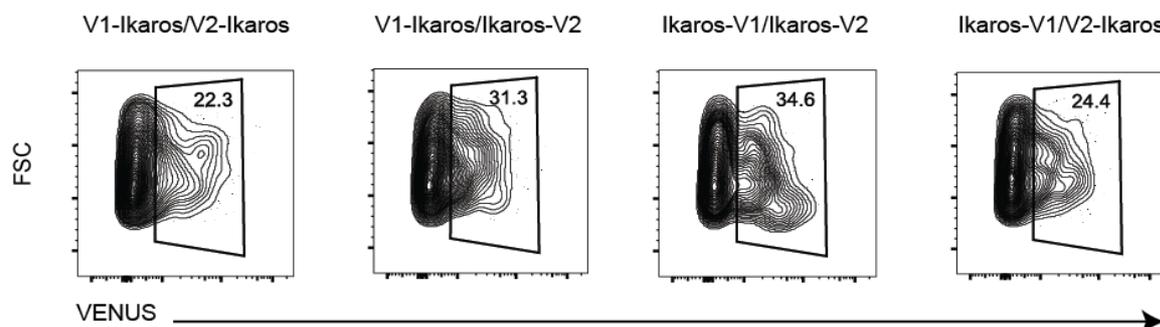


Figure RI.1 Positionnement des fragments V1 et V2 en N- ou C-terminal

Les cellules BH1 ont été transduites avec les vecteurs rétroviraux pMihCD4 et pMiTo portant respectivement les transgènes liés à V1 et à V2. La position de V1 et de V2 dans le nom de la lignée est représentative de la position en N- ou C-terminale. L'analyse par cytométrie de flux est faite trois jours après la transduction. Le marquage du rapporteur hCD4 n'étant pas satisfaisant, l'expression de la fluorescence de Venus par la complémentation V1+V2 est analysée sur les cellules TdTomato+.

B. Création de la lignée BH1-Ikaros-V1/Ikaros-V2

Comme expliqué précédemment, la réexpression d'Ikaros dans les BH1 permet de reprendre le processus de différenciation. Ainsi, pour les études fonctionnelles, il est nécessaire de pouvoir contrôler l'expression des transgènes dans ces cellules. Pour cela j'ai choisi un système d'activation de la transcription par la tétracycline (tétracycline-ON ou tet-ON) ou son analogue, la doxycycline. J'ai dû tester différentes stratégies avant d'obtenir la lignée BH1-Ikaros-V1/Ikaros-V2

1. La lignée BH1-rtTA

Le système tet-ON nécessite la présence, dans la cellule, de son transactivateur (rtTA). Cette molécule agit comme un récepteur nucléaire pour la doxycycline. En présence de doxycycline, il change de conformation et peut alors se fixer sur ses récepteurs spécifiques présents sur le promoteur TRE3G (Tetracycline Responsive Element – 3rd generation) qui contrôle la transcription des transgènes.

J'ai, dans un premier temps, transduit les cellules BH1 grâce à un rétrovirus permettant l'insertion de l'ADNc codant pour rtTA ainsi que celui du gène de résistance à la néomycine. Après sélection, j'ai obtenu une lignée proliférant normalement dans un milieu supplémenté en néomycine, que j'ai considérée comme stable, et que j'ai nommée BH1-rtTA.

2. Première stratégie pour l'obtention de la lignée BH1-Ikaros-V1/ikaros-V2

Pour obtenir une lignée exprimant de façon équivalente les deux nouvelles formes d'Ikaros, je les ai d'abord clonées dans le même vecteur, séparé par un IRES. Le vecteur p259 vient originalement avec le gène Hph codant pour l'Hygromycine B phosphotransférase qui dégrade l'hygromycine B et confère aux cellules la résistance au traitement. Après une semaine de sélection, j'ai activé la transcription des transgènes en traitant les cellules avec la doxycycline puis j'ai analysé la fluorescence de Venus par cytométrie de flux. Bien que les cellules prolifèrent normalement dans un milieu supplémenté en hygromycine, la fluorescence de Venus n'est pas détectable (Figure RI.3 A et B). Nous verrons dans le paragraphe suivant que l'hygromycine B n'est pas un antibiotique efficace dans notre lignée cellulaire. C'est pourquoi j'ai changé le gène de résistance dans mon vecteur p259 Ikaros-V1 IRES Ikaros-V2 par le gène Pac codant pour la Puromycine-N acétyltransférase conférant aux cellules une résistance au traitement à la puromycine. Cependant, j'ai obtenu le même résultat. Il semblerait que les cellules soient sélectionnées par l'antibiotique mais celles-ci n'expriment pas les gènes d'intérêt (Figure RI.3 C et D).

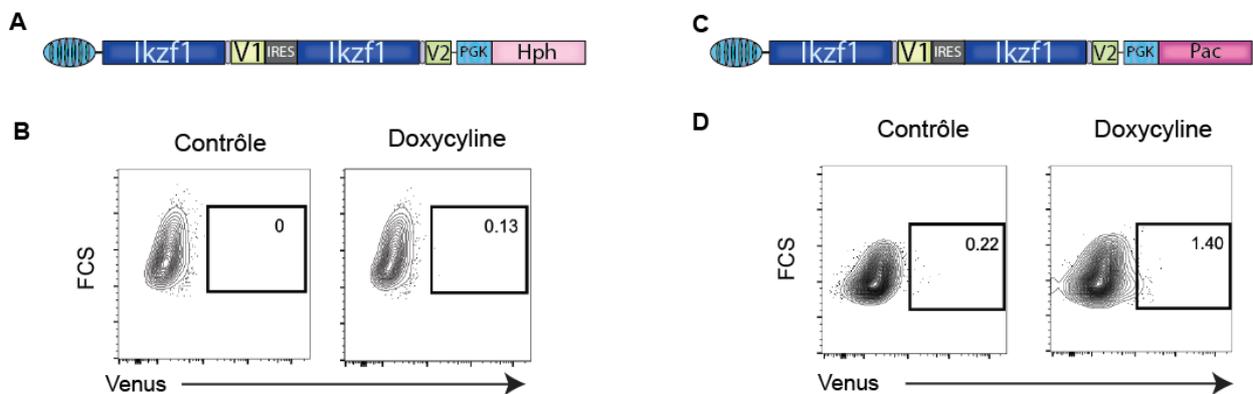


Figure RI.2 Première stratégie pour l'expression des transgènes liés aux fragments V1 et V2
Les cellules BH1 rtTA ont été transduites avec les vecteurs rétroviraux permettant l'insertion dans l'ADN des constructions représentées en A et C. En B et D, les dotplots obtenus par cytométrie de flux pour la fluorescence de Venus 24 heures après le traitement avec 2µg/ml de doxycycline.

3. Deuxième stratégie pour l'obtention de la lignée BH1-Ikaros-V1/ikaros-V2

L'échec de la première stratégie pourrait s'expliquer par la taille de la cassette qui doit être insérée : 8,1kb ce qui est la limite de capacité d'insertion des vecteurs rétroviraux.

J'ai donc décidé de faire exprimer les deux formes d'Ikaros grâce à deux vecteurs différents. Le premier portant d'ADNc codant pour Ikaros-V1 accompagné par le gène Hph : p259-Ikzf1V1-PGK-Hph (Figure RI.3.A). Après transduction des cellules, j'ai vérifié que les cellules expriment bien Ikaros en faisant un marquage intracellulaire et une analyse par cytométrie de flux (Figure RI.3.B). Seulement 34% des cellules expriment Ikaros après dix jours de sélection avec l'hygromycine. Le deuxième vecteur permet l'expression d'Ikaros-V2 et une sélection des cellules transduites avec la puromycine (Figure RI.3.C). J'ai également contrôlé l'expression d'Ikaros dans les cellules et 73% d'entre-elles expriment Ikaros après 10 jours de sélection. J'ai ensuite transduit les cellules BH1-rtTA avec les deux vecteurs à 24 heures d'intervalle puis j'ai commencé la sélection antibiotique le troisième jour. Après 10 jours de sélection pendant lesquels j'ai pu observer l'apparition de clusters cellulaires, les cellules ont commencé à proliférer normalement. J'ai alors testé l'activation de la transcription par la doxycycline et analysé l'expression de VENUS par cytométrie de flux (Figure RI.3.E).

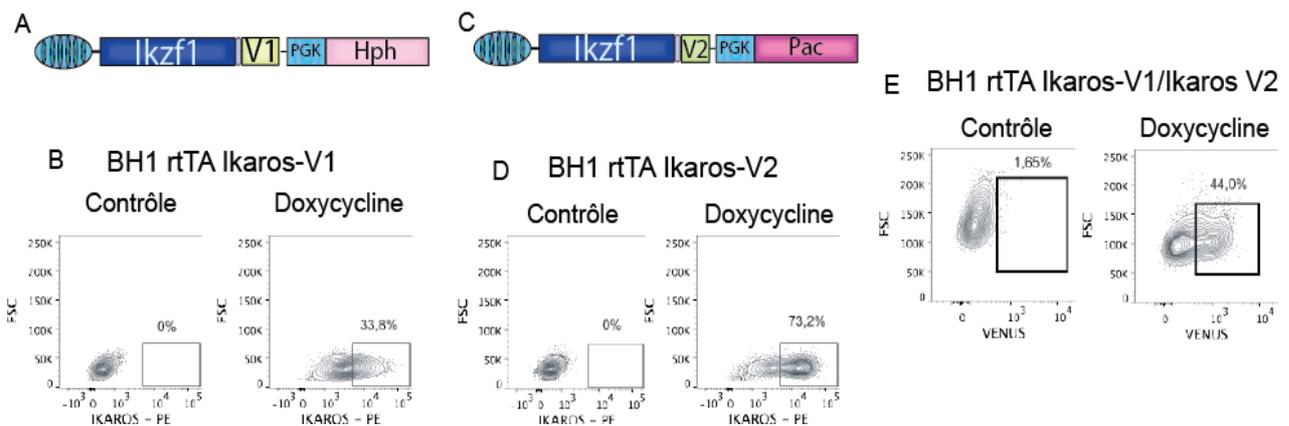


Figure RI.3 Expression de la fluorescence de Venus par complémentarité V1+V2

Les cellules BH1 rtTA ont été transduites avec les vecteurs rétroviraux permettant l'insertion des constructions présentées en A et C. En B et D sont représentés les marquages intracellulaires de la protéine Ikaros grâce à un anticorps polyclonal Lapin anti-Ikaros et un anticorps secondaire anti-lapin-PE et analysé par cytométrie de flux. En E est représentée l'expression de la fluorescence de Venus grâce à la complémentarité V1+V2. Le traitement à la doxycycline est fait pendant 24 heures à une concentration de 2µg/ml.

D'après les résultats obtenus il y a moins de cellules qui expriment Ikaros dans la lignée BH1 rtTA Ikaros-V1 que dans la lignée BH1 rtTA Ikaros-V2. La différence peut s'expliquer par une plus faible efficacité du virus ou par une mauvaise sélection antibiotique. Cependant, cette stratégie de transduction semble être la bonne puisque les cellules BH1 rtTA Ikaros-V1/Ikaos-V2 expriment une fluorescence de Venus après le traitement à la doxycycline.

Malgré la présence constante des antibiotiques dans le milieu de culture la sélection des cellules n'est pas maintenue à un niveau stable (Figure RI.4.A). Il y a

une contre-sélection des cellules qui sont transduites avec Ikaros-V1. Cette absence du maintien de la sélection au cours du temps peut être due au fait que l'hygromycine est un antibiotique qui agit lentement et qui ne tue pas les cellules par un arrêt de l'activité ribosomale mais par l'incorporation anormale d'acides aminés pendant la traduction menant à la production de protéines non conformes.

J'ai donc choisi de changer le gène de résistance à l'hygromycine dans le vecteur portant l'ADNc codant pour les protéines liées à la partie V1 par le gène BSD prodiguant la résistance à la blasticidine qui se montre aussi efficace que la puromycine pour le maintien en culture des cellules transduites (Figure RI.4.B).

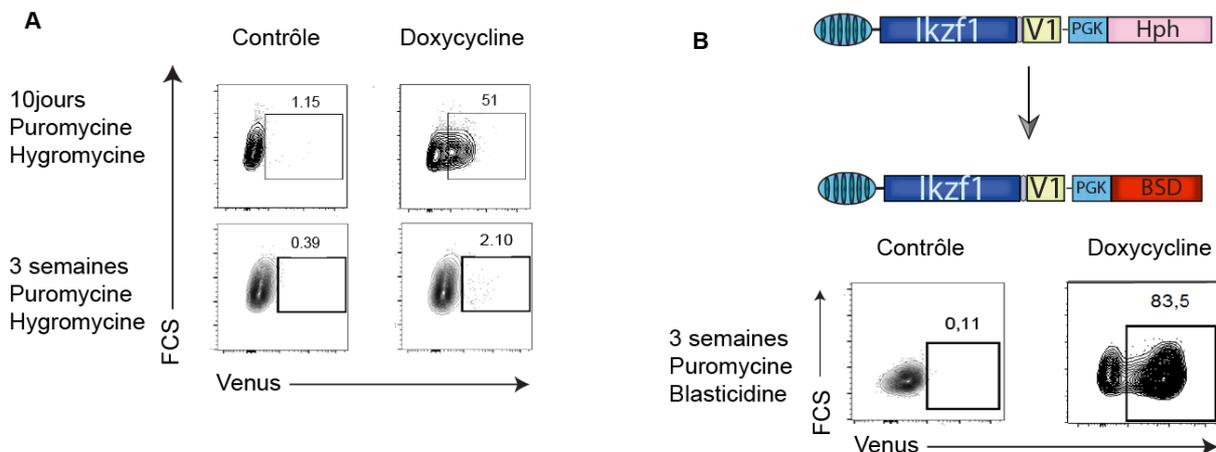


Figure RI.4 Remplacement de l'hygromycine par la blasticidine

(A) Les cellules BH1 ont été traitées pendant 24 heures avec la doxycycline (2µg/ml) après 10 jours et 3 semaines de sélection hygromycine (400mg/ml) / puromycine (2µg/ml), puis la fluorescence de Venus a été mesurée par cytométrie de flux. (B) Le gène de sélection Hph a été remplacé par le gène de résistance à la blasticidine (BSD) puis les cellules transduites ont été sélectionnées trois semaines avec la blasticidine (20µg/ml) et la puromycine (2µg/ml) et traitées avec la doxycycline (2µg/ml) pendant 24 heures. La fluorescence de Venus a été analysée par cytométrie de flux (B).

C. La présence de V1 et V2 n'empêche pas la localisation normale d'IKAROS dans les cellules

Dans les lymphocytes pré-B murins et donc dans les cellules BH1, Ikaros est localisé, dans le noyau, au niveau des *foci* de PC-HC. Pour contrôler la localisation nucléaire d'Ikaros lorsqu'il est lié à V1, V2 ou la protéine totale Venus, j'ai utilisé les trois lignées suivantes : BH1-Ikaros-ires-GFP qui, après le traitement à la doxycycline, exprime la protéine murine Ikaros et la GFP ; BH1-Ikaros-V1/Ikaros-V2 et BH1-Ikaros-Venus (Figure RI.5). La première lignée confirme bien qu'Ikaros se localise au niveau

dans le noyau en foci. Les trois autres lignées montrent que la localisation d'Ikaros ne semble pas perturbée par l'ajout des parties V1 et V2 ni par la protéine totale Venus.

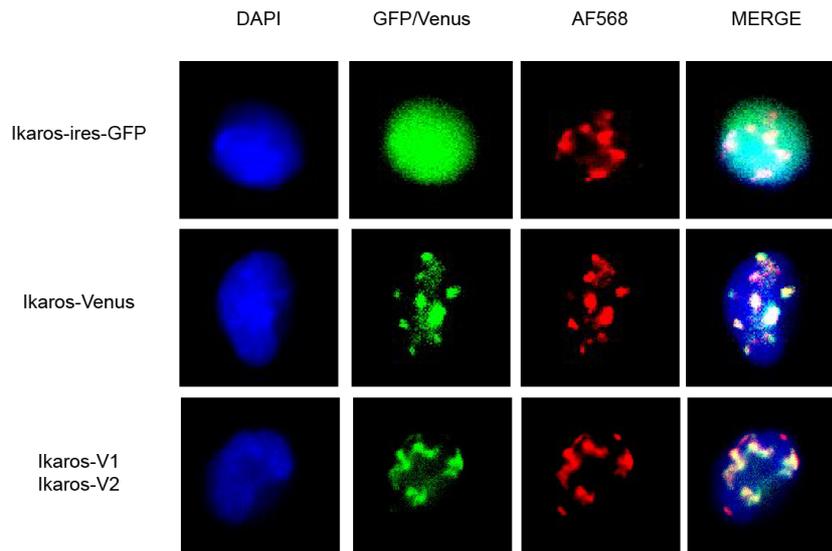


Figure RI.5 Visualisation de la localisation d'Ikaros par immunofluorescence

Les cellules BH1 transduites avec Ikaros-ires-GFP ou Ikaros-Venus ou Ikaros-V1 et Ikaros-V2 ont été traitées pendant 24 heures avec de la doxycycline (2µg/ml). Pour l'analyse par immunofluorescence, Ikaros a été marqué par un anticorps polyclonal de lapin et un anticorps secondaire anti-lapin-AF568 (rouge). La fluorescence de la GFP et de Venus est naturelle (vert)

D. V1 et V2 ne perturbent pas la fonctionnalité d'Ikaros dans les cellules BH1

Les cellules BH1 sont bloquées dans le processus de différenciation au stade pré-B. La réexpression de la forme entière d'Ikaros dans ces cellules permet de restaurer le processus de différenciation et il est alors possible de détecter des cellules au stade immature présentant un BCR après 5 jours.

J'ai donc procédé à une expérience de différenciation *in vitro* pour la lignée BH1-IkarosV1/IkarosV2. Pour cela j'ai cultivé les cellules avec ou sans doxycycline dans un milieu dépourvu d'IL-7. Après 5 jours, j'ai marqué les cellules avec un anticorps dirigé contre la chaîne lourde μ des immunoglobulines et un anticorps dirigé contre la chaîne légère κ , qui est la chaîne légère la plus fréquemment exprimée par les lymphocytes B murins (Figure RI.6). Ainsi, les cellules qui sont doublement positives sont celles qui expriment un BCR à leur surface et qui ont donc repris le processus de différenciation. L'analyse par cytométrie en flux montre que plus de 85% des cellules sont dans ce cas. Cette fonction spécifique d'Ikaros n'est donc pas altérée par la présence des fragments V1 et V2.

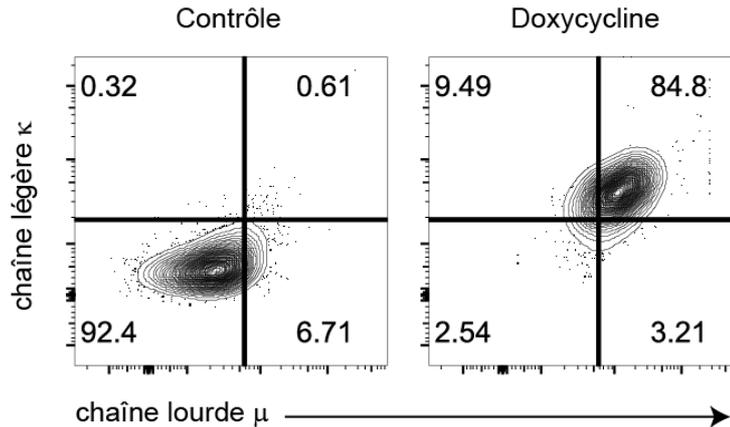


Figure RI.6 Différenciation des cellules BH1-IkarosV1/IkarosV2

Les cellules sont traitées pendant 5 jours avec ou sans doxycycline (2 μ g/ml). Les cellules sont ensuite marquées avec les anticorps dirigés contre les chaînes lourde μ et légère κ , respectivement couplés aux fluorochromes Cy5 et BV421. Les cellules doubles positives sont celles exprimant le BCR.

E. L'anticorps α Venus permet l'immunoprécipitation spécifique des dimères

Dans ce travail, je veux pouvoir étudier les spécificités fonctionnelles des dimères de la famille d'Ikaros. Pour cela, j'utilise la BiFC pour bénéficier de la capacité de l'anticorps α Venus à immunoprécipiter les dimères formés avec les fragments V1 et V2. Pour vérifier que cet anticorps est efficace, j'ai réalisé une expérience d'immunoprécipitation. Pour cela, j'ai utilisé une lignée BH1-IkarosV2 et la lignée BH1-IkarosV1/IkarosV2 (Figure RI.7.A). Cette expérience a été réalisée avec la lignée sélectionnée avec l'hygromycine B, ce qui explique la plus faible présence d'Ikaros-V1 dans l'Input. En revanche, cela permet de bien visualiser son enrichissement dans la condition de l'immunoprécipitation pour atteindre un ratio 1 :1 avec la forme Ikaros-V2 ce qui confirme la précipitation des homodimères. La lignée BH1-IkarosV2 confirme bien la spécificité de l'anticorps pour la protéine Venus complétée car la forme Ikaros-V2 n'est pas immunoprécipitée.

De plus, l'anticorps α Venus peut être utilisé pour immunoprécipiter la protéine Venus entière. Ainsi, j'ai réalisé une immunoprécipitation avec soit l'anticorps α Venus soit l'anticorps α Ikaros sur des extraits protéiques issus de la lignée BH1-Ikaros-Venus (Figure RI.7.B). Dans ce cas-là, les deux anticorps permettent de co-immunoprécipiter Ikaros-Venus avec la protéine Aiolos endogène. Aiolos interagit donc naturellement avec Ikaros et ne peut le faire que lorsque le domaine de dimérisation d'Ikaros est

disponible. Alors que la probabilité de former des hétérodimères entre une protéine surexprimée et une protéine endogène faiblement exprimée est assez faible, Ikaros-Venus se dimérise avec Aiolos endogène. Cela pose la question de l'affinité de dimérisation d'Ikaros avec lui-même ou avec Aiolos.

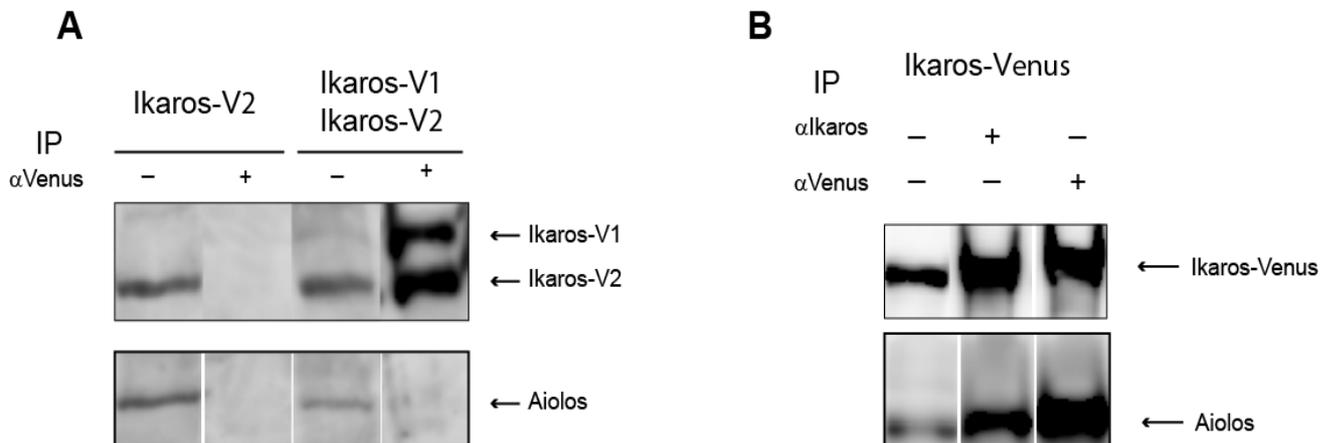


Figure RI.7 L'anticorps α Venus permet l'immunoprécipitation des dimères V1+V2
 (A) Les cellules des lignées BH1-Ikaros-V2 et BH1-Ikaros-V1/Ikaros-V2 ont été traitées pendant 24 heures avec la doxycycline (2 μ g/ml) puis protéines totales ont été extraites. L'immunoprécipitation a été réalisée avec l'anticorps α Venus Le western blot est fait avec des anticorps anti-Ikaros (polyclonal de lapin) et anti-Aiolos (monoclonal de lapin) et un anticorps anti-Lapin-HRP. (B) Immunoprécipitation avec les anticorps α Ikaros et α Venus sur des extraits protéiques de la lignée BH1-IkarosVenus. Le western blot est réalisé avec les mêmes anticorps qu'en A.

F. V1 et V2 ne force pas la dimérisation des facteurs de transcription

Pour terminer cette preuve de concept je me suis intéressée à l'impact de V1 et V2 sur la dimérisation des facteurs. Pour cela, j'ai supprimé, par mutagenèse, les deux doigts de zinc du domaine de dimérisation dans les constructions Ikaros-V1 et IkarosV2 (Figure RI.7.A). J'ai ensuite fait exprimer ces constructions dans des cellules COS-1 et réalisé l'extraction des protéines. Nous avons réalisé une expérience de retard sur gel en utilisant une sonde synthétique BS4 (Molnár and Georgopoulos, 1994) marquée au P32 qui est liée par Ikaros (Figure RI.7.B). Dans ces expériences, les complexes ADN/protéines migrent sur un gel d'acrylamide en fonction de leurs tailles et sont révélés par autoradiographie. Les formes WT d'Ikaros migrent en un complexe majeur alors que les formes mutées migrent en deux complexes distincts. Alors que les formes WT semblent se fixer à l'ADN sous forme de dimère de taille unique, il semble donc que les protéines mutées se fixent à l'ADN sous forme de

monomère, il y a donc deux complexes de tailles différentes car V1, 157 acides aminés, est plus grand que V2, 81 acides aminés. En utilisant l'anticorps α Venus pendant l'incubation des protéines avec la sonde cela alourdit le complexe formé par V1 et V2 qui migrera donc moins vite que les complexes non fixés par l'anticorps. Ainsi, la majorité du complexe migre moins loin que précédemment ce qui prouve bien que la majorité des complexes formés et liant l'ADN sont les dimères formés avec V1 et V2.

Dans le cas des protéines qui n'ont plus le domaine de dimérisation, la migration des complexes reste inchangée en présence de l'anticorps α Venus. Ceci montre que les complexes ne sont pas ciblés par l'anticorps α Venus et ne sont donc pas formés par des dimères comprenant V1 et V2. Les fragments V1 et V2 ne forcent donc pas la dimérisation des protéines.

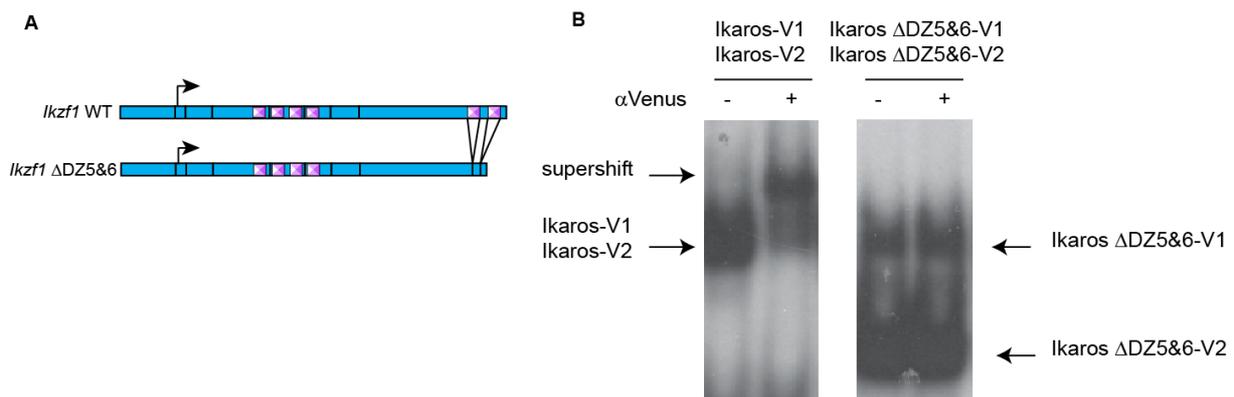


Figure RI.8 V1 et V2 ne forcent pas la dimérisation

(A) Présentation des formes WT et Δ DZ5&6 d'Ikaros. (B) Expérience de retard sur gel avec ou sans l'anticorps α Venus, réalisée par Philippe Kastner avec des extraits protéiques venant de cellules Cos-1 transfectées avec les constructions Ikaros-V1 et Ikaros-V2 WT ou Δ DZ5&6 et la sonde BS4.

Dans cette première partie de mes résultats, j'ai montré que le système de complémentarité bi-moléculaire de fluorescence était transposable dans la lignée de lymphocytes pré-B murins, BH1. Le couplage des fragments V1 et V2 ne semble empêcher ni la localisation ni la fonctionnalité d'Ikaros et donc, certainement, d'Helios et d'Aiolos. De plus, l'anticorps α Venus, permet l'immunoprécipitation spécifique des dimères. Enfin, il semble que V1 et V2 ne favorisent pas la dimérisation.

II. Étude de la liaison à l'ADN des facteurs de transcription de la famille d'Ikaros

L'objectif principal de cette partie du travail est de mettre en évidence des spécificités de liaison à l'ADN des dimères formés par les facteurs de transcription de la famille d'Ikaros. Bien que le système tet-ON permette de contrôler le temps de l'activation de la transcription et de l'expression des transgènes, cela reste un système de surexpression. J'ai, dans un premier temps, voulu vérifier que malgré une surexpression des facteurs de transcription de la famille d'Ikaros, nous serions en mesure de mettre en évidence des spécificités de liaison à l'ADN pour chacun des facteurs.

A. Aiolos a des spécificités de liaison à l'ADN non partagées avec Ikaros et Helios

1. Le profil de liaison à l'ADN d'Aiolos diffère de ceux d'Ikaros et d'Helios

Pour vérifier que la surexpression d'Ikaros, Helios et Aiolos me permettrait tout de même de déterminer certaines spécificités de liaison à l'ADN de chacun des facteurs, j'ai utilisé les lignées BH1-Ikaors-V2, BH1-Helios-V2 et BH1-Aiolos-V2. J'ai comparé les niveaux d'expression entre les traitements contrôle et doxycycline après 24 heures (Figure RII.1.A). Les cellules BH1 sont bien déficientes pour Ikaros, tandis qu'elles expriment Helios et Aiolos. L'expression d'Helios dans les cellules traitées à la doxycycline est augmentée de 3,8 fois et l'expression d'Aiolos de 5,4 fois. J'ai ensuite réalisé une expérience d'immunoprécipitation de chromatine suivie de séquençage haut-débit (ChIP-seq). La grande majorité des pics identifiés pour chaque facteur est commune avec les deux autres. Il y a néanmoins la possibilité d'isoler des pics qui semblent être spécifiques pour chaque facteur (Figure RII.1.B). La répartition générale des pics est la même pour les trois facteurs de transcription. Les pics sont principalement distribués au niveau des promoteurs, des introns et des régions intergéniques (Figure RII.1.C). La comparaison des profils de liaison à l'ADN d'Ikaros, Helios et Aiolos, sur une large portion du génome (5Mb), confirme que la plupart des

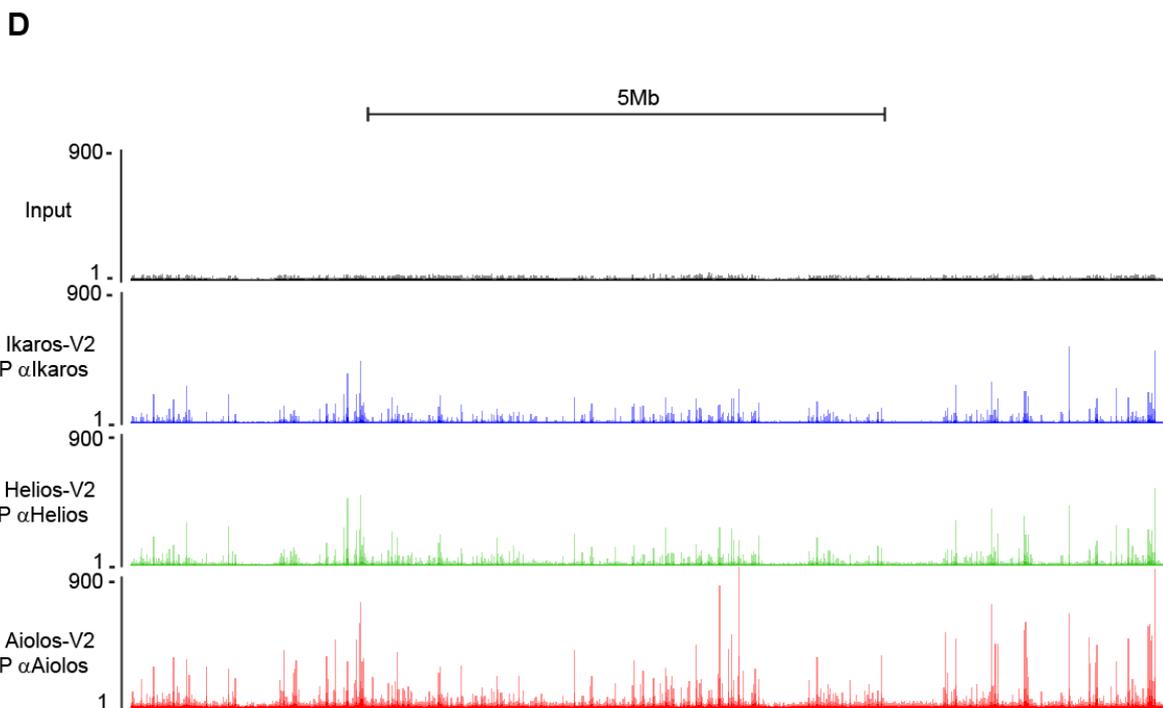
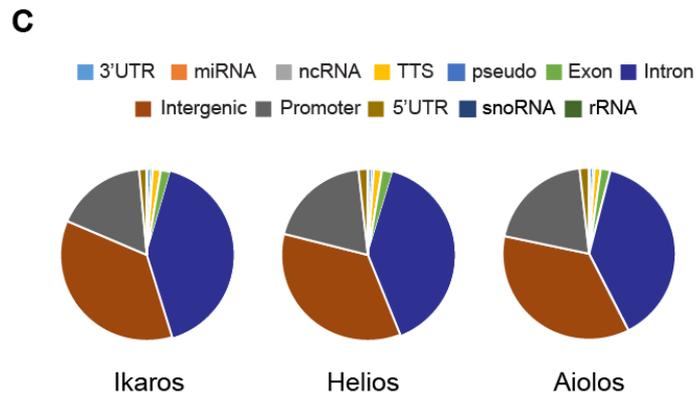
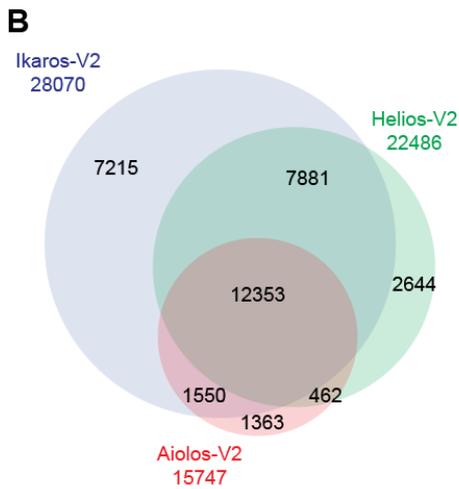
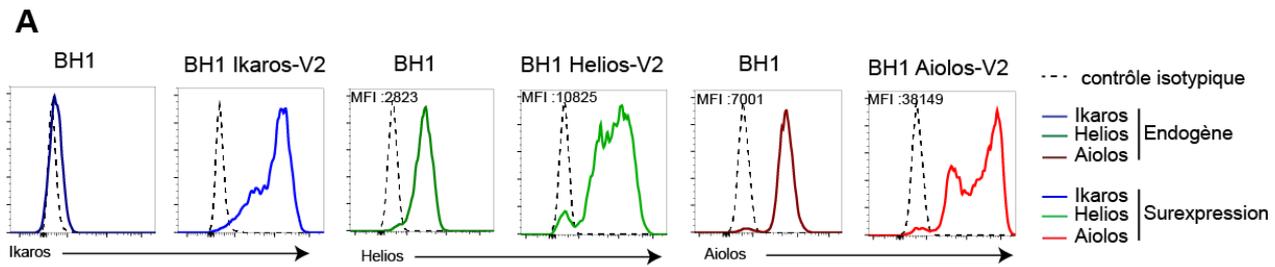


Figure RII.1 Analyse de la liaison à l'ADN d'Ikaros, Helios et Aiolos dans les cellules BH1
 (A) Marquage intracellulaire des cellules BH1 avec les anticorps dirigés contre Ikaros (bleu foncé), Helios (vert foncé) ou Aiolos (rouge foncé) et des cellules BH1-Ikaros-V2 (bleu), BH1-Helios-V2 (vert) et BH1-Aiolos-V2 (rouge) après 24 heures de traitement avec la doxycycline (2 μ g/ml) avec les mêmes anticorps. (B) Diagramme de Venn représentant les intersections entre les trois groupes de pics identifiés par ChIP-seq. (C) Répartition sur le génome des pics de ChIP-seq. (D) Profil de liaison à l'ADN des protéines Ikaros-V2, Helios-V2 et Aiolos-V2. Les flèches représentent l'emplacement de pics spécifiques du profil d'Aiolos.

pics sont communs (Figure RII.1.D). Les profils de liaison à l'ADN d'Ikaros et d'Helios sont pratiquement identiques et semblent s'opposer à celui d'Aiolos qui présente des pics plus forts et des pics spécifiques, absents des profils d'Ikaros et d'Helios.

2. Les gènes spécifiques des lymphocytes B sont des cibles privilégiées d'Aiolos

Nous avons isolé les pics spécifiques d'Aiolos, c'est-à-dire les pics uniquement présents dans le profil d'Aiolos ou beaucoup plus forts par rapport à Ikaros et Helios dans le groupe des pics communs. Nous nous sommes rendus compte qu'un grand nombre de gènes associés à ces pics sont des gènes spécifiques des lymphocytes B. J'ai alors analysé les profils de liaison à l'ADN des trois facteurs à proximité de certains gènes spécifiques des lymphocytes B (Figure RII.2). Certains de ces gènes, tels que *Cish*, *Igll1* ($\lambda 5$) et *Cd40* sont communément liés par les trois facteurs. D'autres comme *Il7ra* et *IL4ra* sont fixés par les trois facteurs mais Aiolos semble avoir une fixation plus forte. Enfin, les régions régulatrices des gènes *Tnfrsf13c* (Baff-R), *Cd19*, *Sh2b3* (LNK) et *Ighj4* sont uniquement liées par Aiolos et ces pics spécifiques sont, pour la plupart, beaucoup plus enrichis, plus forts, que les pics communs.

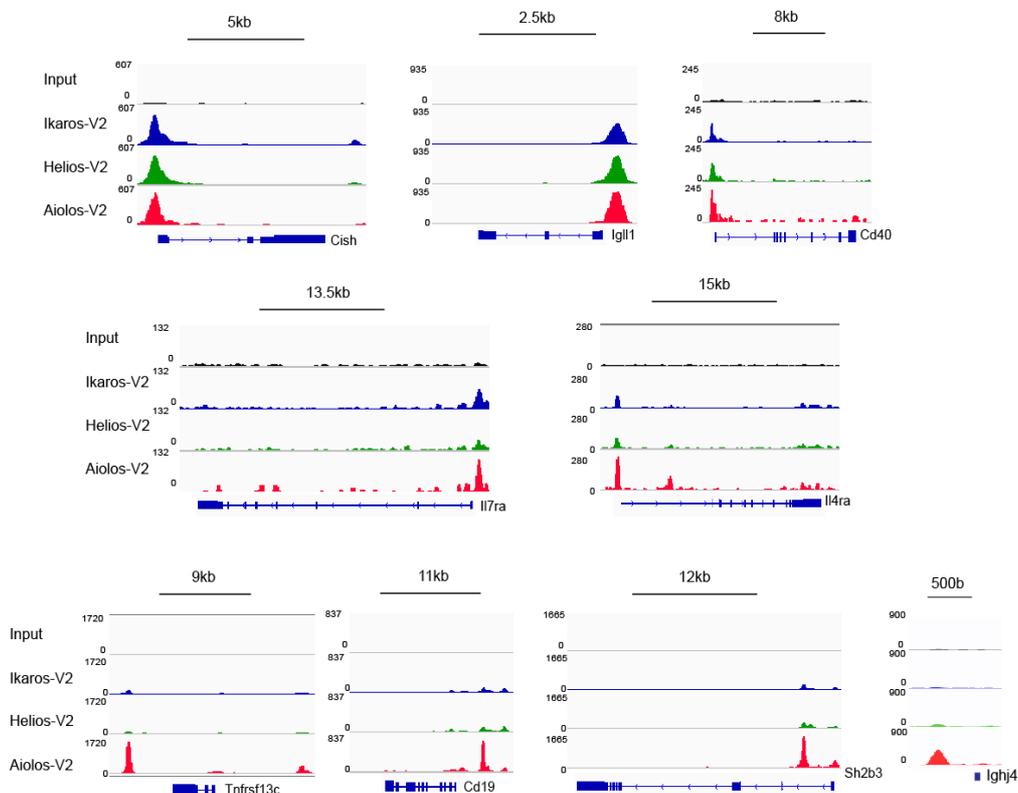


Figure RII.2 Les gènes spécifiques des lymphocytes B sont des cibles privilégiées d'Aiolos
 Profils de liaison à l'ADN des protéines Ikaros-V2, Helios-V2 et Aiolos-V2 à proximité de gènes spécifiques des lymphocytes B.

3. La spécificité de liaison d'Aiolos n'est pas due à sa surexpression

Pour tester si cette spécificité de liaison pouvait être due à la surexpression d'Aiolos, nous avons comparé les capacités de liaison *in vitro* des protéines Ikaros et Aiolos grâce à une expérience de retard sur gel (EMSA – Electrophoretic Mobility Shift Assay). Pour cette expérience, nous avons utilisé comme contrôles de liaison la sonde synthétique BS4 et une sonde dérivée de la séquence génomique liée par Ikaros et Aiolos au promoteur du gène *Cish* (Heizmann et al., 2020). Puis, nous avons comparé les capacités de liaison à l'ADN d'Ikaros et d'Aiolos sur des sondes dérivées des séquences spécifiquement liées par Aiolos au niveau des gènes *Tnfrsf13c*, *Ighj4* et *Sh2b3* (Figure RII.3.A). Pour les trois sondes testées, Ikaros se fixe moins bien qu'Aiolos voire pas du tout, pour la sonde *Ighj4*. Ces résultats montrent que ces différents gènes sont des cibles privilégiées d'Aiolos et que cette fixation n'est pas due à une surexpression de la protéine.

Pour vérifier que ces pics spécifiques d'Aiolos peuvent être observés sans expression ectopique de la protéine, j'ai comparé le profil de liaison à l'ADN d'Aiolos dans les cellules BH1-Aiolos-V2 à celui obtenu dans des lymphocytes B WT murins exprimant les plus forts niveaux d'Aiolos, à savoir les lymphocytes B spléniques (Figure RII.3 B et C). Plus de 50% des pics obtenus dans les splénocytes sont communs avec ceux détectés dans les cellules BH1. Parmi les pics communs, j'ai retrouvé les pics spécifiques découverts précédemment. Ainsi, les pics proches des gènes *Tnfrsf13c*, *Sh2b3* ou *IL4ra* sont tout à fait semblable à ceux obtenus dans les splénocytes. De plus, nous pouvons observer quelques différences dans les profils de liaison. En effet, il y a deux pics au niveau du promoteur *Sh2b3*, le premier est favorisé dans les BH1 et le second dans les splénocytes. Il y a également un pic supplémentaire dans les splénocytes à proximité du gène *IL4ra* et enfin le profil de liaison d'Aiolos dans les splénocytes au niveau des gènes *Ighm* et *Ighj4* montre une préférence pour le promoteur d'*Ighm* dans les splénocytes et pour *Ighj4* dans les BH1. Ainsi, il semble y avoir des signatures « splénocyte » et « BH1 ». Ces différences montrent bien que le profil de liaison à l'ADN d'Aiolos dans la lignée BH1-Aiolos-V2 n'est pas dû à une surexpression de la protéine mais bien à une spécificité de liaison d'Aiolos qui n'est pas partagée avec Ikaros et Helios.

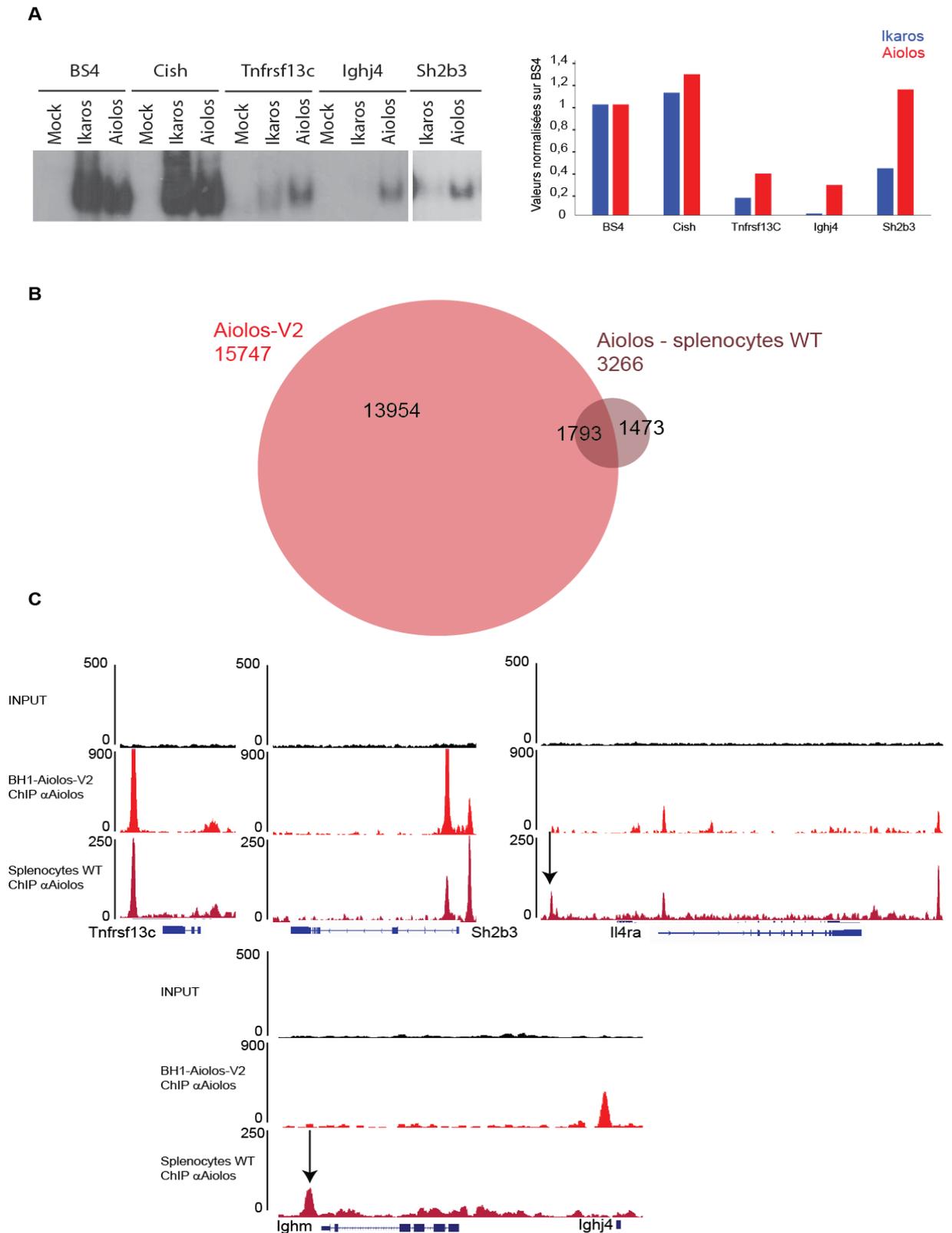


Figure RII.3 Les spécificités de liaison à l'ADN d'Aiolos ne sont pas due à la surexpression

(A) Expérience de retard sur gel et quantification réalisée par P.Kastner, avec des extraits protéiques issus de cellules COS-1 transfectées avec Ikaros ou Aiolos ou vecteur vide. (B) Diagramme de Venn représentant l'intersection entre les groupes de pics issus de ChIP-seq α Aiolos sur des cellules BH1-Aiolos-V2 ou des lymphocytes B spléniques de souris WT. (C) Profils de liaison de la protéine Aiolos dans les cellules BH1-Aiolos-V2 (rouge) et dans les lymphocytes B spléniques (rouge bordeaux) à proximité de gènes spécifiques des lymphocytes B.

4. Les protéines Aiolos et Helios endogènes se colocalisent avec les protéines surexprimées

Les cellules BH1 expriment de faibles niveaux d'Helios et d'Aiolos (Figure RII.1.A). J'ai donc voulu savoir si ces protéines se fixent à l'ADN de la même façon que les protéines surexprimées. J'ai donc réalisé des expériences de ChIP-seq sur des cellules BH1 non transduites avec les anticorps spécifiques des protéines Helios et Aiolos. Près de 80% des pics d'Aiolos endogène sont communs avec Helios endogène, qui lui en partage près de 35% de ses pics avec Aiolos (Figure RII.4.A). Si ces protéines sont retrouvées aux mêmes sites de liaisons, cela pourrait refléter l'existence d'hétérodimère Aiolos/Helios. Bien que les protéines Helios et Aiolos endogènes se fixent communément au promoteur du gène *Igll1* comme les protéines surexprimées, de façon surprenante Aiolos endogène ne se fixe pas au niveau des gènes *Sh2b3* et *Tnfrsf13c* qui sont parmi les cibles les plus fortement liées par la protéine surexprimée (Figure RII.4.B). Aiolos endogène partage près de 80% de ses pics avec la protéine surexprimée Aiolos-V2. Ces pics font partie des pics communs entre les trois protéines Ikaros-V2, Helios-V2 et Aiolos-V2. Les 20% de pics différents d'Aiolos-V2 sont en fait répartis dans les groupes des pics communs Ikaros-V2/Helios-V2 ou les pics isolés d'Ikaros-V2 et d'Helios-V2. Finalement seuls 37 pics semblent être seulement attribués à Aiolos endogène (Figure RII.4.C). De la même façon, 96% des pics d'Helios endogènes sont partagés avec Helios-V2 et la grande majorité fait également partie du groupe des pics communs entre les trois protéines surexprimées. Ces protéines sont exprimées à de faibles niveaux ce qui ne favorise pas la formation des homodimères contrairement à la surexpression. Il est alors possible que les protéines endogènes se fixent à l'ADN sous forme de monomère et non de dimère. Finalement, les protéines endogènes se fixent à l'ADN au niveau de régions également ciblées par les protéines surexprimées. Il est donc possible que ces protéines forment des dimères avec les protéines surexprimées.

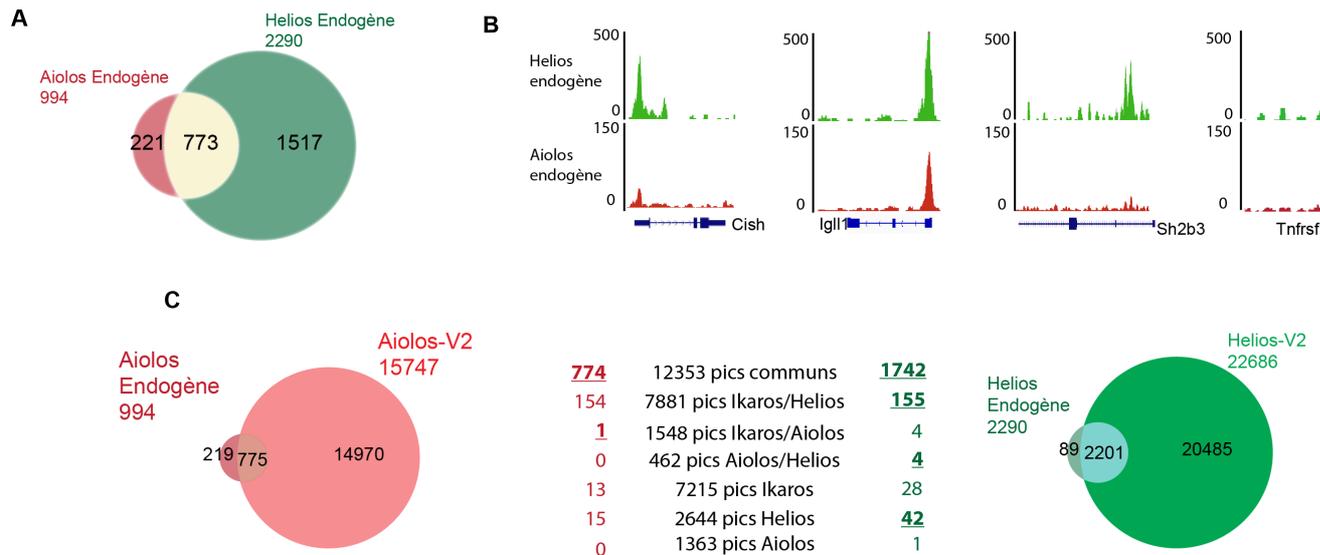


Figure RII.4 Les protéines endogènes Helios et Aiolos ont des sites de liaison à l'ADN communs avec les protéines surexprimées.

(A) Diagramme de Venn représentant l'intersection des groupes de pics identifiés par ChIP-seq pour les protéines endogènes Aiolos et Helios dans les cellules BH1. (B) Profils de liaison à l'ADN des protéines endogènes Helios (vert clair) et Aiolos (rouge bordeaux) à proximité de gènes spécifiques des lymphocytes B. (C) Diagrammes de Venn représentant les intersections entre les groupes de pics identifiés par ChIP-seq pour Aiolos (rouge) dans les cellules BH1 et dans les cellules BH1-Aiolos-V2 et entre les groupes de pics identifiés pour Helios (vert) dans les cellules BH1 et les cellules BH1-Helios-V2. Distribution des pics communs (gras, soulignés) et des pics spécifiques des endogènes parmi les groupes de pics identifiés dans le diagramme de Venn (Figure RII.1.B)

5. Aiolos se fixe à ses séquences cibles sous forme d'homodimère et d'hétérodimère avec Ikaros

Pour m'affranchir des protéines endogènes et pouvoir étudier les spécificités de liaison à l'ADN des dimères de la famille d'Ikaros j'ai alors utilisé la complémentarité bi-moléculaire de fluorescence (BiFC) et l'anticorps spécifique α Venus. En plus de ne plus être perturbé par les protéines endogènes, les immunoprécipitations seront faites avec le même anticorps, et donc la même spécificité. Étant donné qu'Helios a un profil de liaison similaire à celui d'Ikaros et que ce facteur est destiné à être totalement réprimé au cours de la différenciation des lymphocytes B, je me suis intéressée aux dimères formés par Ikaros et Aiolos.

Au sein de la lignée BH1-Ikaros-V1/Ikaros-V2, trois homodimères peuvent se former : IkarosV1/Ikaros-V1, Ikaros-V2/Ikaros-V2 et Ikaros-V1/Ikaros-V2. Seul, le dernier homodimère sera immunoprécipité avec l'anticorps α Venus, ce qui ne révélera donc qu'un tiers du signal total. Il en va de même pour les deux autres lignées.

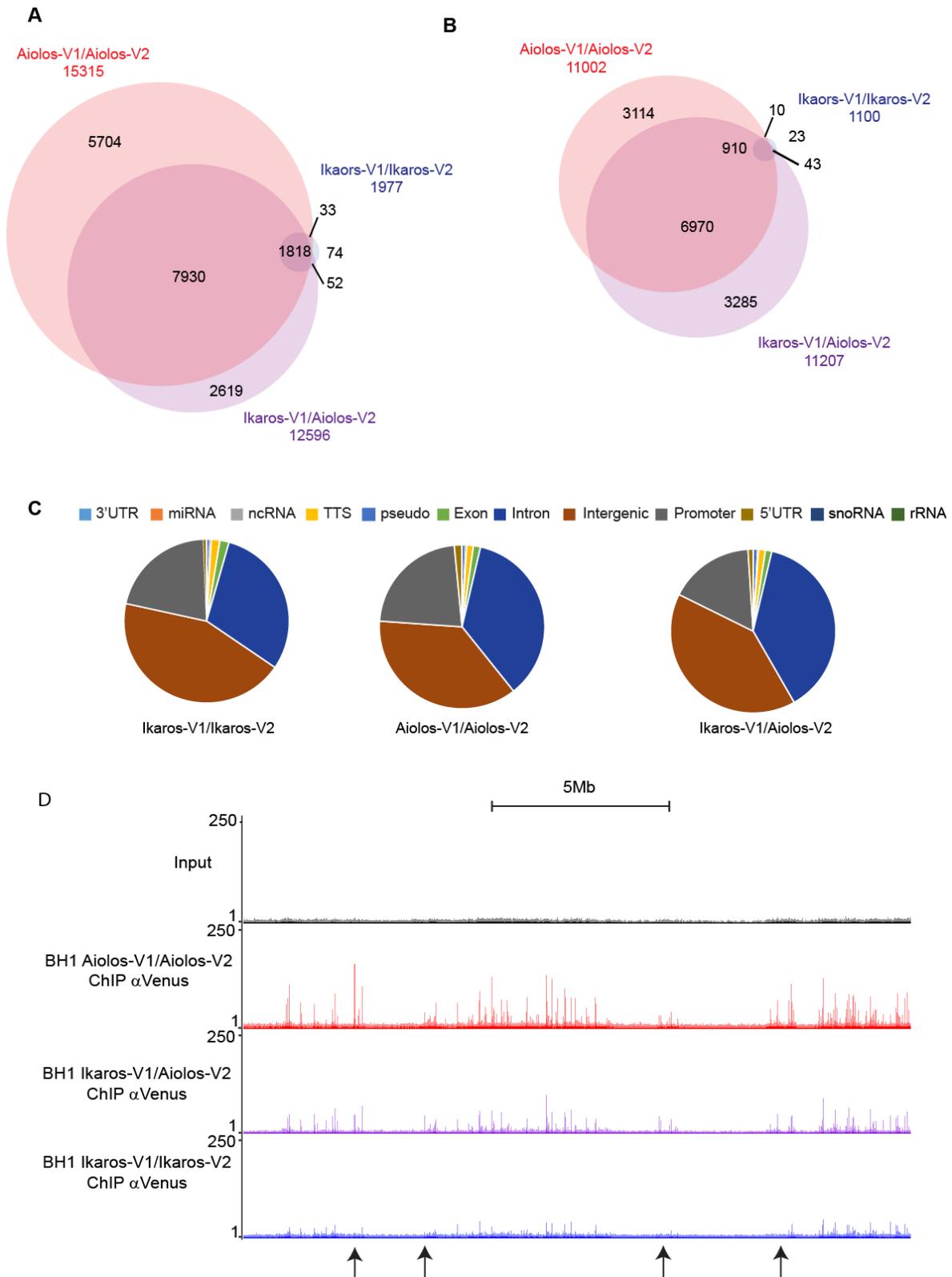


Figure RII.5 L'homodimère d'Ikaros ne se lie pas efficacement à l'ADN

(A) et (B) Diagrammes de Venn représentant les intersections des groupes de pics identifiés par ChIP-seq pour les homodimères d'Aiolos (rouge), les homodimères d'Ikaros (bleu) et les hétérodimères Ikaros/Aiolos (violet) après traitement des trois lignées cellulaires avec la doxycycline pendant 24 heures avec les doses normales (2 μ g/ml) et les doses adaptées (Ikaros-homodimère : 50ng/ml ; Aiolos-homodimères : 35ng/ml et Ikaros/Aiolos-hétérodimères : 40ng/ml) respectivement. (C) Répartition, sur le génome, des pics identifiés pour Ikaros-homodimère, Aiolos-homodimère et Ikaros/Aiolos hétérodimère. (D) Profils de liaison à l'ADN des trois dimères. Les flèches indiquent les pics absents du profil d'Ikaros-homodimère mais présents dans les deux autres profils.

Le résultat le plus surprenant dans cette expérience est la très faible proportion de pics identifiés pour l'homodimère d'Ikaros, seulement 1977, par rapport aux deux autres dimères (Figure RII.5.A). Pour confirmer ce résultat, j'ai réalisé un second ChIP-seq avec les mêmes cellules mais avec une dose réduite de doxycycline, qui sera détaillée plus tard (Figure RII.4.B). La répartition des pics de l'homodimère d'Ikaros est un peu différente de celle qui était observée pour Ikaros-V2 (Figure RII.5.C et RII.1.C). En effet, l'homodimère semble privilégier une liaison dans les zones intergéniques par rapport aux régions introniques. La répartition des pics de l'homodimère d'Aiolos est tout à fait comparable à celle d'Aiolos-V2. Ce qui tend à montrer que la surexpression d'Aiolos dans les cellules BH1 favorise bien l'homodimérisation. En revanche, la répartition des pics de l'hétérodimère a tendance à ressembler à celle de l'homodimère d'Aiolos mais également à celle de la surexpression d'Ikaros. Ceci suggère que certains pics identifiés précédemment pour Ikaros seraient peut-être dûs à une hétérodimérisation avec Aiolos endogène.

Les profils de fixations des homodimères d'Aiolos et des hétérodimères Ikaros/Aiolos sont très similaires entre-eux et différent de celui des homodimères d'Ikaros (Figure RII.5.D). En effet, les pics des homodimères d'Ikaros sont plus petits et l'on peut noter l'absence de certains pics qui sont présents chez les homodimères d'Aiolos et les hétérodimères. Si l'on regarde plus attentivement les profils de liaison à proximité des gènes spécifiques des lymphocytes B, nous pouvons remarquer que l'homodimère d'Ikaros est capable de se lier à l'ADN de façon équivalente des deux autres dimères aux régions régulatrices d'*Igll1* et d'*Ebf1* (Figure RII.6). Cependant il est absent des régions régulatrices des autres gènes spécifiques des lymphocytes B. Il semble donc qu'Ikaros puisse se fixer aux gènes cibles d'Aiolos mais seulement sous forme d'hétérodimère avec celui-ci. Aiolos apparaît alors comme le facteur dominant de l'hétérodimère en permettant sa fixation au niveau de ses gènes cibles.

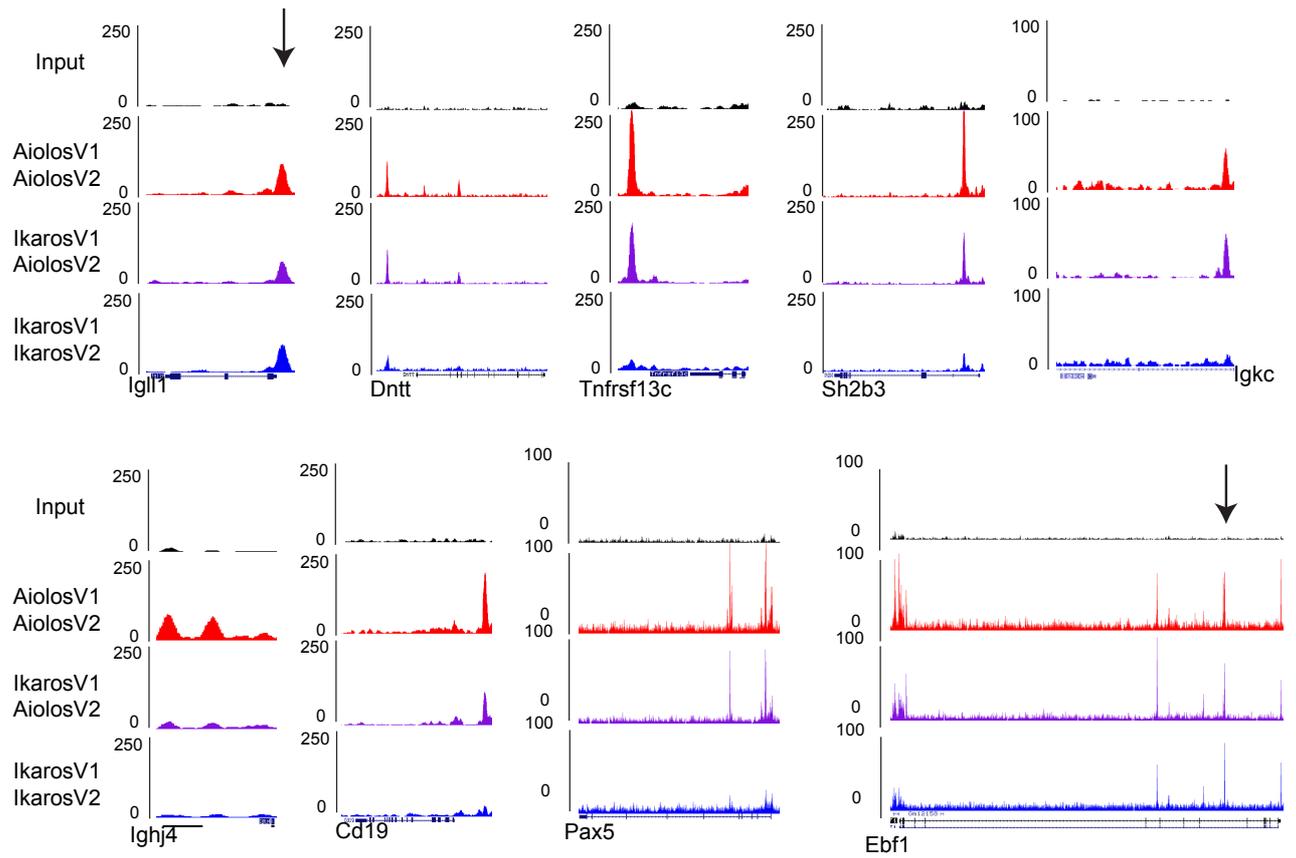


Figure RII.6 Les gènes spécifiques des lymphocytes B sont principalement ciblés par Aiolos
 Profils de liaison à l'ADN des homodimères d'Ikaros (bleu), des homodimères d'Aiolos (rouge) et des hétérodimères Ikaros/Aiolos (violet) à proximité de gènes spécifiques des lymphocytes B. Les flèches représentent des pics présents de façon équivalente dans les trois profils.

B. La spécificité de liaison à l'ADN d'Aiolos est déterminé par sa séquence protéique et son motif de liaison

La spécificité de liaison à l'ADN d'un facteur de transcription est conditionnée par deux paramètres : son domaine de liaison à l'ADN et son motif de liaison sur l'ADN. D'après les expériences de CHIP-seq, Aiolos semble avoir une particularité de liaison à l'ADN qu'il ne partage pas avec Ikaros. Pour comprendre d'où vient cette particularité nous avons testé ces deux paramètres.

1. Le domaine de liaison à l'ADN ne suffit pas à transmettre la spécificité de liaison

Pour étudier l'importance du domaine de liaison d'Aiolos dans sa spécificité de liaison à l'ADN, j'ai échangé entre-eux les exons 4, 5 et 6 portant l'intégralité des domaines de liaison d'Ikaros et d'Aiolos. Puis, nous avons réalisé une expérience d'EMSA en utilisant ces chimères d'Ikaros et d'Aiolos et les formes WT (Figure RII.7).

Les chimères sont toujours capables de se lier à la sonde synthétique contrôle BS4. Cependant, l'échange entre les domaines de liaison à l'ADN ne semble pas conférer à Ikaros la capacité de se lier aussi bien qu'Aiolos aux différentes sondes spécifiques testées. Le domaine de liaison à l'ADN ne semble donc pas être le seul responsable de la spécificité de liaison.

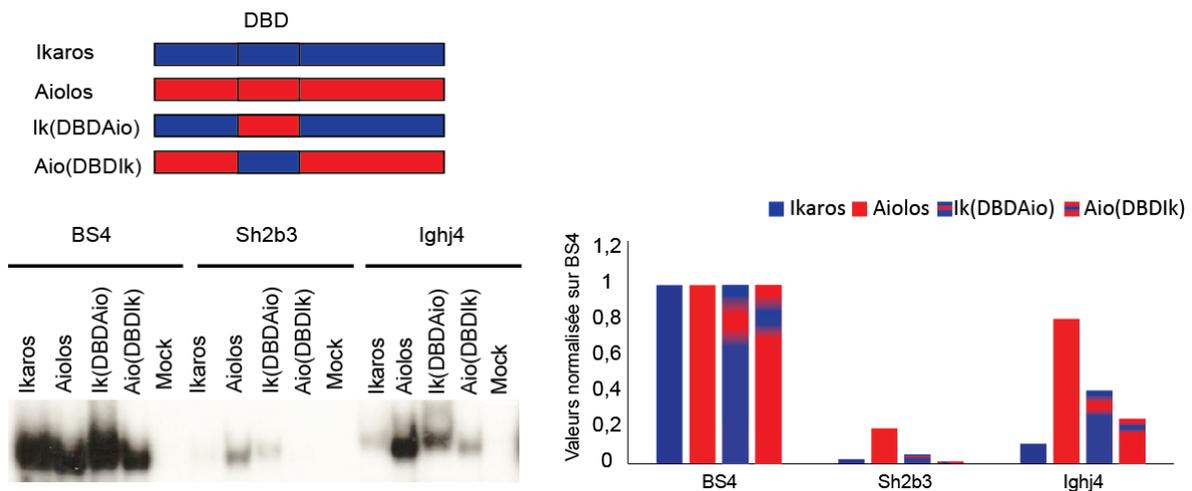


Figure RII.7 Apport du domaine de liaison dans la spécificité d'Aiolos

Expérience de retard sur gel et sa quantification réalisée par Philippe Kastner à partir d'extraits protéiques de cellules COS-1 transfectées avec les formes WT et chimériques (échange des domaines de liaison à l'ADN) des protéines Ikaros et Aiolos.

2. Le motif de liaison à l'ADN d'Aiolos diffère de celui d'Ikaros

Pour comprendre quelle pourrait être la part de la séquence d'ADN dans la spécificité de liaison à 'ADN d'Aiolos, j'ai isolé trois groupes de pics, à partir du premier ChIP-seq, puis j'ai utilisé le logiciel STREME pour identifier les motifs enrichis (Figure RII.8.A). Le motif le plus enrichi dans le groupe des pics communs aux trois facteurs est tout à fait semblable à celui retrouvé dans le groupe des pics spécifiques d'Ikaros.

Le motif enrichi dans les séquences spécifiquement liées par Aiolos diffère légèrement. Il semble que l'acide aminé en -2 du motif GGAA soit plus souvent une guanine qu'une cytosine. Pour vérifier l'importance des nucléotides entourant le motif principal GGAA, nous avons utilisé la sonde Sh2b3 et remplacé les nucléotides en positions -2,-1,+1 et +2 par ceux présents dans les sondes BS4, Cish et Ighj4 (Figure RII.8.B). La première chose que l'on remarque est que la sonde BS4 WT est plus efficacement liée par Ikaros et Aiolos que la sonde Sh2b3 avec le contexte BS4. Cependant la sonde BS4 possède deux motifs GGAA contre un seul pour la sonde Sh2b3. Néanmoins, le contexte BS4 sur la sonde Sh2b3 permet à Ikaros de se lier plus fortement que sur la sonde Sh2b3 WT. Il en va de même pour le contexte Cish. Les sondes BS4 WT et Cish WT étant normalement liées très fortement par les deux facteurs (Figure RII.3.A). En revanche le contexte Ighj4 abolit la liaison d'Ikaros sur la sonde Sh2b3 et rétablit la spécificité d'Aiolos. Ces résultats montrent donc que les deux nucléotides immédiatement en 5' et en 3' du motif GGAA sont déterminants pour spécifier la liaison préférentielle d'Aiolos. À la lueur de ces résultats, j'ai répertorié entre 80 et 130 des pics les plus forts dans chaque groupe (communs, Ikaros spécifique, Aiolos spécifique) et j'ai relevé, sous chaque pic, les nucléotides en position -2,-1,+1 et +2 des motifs GGAA dans une fenêtre de 90 nucléotides sous le sommet du pic et dans une limite de 5 motifs. Ainsi, les motifs générés avec le programme *ggseqlogo* permettent de mettre en évidence l'importance que chaque base dans la définition du motif mais également la distribution des nucléotides pour chaque position (Figure RII.8.C). Ainsi, pour les groupes des pics communs et des pics d'Ikaros la position +2 ne semble pas avoir d'importance. Pour les trois autres positions, la distribution des nucléotides reste assez semblable entre les pics communs et les pics d'Ikaros. D'un autre côté, les motifs relevés pour les pics d'Aiolos, semblent donner une plus grande

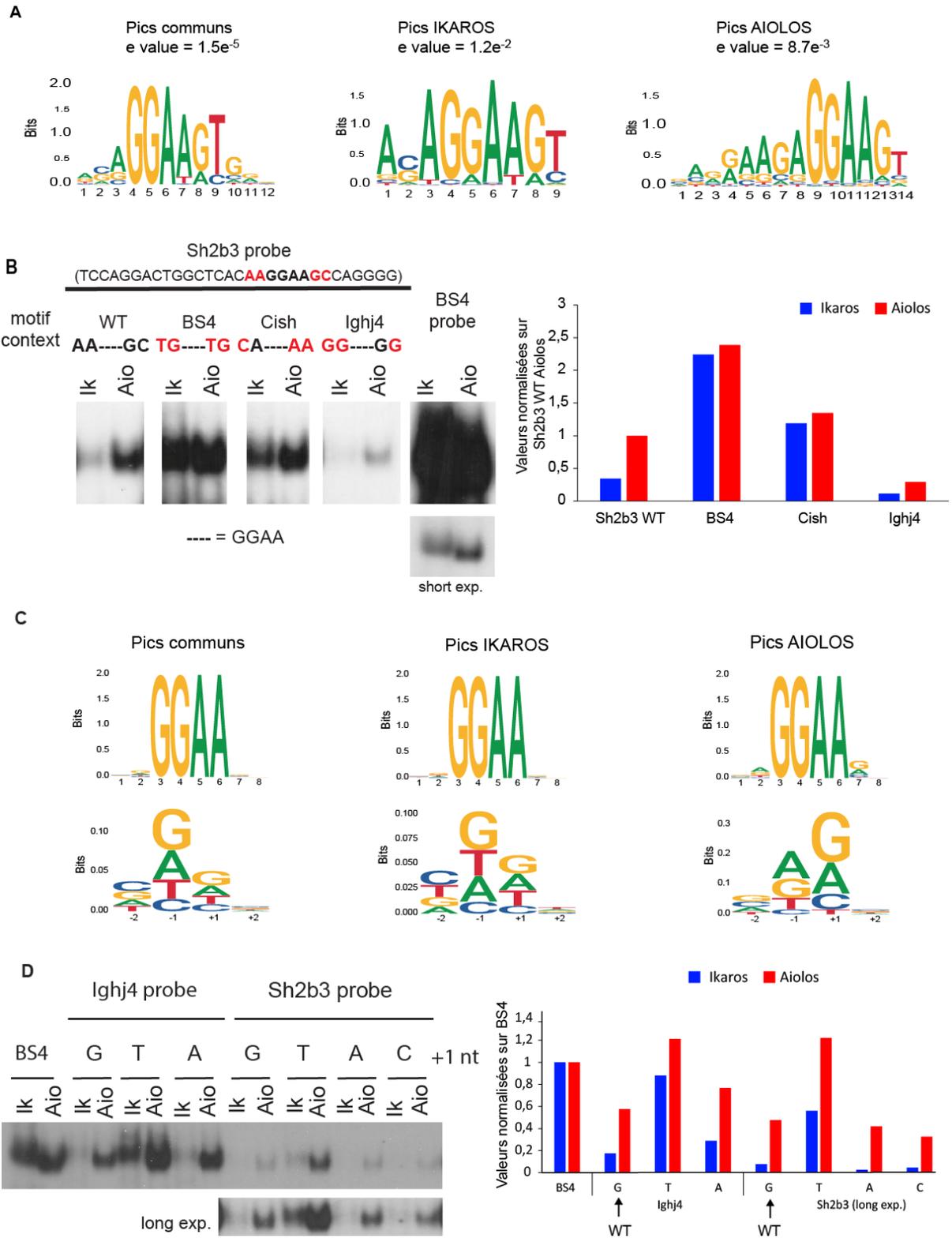


Figure II.8 Le motif de liaison d'Aiolos diffère de celui d'Ikaros

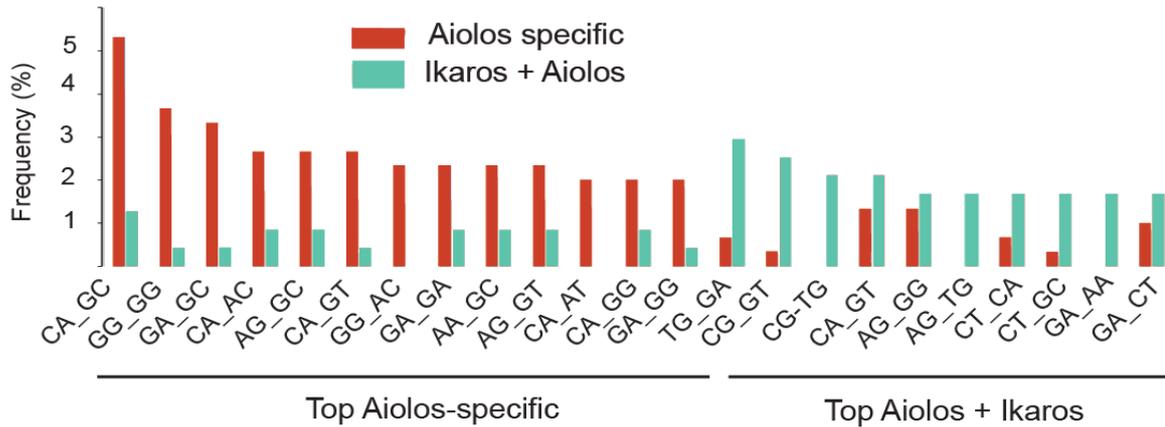
(A) Les motifs de liaison à l'ADN enrichis dans les pics communs, les pics spécifiques d'Ikaros et les pics spécifiques d'Aiolos découvert par STREME. (B). Expérience de retard sur gel réalisée par Philippe Kastner avec substitution des nucléotides aux positions -2, -1, +1 et +2 du GGAA de la sonde Sh2b3 et quantification du signal. (C). Motifs et enrichissement aux positions -2, -1, +1 et +2, dans les pics communs, les pics spécifiques d'Ikaros et les pics spécifiques d'Aiolos, relevés sur les 80 (pics communs) ou 130 (Ikaros et Aiolos) pics les plus forts. (D). Expérience de retard sur gel réalisée par Philippe Kastner avec substitution des nucléotides en position +1 du GGAA sur les sondes lghj4 et Sh2b3. Pour la sonde lghj4, la substitution par un C a engendré la liaison non spécifique d'une protéine nucléaire utilisant toute la sonde et empêchant la fixation de nos protéines d'intérêt.

importance aux positions -1 et +1. À ces positions, nous pouvons remarquer une déplétion de la cytosine en -1 et une déplétion de la thymine en position +1. Ce résultat est très étonnant car le contexte de la sonde BS4 présente un T en +1 et est fortement liée par Aiolos. Ceci suggère qu'Aiolos fait une contre-sélection d'un nucléotide préférentiel pour établir une spécificité de liaison qui ne sera donc pas partagée avec Ikaros. Pour tester cette hypothèse, nous avons remplacé les nucléotides en position +1 dans les sondes Ighj4 et Sh2b3 (Figure RII.8.D). Dans les deux sondes, le contexte WT possède un G en position +1. Cette configuration ne permet pas à Ikaros de se lier sur la sonde. Néanmoins, lorsque l'on place un T en +1, deux phénomènes s'observent. Le premier est que dans ce nouveau contexte Ikaros est capable de se lier sur les sondes Ighj4 et Sh2b3. Le deuxième est que la fixation d'Aiolos est plus forte que dans le contexte WT. De plus, le remplacement G>A et G>C, ne semble pas perturber la fixation d'Aiolos. Ceci tend à démontrer que l'absence de T en +1 joue un rôle important pour assurer la spécificité d'Aiolos bien que ce nucléotide améliore sa fixation à l'ADN.

Bien qu'Ikaros puisse normalement se fixer lorsqu'il y a un A en +1, comme dans la sonde Cish, dans le contexte Sh2b3, Ikaros reste incapable de se fixer. Mais la position +1 n'est pas la seule position qui diffère entre Cish et Sh2b3. D'après les motifs observés, la position +1 semble être décisive pour Aiolos mais pour Ikaros il s'agirait surtout de la position -1. Nous avons, alors, cherché à savoir si les combinaisons des quatre positions pouvaient avoir une importance. Pour cela nous avons relevé les combinaisons des quatre positions les plus fréquentes sur les pics Aiolos et sur les pics communs Ikaros/Aiolos (Figure RII.9.A). Ainsi il semble que certaines combinaisons soient enrichies dans les pics Aiolos par rapport aux pics communs entre Ikaros et Aiolos. Dans l'ensemble il semble bien y avoir des combinaisons prédominantes dans les pics d'Aiolos et d'autres qui favorisent la fixation des deux facteurs. Une nouvelle fois, ces résultats montrent que la présence du T en +1 fait uniquement partie des combinaisons partagées par Ikaros et Aiolos. Comme précédemment, nous avons testé ces différentes combinaisons en les introduisant comme contexte du motif GGAA de la sonde Sh2b3 (Figure RII.9.B). Cette expérience confirme qu'Ikaros ne se fixe pas, ou beaucoup moins bien, aux sondes comportant les combinaisons enrichies dans les pics d'Aiolos. De plus, les deux facteurs sont bien capables de se fixer de façon comparable sur les sondes comportant les combinaisons enrichies dans les pics communs Ikaros/Aiolos.

Ce travail a permis de mettre en évidence une particularité d'Aiolos. En effet, il semble qu'Aiolos fasse une contre sélection d'un nucléotide favorisant sa fixation, le T en +1, pour établir sa spécificité de liaison à l'ADN. Cependant, il reste difficile de prédire, en regardant la séquence d'ADN, s'il s'agit d'un motif commun, ou spécifique d'Ikaros ou d'Aiolos.

A



B

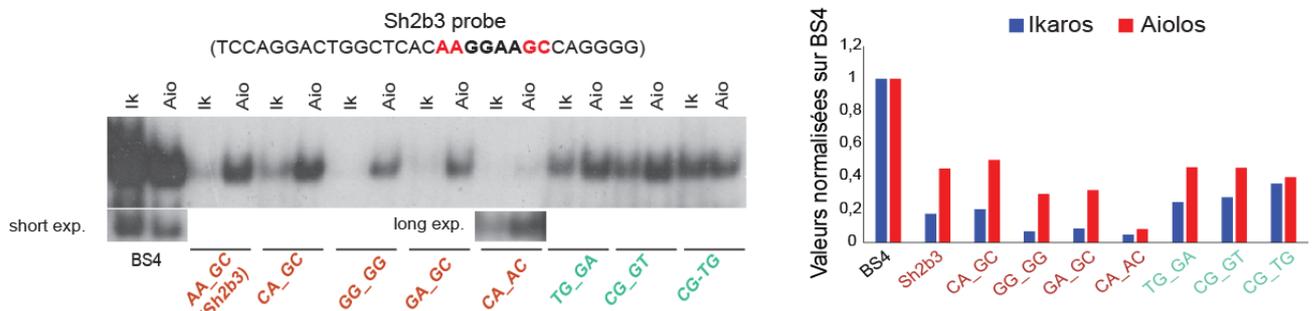


Figure II.9 Enrichissement des combinaisons de nucléotides entourant le motif GGAA dans les pics spécifiques Aiolos et les pics communs Ikaros/Aiolos

(A) Représentation des enrichissements des différents combinaisons les plus fréquentes dans les motifs relevés sous les pics spécifiques d'Aiolos et les pics communs Ikaros/Aiolos. (B) Expérience de retard sur gel et sa quantification réalisée par Philippe Kastner montrant la liaison d'Ikaros et d'Aiolos en fonction des combinaisons testées.

III. Étude de la fonctionnalité des dimères de la famille d'Ikaros

Pour tenter de mettre en évidence des spécificités fonctionnelles des facteurs de transcription de la famille d'Ikaros, j'ai analysé trois fonctions de ces facteurs. Tout d'abord, leur rôle dans la régulation du statut de la chromatine, puis la régulation de l'expression de leurs gènes cibles et enfin leur capacité à permettre la différenciation des cellules BH1. Mais dans un premier temps, j'ai adapté le traitement à la doxycycline pour essayer de me rapprocher d'une expression physiologique d'Ikaros et d'Aiolos.

A. Comparaison de l'expression d'Ikaros et d'Aiolos dans le système tet-ON avec les cellules primaires

1. Les facteurs de transcription sont fortement surexprimés

Pour comparer l'expression d'Ikaros et d'Aiolos dans les BH1 avec leurs expressions *in vivo*, nous avons isolé les lymphocytes pro-B, pré-B et immatures de la moelle osseuse ainsi que ceux de la rate d'une souris WT. Alors que les deux facteurs de transcription atteignent leurs niveaux les plus élevés dans les splénocytes, leurs expressions dans les BH1 avec le système tet-ON induit une expression encore plus forte après 24 heures de traitement avec la dose normale de doxycycline. (Figure RIII.1.A). Pour réduire cette différence entre la physiologie et le système tet-ON, j'ai adapté deux paramètres : la dose de doxycycline et le temps de traitement.

2. Adaptation de la dose de doxycycline

Le premier paramètre du système tet-ON que j'ai fait varier pour me rapprocher des expressions physiologiques est la dose de doxycycline. J'ai donc drastiquement diminué les doses de doxycycline et je les ai adaptées pour chacune des lignées (Figure RIII.1.B). Les doses adaptées de doxycycline sont donc de 50ng/ml pour la lignée Ikaros, de 35ng/ml pour la lignée Aiolos et de 40ng/ml pour la lignée Ikaros/Aiolos. Les expressions d'Ikaros et d'Aiolos avoisinent maintenant les

expressions trouvées dans les splénocytes. Bien qu'il soit possible de se rapprocher encore plus des expressions physiologiques, la diminution de la dose de doxycycline rend la détection de la fluorescence de Venus impossible en cytométrie. C'est pourquoi j'ai décidé de ne pas descendre en dessous des doses mentionnées ici.

3. Cinétique de formation des dimères

Le deuxième paramètre que j'ai fait varier avant d'étudier le rôle des facteurs sur la régulation du statut de la chromatine et de la transcription de leurs gènes cibles est le temps d'exposition à la doxycycline. J'ai tout d'abord vérifié la localisation des dimères au niveau des *foci* de PC-HC (Figure RIII.1.C). Les trois dimères ont une localisation concordante avec celle attendue et les facteurs de transcription colocalisent parfaitement avec la fluorescence de Venus. J'ai ensuite réalisé une cinétique de traitement à la doxycycline et j'ai observé par microscopie l'apparition du « pattern » de localisation des dimères au niveau des *foci* de PC-HC (Figure RIII.1.D). Pour cela j'ai utilisé la dose normale de doxycycline (2µg/ml) mais aussi les doses adaptées de doxycycline. Malgré l'apparition du pattern après 6 heures de traitement avec la dose normale de doxycycline pour les homodimères d'aiolos puis à 14 heures avec la dose adaptée de doxycycline et pour les homodimères d'ikaros avec la dose normale de doxycycline, il est absent de la lignée des hétérodimères avant 16 heures d'activation. Après 16 heures de traitement avec les deux doses de doxycycline, le pattern de localisation des dimères est détecté pour les trois lignées. J'ai donc choisi ce temps d'activation pour étudier la fonctionnalité des facteurs de transcription et pouvoir comparer les lignées entre elles.

Ainsi, pour que l'expression des facteurs de transcription de la famille d'Ikaros dans les cellules BH1 soit proche d'une expression existante physiologiquement, j'ai utilisé des doses adaptées de doxycycline pour chacune des trois lignées : BH1-Ikaros-V1/Ikaros-V2 (50ng/ml), BH-Aiolos-V1/Aiolos-V2 (35ng/ml) et BH1-Ikaros-V1/Aiolos-V2 (40ng/ml). Puis, pour étudier l'effet le plus précoce sur le statut de la chromatine, j'ai limité le traitement avec la doxycycline à 16 heures.

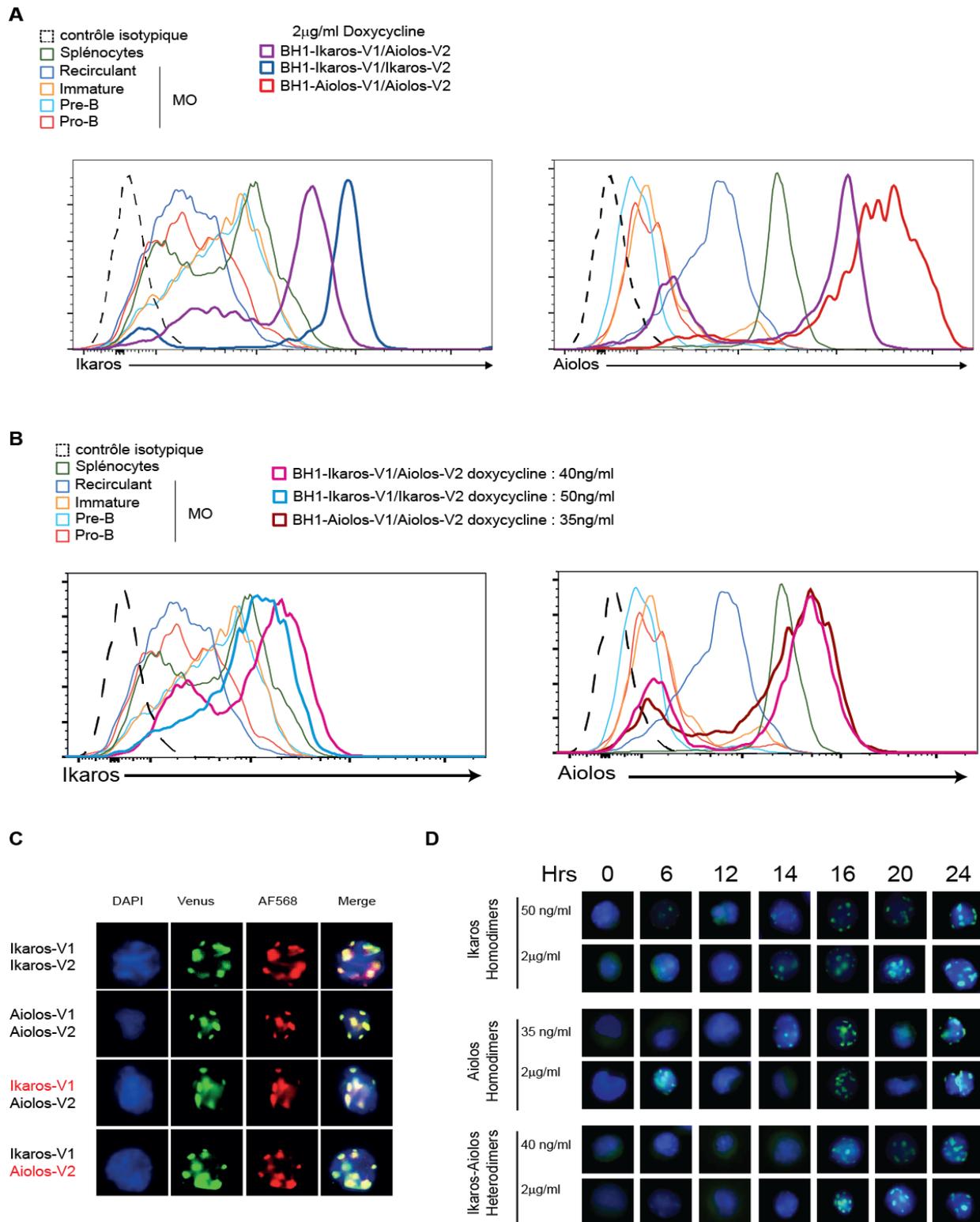


Figure RIII.1 Adaptation de la dose de doxycycline pour une expression plus physiologique (A) Comparaison des expressions d'Ikaros et d'Aiolos dans les cellules BH1-Ikaros-V1/Ikaros-V2 et BH1-Aiolos-V1/Aiolos-V2 et BH1-Ikaros-V1/Aiolos-V2 traitées avec la dose normale de doxycycline (2µg/ml) par rapport aux cellules primaires de la moelle osseuse et de la rate d'une souris WT et. (B) Comme A mais traitement avec des doses adaptées de doxycycline. (C) Immunofluorescence réalisée avec les cellules BH1 des lignées Ikaros-V1/Ikaros-V1, Aiolos-V1/Aiolos-V2 et Ikaros-V1/Aiolos-V2. Ikaros et Aiolos sont marqués avec des anticorps purifiés de lapin et l'anticorps secondaire est couplé à AF568. Pour la lignée Ikaros-V1/Aiolos-V2 l'écriture rouge indique la protéine marquée. (D) Les cellules BH1 des lignées Ikaros-V1/Ikaros-V2, Aiolos-V1/Aiolos-V2 et Ikaros-V1/Aiolos-V2 ont été observées à 0h, 6h, 12h, 14h, 16h, 20h et 24h après le début du traitement avec la doxycycline à 2µg/ml ou avec une dose adaptée : 50ng/ml, 35ng/ml et 40ng/ml respectivement.

B. L'expression d'Ikaros et d'Aiolos a un effet d'ouverture sur la chromatine

Les facteurs de transcription de la famille d'Ikaros sont capable de réguler l'ouverture et la fermeture de la chromatine à proximité de leurs gènes cibles en recrutant différents complexes de remodelage de la chromatine (NuRD, SWI/Snf, PRC2). Pour étudier l'effet de l'expression des facteurs de transcription Ikaros et Aiolos indépendamment sur le statut de la chromatine j'ai réalisé des essais d'accessibilité de la chromatine à la transposase suivis de séquençage haut-débit (ATAC-seq : Assay for Transposase-Accessible Chromatine) après 16 heures de traitement à la doxycycline avec les doses adaptées de doxycycline dans les lignées BH1-Ikaros-V1/Ikaros-V2 et BH1-Aiolos-V1/Aiolos-V2. Les distributions des Log2FC entre les conditions Ikaros et contrôle et Aiolos et contrôle sont centrés sur 0 ce qui signifie que les expressions d'Ikaros et d'Aiolos ne changent pas l'état général de la chromatine (Figure RIII.2.A).

J'ai ensuite voulu savoir quel est l'impact de la liaison des homodimères sur la chromatine. Pour cela j'ai comparé les pics d'ATAC-seq avec les pics de CHIP-seq détecté avec l'anticorps α Venus. Les régions d'ATAC-seq qui ne sont pas corrélées avec la fixation de l'homodimère d'Ikaros sont plutôt des régions qui ont une tendance à la fermeture (Figure RII.B). L'homodimère d'Ikaros a peu de pics de CHIP-seq mais ceux-ci sont, pour la majorité, corrélés avec une ouverture de la chromatine (Figure RIII.2.C). Les gènes à proximité des pics de CHIP-seq corrélés avec la fermeture de la chromatine font partie de la signalisation des MAPK et de la signalisation par les cytokines du système immunitaire. Il pourrait s'agir alors de la voie de signalisation de l'IL7 qui doit être réprimée pour permettre la différenciation des lymphocytes pré-B (Figure RIII.2.D). Les gènes ciblés par l'homodimère d'Ikaros qui se trouvent dans une région ouverte de la chromatine sont impliqués dans les voies moléculaires de la régulation négative de la prolifération, ce qui est nécessaire à la différenciation, mais aussi de la régulation de la production de l'IL6. Cependant les lymphocytes B exprimant l'IL6 sont impliqués dans la formation précoce de centres germinatifs dans des pathologies auto-immunes telles que le lupus systémique érythémateux (SLE) (Arkatkar et al., 2017).

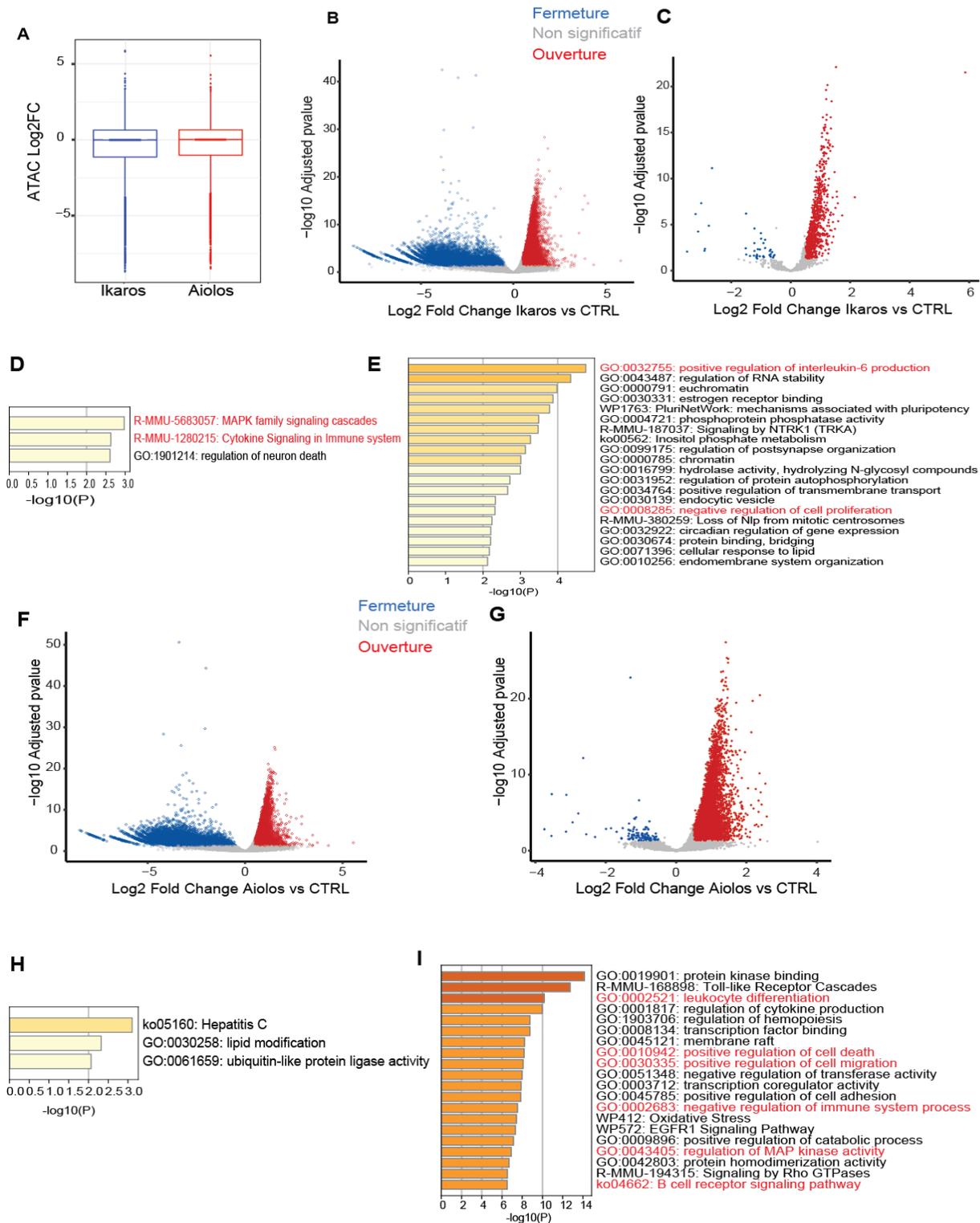


Figure RIII.2 La fixation des facteurs de transcription Ikaros et Aiolos favorise l'ouverture de la chromatine.

(A) Boxplots représentant la répartition des Log2FC des pics identifiés par ATAC-seq dans les conditions Ikaros et Aiolos par rapport au contrôle. (B) Volcanoplot représentant la significativité en fonction des Log2FC des pics identifiés par ATAC-seq dans la condition Ikaros qui ne sont pas corrélés avec des pics de ChIP-seq de Ikaros-homodimère. (C) Comme (B) mais pour des pics d'ATAC-seq corrélés avec les pics de ChIP-seq. (D) Voies moléculaires enrichies dans la liste des gènes associés à une fixation de Ikaros-homodimère et une fermeture de la chromatine. (E) Comme (D) pour des gènes associés à une ouvertures de la chromatine. (F) et (G) Comme (B) et (C) pour Aiolos-homodimère. (H) et (I) Comme (D) et (E) pour Aiolos-homodimère. **Fermeture** représente des pics d'ATAC-seq avec pvalue ajustée $< 0,05$ et le Log2FC ≤ 1 , **Non significatif** représente les pics d'ATAC-seq avec pvalue ajustée $> 0,05$ et **Ouverture** représente les pics d'ATA-seq avec pvalue ajustée $< 0,05$ et Log2FC ≥ 1

De la même façon que pour les homodimères d'Ikaros, les pics d'ATAC-seq qui ne sont pas corrélés avec des pics de CHIP-seq de l'homodimère d'Aiolos tendent vers une fermeture de la chromatine (Figure RIII.2.E). La liaison de l'homodimère d'Aiolos, comme la liaison de l'homodimère d'Ikaros, est corrélée avec une ouverture de la chromatine (Figure RIII.2.F). Les gènes cibles de l'homodimère d'Aiolos qui se trouvent dans une région plus fermée de la chromatine, par rapport à la condition contrôle, ne font pas partie de voies moléculaires spécifiques des lymphocytes B (Figure RIII.2.H). En revanche les gènes cibles qui se trouvent dans une région plus ouverte de la chromatine font partie des voies moléculaires impliquées dans la différenciation leucocytaire, la régulation du programme de mort cellulaire, la migration cellulaire, la régulation de l'activité des MAPK ou encore la voie de signalisation du BCR.

Ces résultats montrent que la fixation des facteurs de transcription de la famille d'Ikaros est plutôt corrélée avec une ouverture de la chromatine dans les lymphocytes pré-B murins. D'un côté l'homodimère d'Ikaros semble réprimer les voies de signalisation induites par des cytokines, dont l'IL7, et la prolifération cellulaire et favoriser la production de l'IL6 qui est une cytokine pro-inflammatoire. D'un autre côté l'homodimère d'Aiolos semble davantage impliqué dans la régulation de la différenciation des lymphocytes B en régulant différents aspects tels que la migration, la mort cellulaire ou la voie de signalisation du BCR.

C. Aiolos régule l'expression des gènes spécifiques des lymphocytes B

Pour analyser le rôle d'Ikaros et d'Aiolos dans la régulation de la transcription de leurs gènes cibles, j'ai réalisé une extraction des ARN totaux suivie d'un séquençage haut-débit (RNA-seq) après 24 heures de traitement avec les doses adaptées de doxycycline sur la lignée BH1 rtTA et sur trois lignées issues de transductions indépendantes BH1-Ikaros-V1/Ikaros-V2 et BH1-Aiolos-V1/Aiolos-V2. L'analyse en composantes principales montre que les réplicats d'une même lignée sont assez proches et que les lignées BH1-Ikaros-V1/Ikaros-V2 et les lignées BH1-Aiolos-V1/Aiolos-V2 sont différentes de la lignée contrôle et entre-elles (Figure RIII.3.A).

L'expression d'Ikaros ou d'Aiolos n'a donc pas le même impact sur la transcription. En effet, il y a plus de gènes dont l'expression est régulée, réprimée ou augmentée, dans la condition Aiolos par rapport au contrôle que dans la condition Ikaros par rapport au contrôle (5885/1862) (Figure RIII.4. B et C). De plus, parmi les gènes dont l'expression est augmentée, dans la condition Ikaros par rapport au contrôle, seulement 11% (96/807) le sont uniquement dans la condition Ikaros. De la même façon, parmi gènes dont l'expression est réprimée 34,5% (364/1055) sont réprimés uniquement dans la condition Ikaros. Les autres gènes sont également régulés par Aiolos et de façon plus forte. D'un autre côté, près de 53% (1591/3043) des gènes dont l'expression est augmentée dans la condition Aiolos le sont uniquement dans cette condition et 71% (2029/2842) sont uniquement réprimés dans la condition Aiolos. Ainsi, Aiolos semble être un régulateur plus fort qu'Ikaros ce qui est concordant avec les résultats présentés par Morgan et collaborateurs en 1997. En effet, les auteurs ont co-transduit des fibroblastes avec le rapporteur tkCAT sous le contrôle d'une séquence comportant quatre copies du site de liaison de forte affinité d'Ikaros : GGGAAA. La mesure de l'activité de l'enzyme chloroamphenicol acetyltransferase (CAT) a permis de démontrer qu'Aiolos stimule deux fois plus l'activité de CAT qu'Ikaros (Morgan et al., 1997).

Je me suis ensuite intéressée à la régulation directe des gènes spécifiques des lymphocytes B. Pour cela, j'ai associé la liaison à l'ADN des homodimères à

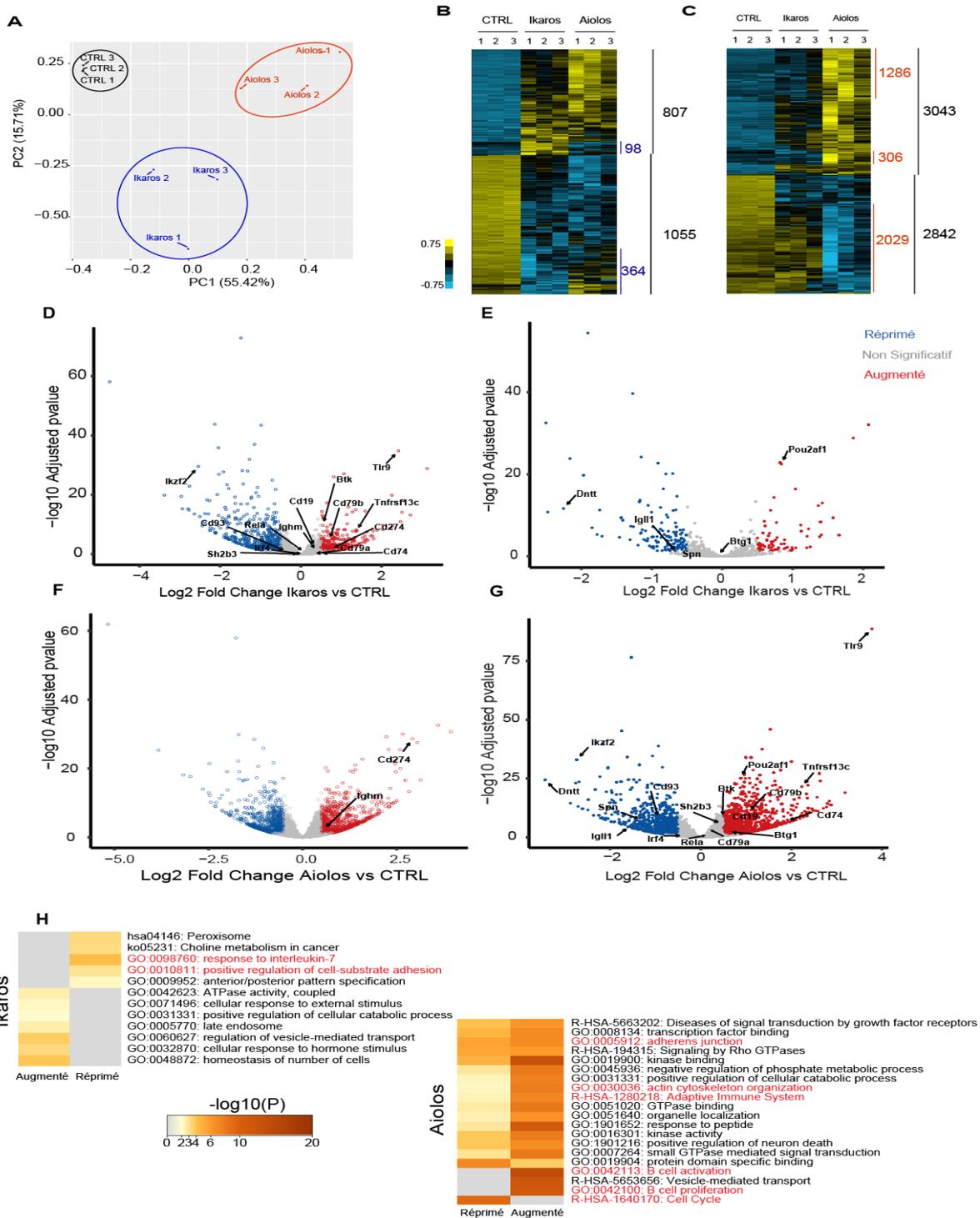


Figure RIII.3 Les gènes spécifiques des lymphocytes B dont directement régulés par Aiolos (A) Analyse des composantes principale pour les trois réplicats de chaque lignée BH1-rtTA (CTRL), BH1-IkarosV1/Ikaros-V2 et BH1-Aiolos-1/Aiolos-V2 traitées pendant 24 heures avec les doses adaptées de doxycycline. 5B) et (C) Heatmap représentant les valeurs de RNA-seq avec pvalue ajustée <0,05 dans les comparaisons Ikaros vs Contrôle et Aiolos vs Contrôle respectivement. (D) Volcanoplot représentant la significativité en fonction des Log2FC des valeurs du RNA-seq dans la condition Ikaros vs Contrôle qui ne sont pas corrélés avec des pics de ChIP-seq de Ikaros-homodimère. (E) Comme (D) pour des valeurs de RNA-seq corrélées avec des pics de ChIP-seq. (F) et (G) Comme (D) et (E) pour Aiolos-homodimère. (H) Comparaison entre les voies moléculaires enrichies dans les listes de gènes dont l'expression est réprimée ou augmentée dans la condition Ikaros vs Contrôle et corrélées avec des pics de ChIP-seq. (I) Comme (H) pour Aiolos-homodimère. Réprimé représente des valeurs de RNA-seq avec pvalue ajustée < 0,05 et le Log2FC ≤ 1 , Non significatif représente les valeurs de RNA-seq avec pvalue ajustée > 0,05 et Augmenté représente les valeurs de RNA-seq avec pvalue ajustée < 0,05 et Log2FC ≥ 1

l'expression des gènes à proximité des pics de ChIP-seq (Figure RIII.3 D, E, F et G). Les gènes spécifiques des lymphocytes B sont bien régulés de la même façon dans les conditions Ikaros et Aiolos, cependant ils ne semblent pas être des cibles directes de l'homodimère d'Ikaros mais plutôt de l'homodimère d'Aiolos (Figure RIII.3 D et G). Comme nous l'avons vu précédemment ces gènes sont des cibles directes de l'homodimère d'Aiolos et de l'hétérodimère Ikaros/Aiolos. Il serait alors possible que l'expression de ces gènes dans la condition Ikaros soit régulée par l'hétérodimère Ikaros/Aiolos formé avec la protéine Aiolos endogène.

Enfin, j'ai comparé les voies moléculaires enrichies entre les cibles directes réprimées et augmentées par l'homodimère d'Ikaros et celles réprimées et augmentées par l'homodimère d'Aiolos (Figure RIII.3.H). L'homodimère d'Ikaros réprime les gènes impliqués dans la voie de signalisation de l'IL7 ce qui est en accord avec les résultats d'ATAC-seq montrant une fermeture de la chromatine à proximité des gènes impliqués dans les voies de signalisation par les cytokines. Ceci est également concordant avec des résultats précédemment obtenus dans le laboratoire montrant une compétition de liaison à l'ADN entre Ikaros et STAT5 au cours de la différenciation des lymphocytes pré-B (Heizmann et al., 2020).

La seconde voie importante réprimée par l'homodimère d'Ikaros est l'adhésion cellulaire au substrat. En effet, comme expliqué dans l'introduction, pour la poursuite de leur différenciation, les lymphocytes pré-B quittent une niche hématopoïétique riche vers une niche pauvre en IL7, grâce notamment à l'expression de CXCR4 et la répression des intégrines $\alpha4\beta1$. J'ai pu effectivement constater que l'expression de *Cxcr4* est augmentée uniquement dans la condition Ikaros mais ne semble pas en être une cible directe, l'homodimère ne se fixe pas sur les régions régulatrices du gène. Le gène codant pour l'intégrine $\beta1$, *Itgb1*, est ciblé par l'homodimère d'Ikaros et son expression est réprimée uniquement dans la condition Ikaros. D'un autre côté, l'homodimère d'Aiolos semble réguler directement les voies moléculaires de l'activation et de la prolifération des lymphocytes B mais aussi de la jonction cellulaire et l'organisation du cytosquelette. Ceci est en accord avec les résultats de l'ATAC-seq montrant une ouverture à proximité des gènes impliqués dans la différenciation leucocytaire et la voie de signalisation du BCR mais aussi dans la migration cellulaire.

Ensemble, ces résultats montrent que l'homodimère d'Aiolos est tout à fait capable de réguler l'expression des gènes impliqués dans la différenciation des lymphocytes B mais également que l'homodimère d'Ikaros semble avoir une

fonctionnalité plus restreinte. Les lymphocytes pré-B déficients pour Ikaros et bloqués dans leur différenciation expriment de faibles niveau d'Aiolos. A ce niveau d'expression, Aiolos n'est donc pas capable de soutenir la différenciation des lymphocytes pré-B, il est, cependant, envisageable qu'à des niveaux d'expression plus élevés, lui permettant de former des homodimères, ou sous la forme d'hétérodimère avec Ikaros, il soit en mesure de réenclencher le programme de différenciation des lymphocytes pré-B déficients pour Ikaros.

D. Les cellules exprimant Aiolos se différencient plus vite

La réexpression d'Ikaros dans les cellules BH1 permet de lever le blocage de différenciation et d'obtenir en 5 jours des cellules exprimant le BCR *in vitro*. La suppression de l'IL7 dans le milieu de culture agit en synergie avec l'expression d'Ikaros. J'ai alors exposé les lignées cellulaires BH1-Ikaros-1/Ikaros-V2, BH1-Aiolos-V1/Aiolos-V2 et BH1-IkarosV1/Aiolos-V2 à la dose normale de doxycycline (+++) et aux doses adaptées de doxycycline (+). Pour suivre le processus de différenciation dans le temps, j'ai analysé l'expression intracellulaire de $\lambda 5$ et l'expression membranaire du BCR (Ig κ et Ig μ) après trois jours et après cinq jours de doxycycline (Figure RIII.4).

Après trois jours toutes les lignées cellulaires présentent une diminution de l'expression intracellulaire de la chaîne légère de substitution $\lambda 5$ et cette diminution est encore plus prononcée après cinq jours. De même, les cellules qui répriment l'expression de $\lambda 5$ sont plus petites. La baisse de ces deux paramètres est le premier signe de la reprise de la différenciation des cellules BH1. Enfin, les cellules traitées avec la doxycycline présentent toutes une population co-exprimant la chaîne lourde μ des Ig et la chaîne légère κ à la surface. Cette co-expression définit la présence du BCR. Le retrait de l'IL7 pendant la différenciation permet bien d'augmenter le pourcentage de cellules différenciées après trois et cinq jours. Ce résultat confirme ceux précédemment obtenus dans l'équipe (Heizmann et al., 2013). Ce qui est nouveau ici est que les cellules exprimant Aiolos ont une cinétique de différenciation accélérée par rapport aux cellules qui n'expriment qu'Ikaros. En effet, après trois jours de différenciation, plus de 95% des cellules de la lignée BH1-Aiolos-V1/Aiolos-V2 expriment le BCR à leur surface même avec la dose faible de doxycycline. La lignée BH1-Ikaros-V1/Aiolos-V2 présentent au minimum 60% de cellules exprimant le BCR dans la condition IL7+/Dox+ et jusqu'à 93% dans la condition de différenciation la plus favorable IL7-/Dox+++. D'un autre côté, les cellules de la lignée BH1-IkarosV1/Ikaros-V2 se différencient moins vite et en plus faible proportion. En effet après trois jours, dans la condition IL7-/Dox+++, seulement 24% des cellules expriment le BCR et 63% après cinq jours. La lignée exprimant les deux facteurs a alors une cinétique intermédiaire entre les deux autres lignées.

JOUR 3

JOUR 5

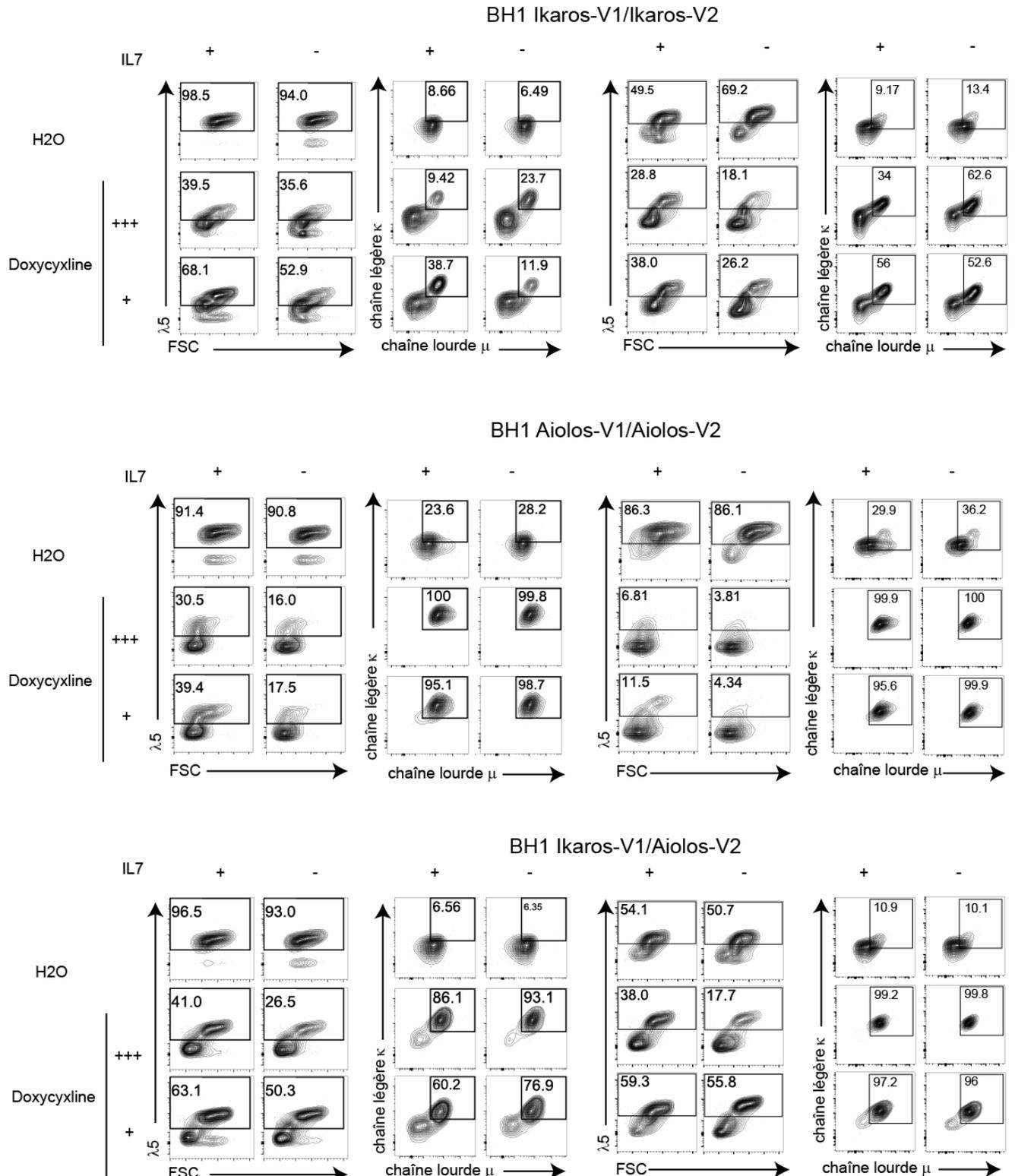


Figure RIII.4 L'expression d'Aiolos permet aux cellules de se différencier plus rapidement

Les cellules sont traitées pendant 5 jours, en présence ou en absence d'IL7, avec deux doses de doxycycline (+++ = 2μg/ml ; + = 50ng/ml pour la lignée BH1-Ikaros-V1/Ikaros-V2, 35ng/ml pour la lignée BH1-Aiolos-V1/Aiolos-V2 et 40ng/ml pour la lignée BH1-Ikaros-V1/Aiolos-V2). L'expression intracellulaire de λ5 dans toutes les cellules et l'expression membranaire de Igκ et Igμ, sur les cellules Venus+ pour les cellules traitées à la doxycycline et sur les cellules Venus- pour les contrôles, sont analysées après trois et cinq jours.

Ce résultat montre que la fonction d'Ikaros dans la différenciation des lymphocytes pré-B n'est pas une fonction unique mais qu'Aiolos est également capable de sauver le phénotype de blocage de la différenciation dû à l'absence d'Ikaros. Cependant, l'absence d'Ikaros mène bien à un blocage de la différenciation alors qu'Aiolos est exprimé dans les cellules. Nous avons vu dans la deuxième partie des résultats que cette protéine Aiolos endogène est faiblement exprimée et ne se fixe pas aussi efficacement à l'ADN que lorsqu'elle est présente à de plus fortes concentrations et particulièrement au niveau des régions régulatrices des gènes spécifiques des lymphocytes B. Il est possible que cette protéine faiblement exprimée ne parviennent pas à former efficacement des homodimères et se fixe à l'ADN sous forme de monomère. Mais, nous avons vu dans la première partie des résultats que la protéine Ikaros-Venus surexprimée est capable d'interagir avec la protéine Aiolos endogène pour former des hétérodimères. Dans la lignée cellulaires BH1-Ikaros-V1/Aiolos-V2, les deux facteurs Ikaros et Aiolos sont surexprimés favorisant leur hétérodimérisation. Dans cette lignée, le processus de différenciation est plus rapide et plus efficace que dans la lignée ne surexprimant qu'Ikaros.

Ces résultats suggèrent que la protéine Ikaros est en mesure de former des hétérodimères avec Aiolos dès que celui-ci commence à être exprimé et que ces hétérodimères pourraient être les réelles entités régulatrices de la différenciation des lymphocytes pré-B.

E. La co-expression d'Ikaros et d'Aiolos est plus efficace dans la régulation de l'expression des gènes spécifiques des lymphocytes B

Au regard des résultats précédents, j'ai voulu savoir dans quelle mesure les gènes spécifiques des lymphocytes B étaient régulés dans les trois lignées cellulaires en fonction de la dose de doxycycline et donc de la concentration des facteurs exprimés dans les cellules.

Pour cette expérience, j'ai traité les cellules des trois lignées BH1-Ikaros-V1/Ikaros-V2, BH1-Aiolos-V1/Aiolos-V2 et BH1-Ikaros-V1/Aiolos-V2 avec des concentrations croissantes de doxycycline : 35, 50, 100, 500, 1000 et 2000 ng/ml pendant 24 heures. J'ai ensuite extrait les ARN totaux et réalisé des RT-qPCR pour suivre l'expression des transcrits de différents gènes spécifiques des lymphocytes B qui sont différemment régulés, en RNA-seq, entre les conditions Ikaros et Aiolos (Figure RIII.5). J'ai, dans un premier temps, vérifié que les transcrits d'*Ikzf1* et *Ikzf3* étaient bien augmentés en fonction de la concentration de doxycycline et également si les niveaux sont comparables entre les lignées des homodimères et la lignée de l'hétérodimère. La lignée BH1-Ikaros-V1/Aiolos-V2 semble exprimer de plus faibles niveaux d'*Ikzf1* que la lignée BH1-Ikaros-V1/Ikaros-V2 mais des niveaux équivalents d'*Ikzf3* à ceux de la lignées BH1-Aiolos-V1/Aiolos-V2.

Les résultats obtenus dans cette expérience sont tout à fait concordants avec ceux obtenus dans l'analyse du RNA-seq. Nous pouvons voir, ici, qu'avec la concentration la plus élevée de doxycycline, les niveaux de transcrits de *Tnfrsf13c* et *Tlr9* dans la lignée BH1-Ikaros-V1/Ikaros-V2 n'atteignent pas ceux de la lignée BH1-Aiolos-V1/Aiolos-V2 avec les concentrations les plus faibles de doxycycline. Cela confirme le fait que ces gènes sont des cibles privilégiées d'Aiolos. Ce qui est surprenant dans cette expérience est la régulation de ces transcrits dans la lignée BH1-Ikaros-V1/Aiolos-V2. Lorsque la régulation dans la condition Ikaros et Aiolos a la même tendance, augmentation pour *Tnfrsf13c* et *Tlr9* ou répression pour *Igll1*, alors l'augmentation ou la répression est plus forte dans la condition Ikaros/Aiolos. Lorsque la tendance est inversée entre Ikaros et Aiolos par exemple pour *Slamf7* dont la

Ikaros-V1/Ikaros-V2 Aiolos-V1/Aiolos-V2 Ikaros-V1/Aiolos-V2

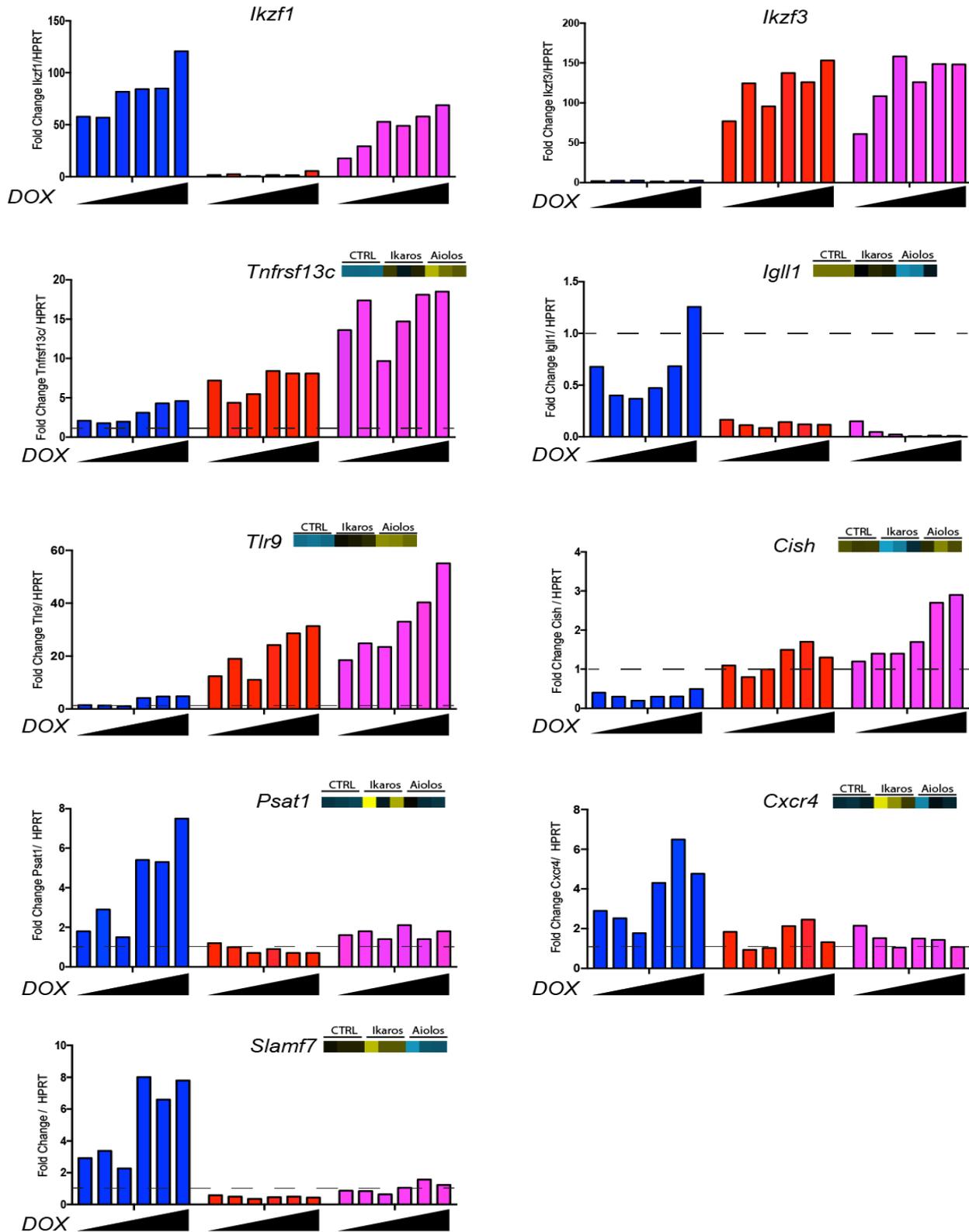


Figure RIII.5 La co-expression d'Ikaros et d'Aiolos permet une régulation plus forte de la transcription

Représentation des $\Delta\Delta Ct$ /HPRT issus des PCR quantitatives après retrotranscription des ARN extraits des lignées cellulaires BH1-Ikaros-V1/Ikaros-V2 (Bleu), BH1-Aiolos-V1/Aiolos-V2 (rouge) et BH1-Ikaros-V1/Aiolos-V2 (rose) traitées avec des doses croissantes de doxycycline : 0, 35, 50, 100, 500, 1000 et 2000 ng/ml. Les données ont été normalisées sur la condition contrôle 0ng/doxycycline, la ligne pointillée représente donc le niveau d'expression dans la condition contrôle. Pour chaque gène testé, la représentation du heatmap issu du RNA-seq est adjoint.

transcription est favorisée dans la condition Ikaros et réprimée dans la condition Aiolos, alors le transcrit reste stable dans la condition Ikaros/Aiolos. Lorsque la transcription reste stable dans la condition d'Aiolos et est réprimée (Cish) ou augmentée (Psat1, Cxcr4) dans la condition Ikaros, alors la transcription aura tendance à rester stable dans la condition Ikaros/Aiolos. Ces résultats suggèrent qu'Aiolos a un effet dominant et que les deux facteurs peuvent agir en synergie lorsqu'ils régulent la transcription de la même façon.

L'ensemble de ces résultats suggèrent qu'Aiolos a un rôle prépondérant dans la différenciation des lymphocytes pré-B. Il permettrait à Ikaros de cibler les gènes impliqués dans les différentes voies moléculaires de la différenciation en formant des hétérodimères. Bien que ces deux facteurs semblent pouvoir agir en synergie, la fonctionnalité d'Aiolos paraît dominante. Finalement, l'hétérodimère Ikaros/Aiolos apparaît comme pouvant être le réel régulateur de la différenciation des lymphocytes pré-B.

Discussion

Dans ce travail de thèse nous nous sommes intéressés aux spécificités de liaison à l'ADN et à la fonctionnalité des dimères des facteurs de transcription de la famille d'Ikaros au cours du développement des lymphocytes B. Grâce à la complémentation bi-moléculaire de fluorescence importée dans les cellules BH1, nous avons pu étudier, dans un même système expérimental, les homodimères et hétérodimères formés par Ikaros et Aiolos. L'anticorps α Venus nous a, quant à lui, permis d'analyser les propriétés de liaison à l'ADN des différents dimères sans être contraints par les différences d'efficacité des anticorps spécifiques des protéines. Nous avons pu mettre en évidence un large répertoire de gènes préférentiellement ciblés par l'homodimère d'Aiolos et l'hétérodimère Ikaros/Aiolos. En effet, il apparaît qu'Aiolos peut directement réguler l'expression des gènes spécifiques des lymphocytes B, lorsqu'il est exprimé à de plus forts niveaux que dans les lymphocytes pré-B. Il est également capable de réenclencher le programme de différenciation des lymphocytes déficients pour Ikaros et bloqués au stade pré-B, une fonction qui apparaissait comme uniquement attribuée à Ikaros. Finalement, mes données suggèrent que l'hétérodimère Ikaros/Aiolos pourrait jouer un rôle prépondérant dans cette différenciation.

I. Les propriétés de liaison à l'ADN des facteurs de transcription de la famille d'Ikaros

Dans les résultats présentés ici, il semble qu'Ikaros et Helios partagent une grande similarité dans leurs propriétés de liaison à l'ADN alors qu'Aiolos se démarque en présentant un répertoire plus vaste. Cela a aussi été mis en évidence par l'étude de l'homodimère d'Aiolos qui se fixe à l'ADN de façon plus efficace que l'homodimère d'Ikaros. Nous avons pu aussi montrer que les propriétés de liaison de l'hétérodimère Ikaros/Aiolos sont plus proches de celles de l'homodimère d'Aiolos que de celui d'Ikaros. Nous avons donc tenté de comprendre d'où provient cette spécificité de liaison à l'ADN d'Aiolos.

Nous avons répertorié les motifs présents aux régions spécifiquement liées par Aiolos ou conjointement liées par Ikaros et Aiolos et avons comparé la distribution des bases à proximité immédiate du motif central GGAA. Aiolos semble effectuer une sélection négative dans les bases composant son motif de liaison pour assurer sa spécificité. Ainsi, Aiolos est capable de lier des motifs ne possédant pas de T en position +1 du GGAA alors qu'Ikaros ne l'est pas. Les motifs ainsi composés sont des motifs de faible affinité pour Aiolos ; en effet le remplacement du nucléotide par un T accentue sa capacité de liaison. La régulation de l'expression génique passe par l'utilisation, par les facteurs de transcription, de sites de liaison de haute ou de faible affinité (Wang et al., 2015). Comme expliqué par Crocker et al., les sites de liaison de faible affinité sont nécessaires à la régulation de l'expression génique (Crocker et al., 2016). En effet, ils sont essentiels à une régulation fine et précise de l'expression génique au cours du développement, permettant une activation de l'expression des gènes cibles modulée par la forte ou faible présence des facteurs de transcription spécifiques. C'est ce que j'observe pour Aiolos. Les faibles niveaux d'expression d'Aiolos dans les cellules BH1 ne permettent pas sa liaison aux sites spécifiques mis en évidence par la surexpression. Les motifs de liaison analysés pour la protéine Aiolos endogène révèlent une forte utilisation du T en +1 et donc des sites de liaison de haute affinité (Figure D.I.1). En se liant à ces motifs de faible affinité, Aiolos se présente comme un régulateur de la régulation génique plus important qu'Ikaros.

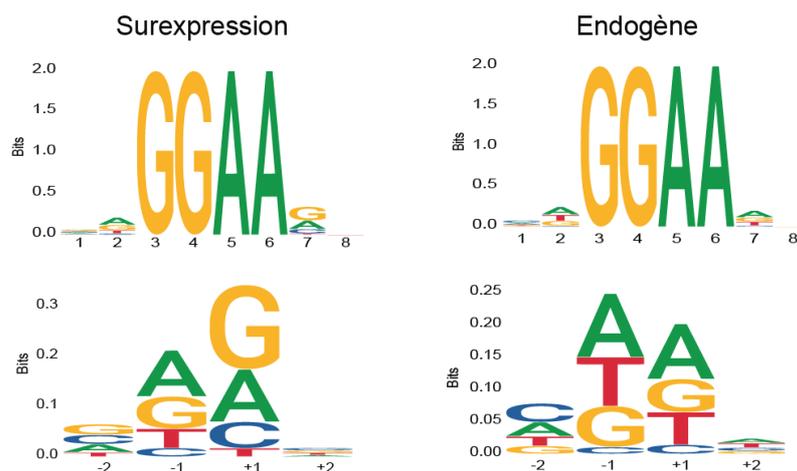


Figure D.I.1 Comparaison des motifs de liaison à l'ADN de la protéine Aiolos endogène ou surexprimée

Nous avons également relevé les combinaisons de nucléotides à proximité immédiate du GGAA. Les motifs présents dans les régions spécifiquement ciblées par

Aiolos présentent des combinaisons de ces quatre nucléotides aux positions -2, -1, +1 et +2 du GGAA très bien conservées et spécifiquement enrichies. En effet, des combinaisons différentes sont retrouvées dans les motifs des régions communément liées par Ikaros et Aiolos. Malgré leur conservation importante, ces combinaisons de nucléotides restent assez variables entre-elles ce qui explique l'absence de motif fiable pour les facteurs de transcription de la famille d'Ikaros dans les bases de données. Les domaines de liaison à l'ADN (DBD) des facteurs de transcription de la famille d'Ikaros sont composés de quatre doigts de zinc (DZ). Dans une étude de 2019, les auteurs ont utilisé un modèle de prédiction de la spécificité de liaison à l'ADN de chaque doigt de zinc qui compose le DBD d'un facteur de transcription (Dogan et al., 2019). Dans les facteurs étudiés, il y avait Aiolos. Les auteurs ont conclu qu'Aiolos n'utilise que les deux DZ centraux pour une liaison à un motif GGAA. Les DZ1 et 4 n'auraient alors pas de rôle dans la spécificité de liaison à l'ADN. Cependant, il a été montré que WT1 (Wilms tumor 1) qui possède aussi quatre DZ dans son domaine de liaison a besoin de trois de ses DZ (DZ 2, 3 et 4) pour établir une liaison à l'ADN spécifique. Le DZ1 n'influence pas l'affinité de liaison mais plutôt la cinétique d'association/dissociation (Nurmemmedov et al., 2009). Il a été montré que les DZ1 et DZ4 d'Ikaros sont impliqués dans la régulation de différents groupes de gènes et donc dans différentes fonctions d'Ikaros (Schjerven et al., 2013). Le DZ1 permet à Ikaros d'assumer son rôle dans la différenciation des lymphocytes B et le DZ4 son rôle de suppresseur de tumeur. Il serait alors possible que suivant les gènes cibles et donc les motifs de liaison à l'ADN, les facteurs de transcription de la famille d'Ikaros utilise alternativement leur DZ1 et DZ4 pour assurer leur liaison à l'ADN. Cela rend alors difficile l'établissement d'un motif particulier permettant une prédiction de liaison pour chacun des facteurs.

Après avoir analysé l'apport de la séquence nucléotidique pour la spécificité de liaison d'Aiolos à l'ADN, nous avons voulu savoir quel était le rôle du domaine de liaison à l'ADN (DBD) d'Aiolos dans la définition de cette spécificité. Ainsi, nous avons échangé les DBD entre Ikaros et Aiolos et nous avons confronté ces chimères à des motifs de liaison spécifiques d'Aiolos. Le DBD d'Aiolos ne confère pas à Ikaros une totale capacité de liaison sur ces motifs spécifiques. Le contexte protéique semble alors jouer un rôle dans la liaison à l'ADN. Dans une récente étude, les auteurs ont montré que, pour certains facteurs de transcription de la levure, d'autres domaines

que le DBD peuvent jouer un rôle dans la spécificité de liaison à l'ADN. En effet, les domaines intrinsèquement désordonnés (IDR : Intrinsically Disordered Regions) des facteurs de transcription semblent avoir un rôle important dans la reconnaissance du motif de liaison à l'ADN (Brodsky et al., 2020). Les IDRs sont composés d'acides aminés polaires ou chargés empêchant l'établissement d'une structure tridimensionnelle stable. Cela permet l'adaptation du facteur de transcription à la structure de son ADN cible. Il est possible de visualiser les IDRs des protéines grâce à Iupred2a (<https://iupred2a.elte.hu/>). Les protéines de la famille d'Ikaros possèdent un domaine N-terminal désordonné. Si nous confrontons l'alignement des protéines Ikaros, Helios et Aiolos à la prédiction de structures désordonnées, nous observons une concordance entre les domaines désordonnés et les domaines non conservés entre les trois protéines (Figure D.I.2). L'étude de ce domaine N-terminal pourrait alors aider à la compréhension des spécificités de liaison à l'ADN des facteurs de transcription de la famille d'Ikaros.

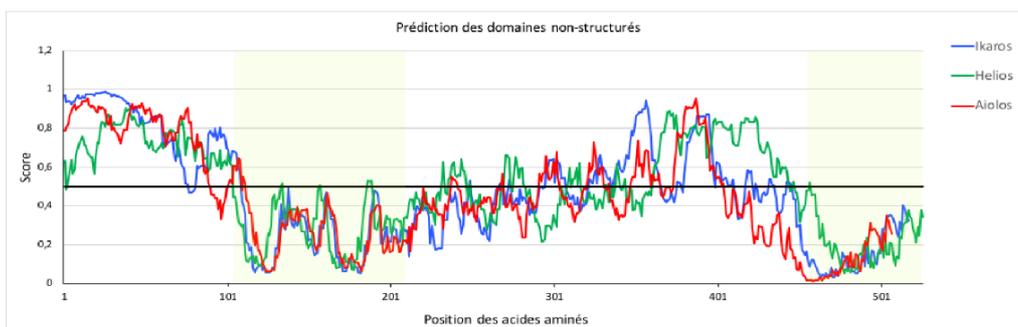
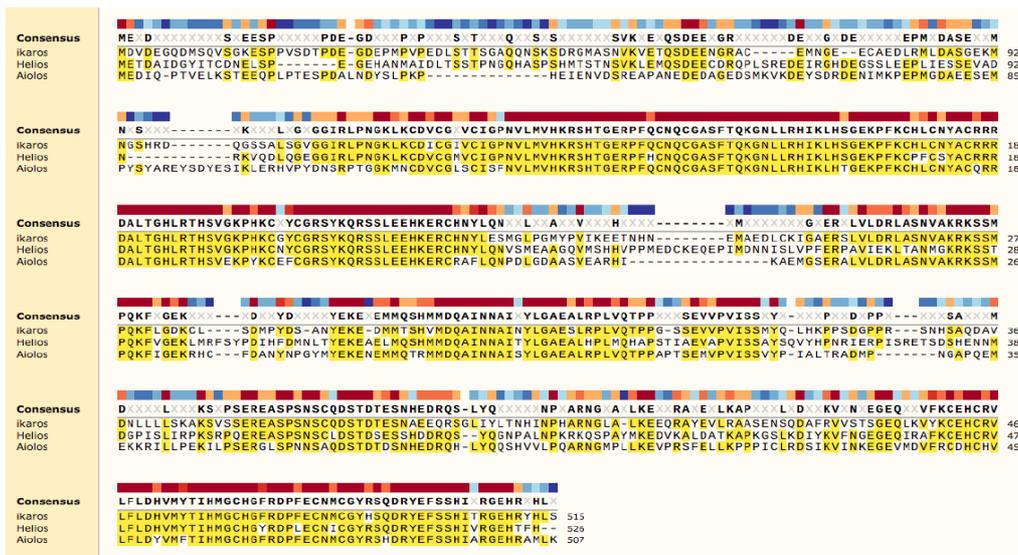


Figure D.I.2 Alignement d'Ikaros, Helios et Aiolos et analyse des domaines désordonnés. Un domaine est considéré comme désordonné lorsque son score est supérieur à 0.5. Les domaines en vert sont les domaines de liaison à l'ADN et de dimérisation.

La modulation des spécificités de liaison à l'ADN au sein d'une même famille de facteurs de transcription dont les domaines de liaison à l'ADN sont très conservés entre eux peut également provenir de variations dans la séquence d'acides aminés. Par exemple dans la famille RUNX, bien que les résidus 123 et 168 ne soient pas en contact direct avec l'ADN, la présence d'une valine à ces sites permet à Runx3 d'établir une liaison de plus forte affinité avec CBF β permettant une liaison de haute affinité avec l'ADN. Cela permet à Runx3 et non à Runx1 de réguler l'activation des lymphocytes T régulateurs (Treg) (Bruno et al., 2019). Grâce à l'anticorps α Venus, nous pourrions étudier les partenaires protéiques des différents dimères des facteurs de transcription de la famille d'Ikaros par immunoprécipitation suivie de spectrométrie de masse.

II. La fonctionnalité des facteurs de transcription de la famille d'Ikaros

Durant ma thèse, je me suis également intéressée à la fonctionnalité des facteurs de transcription de la famille d'Ikaros au cours du développement B. Pour cela j'ai regardé deux aspects : la régulation de la transcription de leurs gènes cibles et leur faculté à soutenir la différenciation des lymphocytes pré-B en culture. Ainsi j'ai montré, *in vitro*, que l'expression d'Aiolos permet une différenciation plus rapide des cellules BH1 que l'expression d'Ikaros. De plus, l'homodimère d'Aiolos régule directement les gènes impliqués dans la régulation de la différenciation, de l'apoptose et de la migration cellulaire tandis que l'homodimère d'Ikaros ne semble pouvoir réguler que la voie de signalisation de l'IL7 et l'adhésion cellulaire.

J'ai montré, par exemple qu'Aiolos cible de façon privilégiée les gènes spécifiques des lymphocytes B. En 2017, Nakayama et al. ont associé la surexpression d'Aiolos au développement d'une forme de lupus systémique érythémateux dépendant de la signalisation de BAFF (Nakayama et al., 2017). Les auteurs ont également montré que la modulation de l'activité du complexe CRL4^{CRBN} permet de réduire les niveaux d'Aiolos dans ces cellules et réprime leur différenciation

en plasmablastes. Cependant le mécanisme reliant l'expression d'Aiolos à la sensibilité des cellules pour BAFF n'avait pas été exploré. J'ai montré, dans ce travail, qu'Aiolos semble directement réguler la transcription de *Tnfrsf13c* codant pour le récepteur de BAFF (BAFF-R). Ainsi les cellules surexprimant Aiolos auront une tendance à exprimer plus de BAFF-R ce qui expliquerait leur sensibilité accrue pour cette voie de signalisation et l'inhibition de la différenciation en plasmablaste par la modulation l'expression d'Aiolos via l'activité du complexe CRL4^{CRBN}. En 2016, deux études ont montré l'importance de la régulation de *Sh2b3* codant pour la protéine LNK dans la régulation de la voie de signalisation de l'IL7 et dans le développement de leucémies aiguës lymphoblastiques (Cheng et al., 2016; Ge et al., 2016). En effet il semble que la protéine LNK puisse interagir directement avec JAK3 pour atténuer la voie de signalisation IL7/JAK/STAT. D'un autre côté, les faibles taux d'expression de *Sh2b3* ont été corrélés avec une surexpression du récepteur de l'IL7. Dans mon travail j'ai montré que *Sh2b3* est une cible privilégiée d'Aiolos et que son expression est augmentée seulement dans la condition Aiolos. J'ai également montré que l'expression d'Aiolos pousse les cellules BH1 dans la différenciation. Aiolos est également capable de réprimer le cycle cellulaire des cellules Nalm-6 qui sont des cellules issues d'une leucémie aiguë lymphoblastique (Zhuang et al., 2014). Le contrôle de l'expression d'Aiolos dans les B-ALL pourrait alors être un outil thérapeutique dans le traitement de cette pathologie.

Bien que j'aie montré qu'Aiolos semble avoir un rôle important dans la différenciation des lymphocytes pré-B, les souris déficientes pour Aiolos ne présentent pas de blocage de la différenciation des lymphocytes B (Wang et al., 1998). Cependant, les lymphocytes B de ces souris produisent plus d'auto-anticorps et développent des lymphomes ce qui dénote d'un défaut dans la tolérance centrale. J'ai montré dans mon travail qu'Aiolos régule le programme de mort cellulaire. Il est donc possible qu'Aiolos soit le régulateur de la tolérance centrale des lymphocytes B en régulant les voies de l'apoptose. Il y a donc une ambivalence dans le rôle d'Aiolos où d'un côté il régule l'expression de BAFF-R qui prodigue un signal de survie cellulaire, et de l'autre il régule le programme de l'apoptose régulant ainsi la tolérance centrale. Ceci pourrait expliquer que son expression soit croissante à partir du moment où elle est induite au stade pré-B. Ces résultats ont été obtenus dans un système de surexpression où le niveau d'Aiolos est comparable à celui des lymphocytes B spléniques. J'ai ainsi pu mettre en évidence certains rôles spécifiques mais en

comparaison à la physiologie, il est possible que la régulation de ses gènes cibles soit différente suivant le stade de différenciation.

Malgré la surexpression, je n'ai pas observé autant de spécificités pour l'homodimère d'Ikaros que pour celui d'Aiolos. Les rôles que j'ai mis en évidence dans mon travail avaient déjà été mentionnés par différents laboratoires dont le nôtre. L'homodimère d'Ikaros permet la répression de la voie de signalisation de l'IL7. Nous avons montré dans le laboratoire qu'Ikaros permet l'atténuation de la voie de signalisation de l'IL7 par compétition de la liaison à l'ADN avec STAT5 (Heizmann et al., 2013, 2020). La deuxième fonction principale de l'homodimère d'Ikaros, mise en évidence dans mon travail, est la régulation de l'adhésion cellulaire. J'ai effectivement pu noter dans l'analyse du RNA-seq que les gènes *Itga1*, *Itga5* et *Itgb1* codant pour les intégrines et les gènes codants pour d'autres composants de l'adhésion focale (*Actn1*, *Rock1*, *Rogdi*) étaient réprimés dans la condition Ikaros. Il est à noter que certains d'entre eux le sont aussi dans la condition Aiolos. La perte d'adhésion permet la migration des cellules d'une niche hématopoïétique riche en IL7 vers une niche pauvre en IL7 ce qui promeut la différenciation (Fistonich et al., 2018). Il a été montré que la perte de fonction d'Ikaros dans la lignée lymphoïde B entraîne une surexpression de ces différents gènes et ainsi une dérégulation de l'adhésion cellulaire menant à une sensibilisation des cellules au développement d'une leucémie aiguë lymphoblastique (Joshi et al., 2014).

III. L'hétérodimère Ikaros/Aiolos est-il la clé du mystère ?

Un aspect totalement nouveau de mon travail de thèse est la mise en évidence du rôle de l'hétérodimère Ikaros/Aiolos dans la différenciation des lymphocytes pré-B. En effet, j'ai pu montrer que d'une part l'hétérodimère a un répertoire de liaison à l'ADN plus vaste que celui de l'homodimère d'Ikaros et semblable à celui de l'homodimère d'Aiolos. Ce qui montre une dominance d'Aiolos dans le dimère. D'autre part, j'ai montré que la protéine Ikaros surexprimée se dimérise avec la protéine Aiolos endogène mettant en avant une affinité de dimérisation plus importante d'Ikaros avec Aiolos que d'Ikaros avec lui-même. Cependant, les propriétés de dimérisation des protéines restent à étudier. Idéalement, il faudrait avoir accès aux protéines purifiées pour utiliser une technique de Bio-Layer Interferometry (BLI) permettant de mesurer les constantes d'association et de dissociation de deux protéines. Au niveau de la fonctionnalité, il semble que la co-expression des deux facteurs Ikaros et Aiolos permette une différenciation des cellules BH1 plus rapide qu'avec Ikaros seulement et la régulation des gènes spécifiques des lymphocytes B semble montrer une synergie entre Ikaros et Aiolos. Ces résultats sont concordants avec ceux présentés par Morgan et al., dans lesquels ils montrent que la co-expression d'Ikaros et d'Aiolos augmente, de façon plus importante que Ikaros ou Aiolos seulement, l'activité de la CAT (Chloramphenicol AcetylTransferase), dont l'expression est contrôlée par un promoteur portant quatre répétitions du site de liaison de haute affinité pour Ikaros : GGGAAA (Morgan et al., 1997). Cependant, mes résultats ne prouvent pas l'implication directe de l'hétérodimère car les trois dimères sont présents dans les cellules BH1-Ikaros-V1/Aiolos-V2. De la même façon, la lignée BH1-Ikaros-V1/Ikaros-V2 exprime toujours la protéine Aiolos endogène et nous ne pouvons pas exclure que certains hétérodimères ne soient pas à l'origine des effets sur la différenciation ou la régulation de la transcription provoqués par la ré-expression d'Ikaros. Ainsi, pour étudier spécifiquement la fonctionnalité des homodimères, nous avons généré, dans le laboratoire, une lignée BH1 déficiente pour les trois facteurs de transcription Ikaros, Aiolos et Helios. Pour l'étude de l'hétérodimère nous pourrions utiliser un système de dimérisation inducible où nous remplacerions les domaines de dimérisation des

protéines par des domaines absolument complémentaires qui ne permettent la formation que des hétérodimères (Fernandez-Rodriguez and Marlovits, 2012).

Dans la littérature, certains faisceaux d'indices suggèrent que l'hétérodimère pourrait avoir un rôle important. Il a été montré qu'Ikaros est capable de recruter Aiolos au niveau des super-enhancers permettant ainsi la régulation de l'expression des gènes de la différenciation des cellules pré-B (Hu et al., 2016). Il a été également montré qu'Ikaros et Aiolos collaborent pour la répression de l'expression de *c-Myc*, permettant la régulation du cycle cellulaire, nécessaire à la poursuite de la différenciation des lymphocytes pré-B (Ma et al., 2010).

De plus, bien que la différenciation des lymphocytes B ne soit pas bloquée chez les souris déficientes pour Aiolos, montrant qu'Ikaros est capable à lui seul d'assumer cette fonction, les souris *Ikzf1^{+/-}Ikzf3^{-/-}*, présentent, elles, un blocage dans la différenciation au stade pré-B (Joshi et al., 2014; Wang et al., 1998). Cela suggère un rôle important d'Aiolos à ce stade la différenciation et donc potentiellement de l'hétérodimère. Les profils de liaison à l'ADN d'Helios et d'Ikaros sont très similaires. Cela pourrait laisser penser que ces deux facteurs forment des hétérodimères Ikaros/Helios. De plus, la surexpression d'Helios dans le lymphocytes B mène au développement d'un phénotype proche de celui des souris déficientes pour Aiolos (Dovat et al., 2005). La surexpression d'Helios empêcherait donc la formation de l'hétérodimère Ikaros/Aiolos et ainsi la régulation de la tolérance centrale des B. La répression d'Helios au cours de la différenciation des lymphocytes B prend alors tout son sens.

Dans une récente étude, les auteurs ont décrit des mutations hétérozygotes non-sens et faux-sens dans *IKZF1* entraînant la perte du domaine de dimérisation (R213*, S427*) ou seulement de sa fonction (C467R). Les patients se présentent avec des pathologies diverses telles que pneumonies, anémies, thrombocytopénie, hypogammaglobulinémie, T-ALL, production d'auto-anticorps et lymphome de Burkitt. Les protéines Ikaros produites se lient à l'ADN sous forme de monomère (Kuehn et al., 2021). Ainsi la perte de dimérisation d'Ikaros conduit au développement de lymphomes B et à la production d'auto-anticorps. Ces différentes hémopathies humaines sont retrouvées chez les souris déficientes pour Aiolos alors que dans ce cas précis Aiolos n'est pas affecté par les mutations. La perte de dimérisation d'Ikaros pourrait alors affecter la fonctionnalité d'Aiolos. Ceci tend à montrer que l'hétérodimère

Ikaros/Aiolos pourrait être la réelle entité assurant les fonctions d'Aiolos dans les lymphocytes B.

Quels seraient alors les mécanismes par lesquels l'hétérodimère pourrait assurer la fonctionnalité d'Ikaros et d'Aiolos ? Un aspect important de la fonctionnalité des facteurs de transcription de la famille d'Ikaros qui n'a pas été abordé dans ce travail est leur capacité de recruter des partenaires protéiques. Il est alors envisageable que chaque partenaire du dimère soit capable de recruter des partenaires particuliers. Pour étudier cette possibilité nous pourrions utiliser l'anticorps α Venus dans une immunoprécipitation suivie d'une analyse par spectrométrie de masse.

En conclusion, mon travail a permis d'une part de mettre en évidence des différences fonctionnelles entre les facteurs de transcription de la famille d'Ikaros. Et d'autre part, il pose la question de la fonctionnalité des dimères et surtout des hétérodimères. Ces derniers ne seraient-ils pas les réelles entités régulatrices, garantes des fonctions des facteurs de transcription de la famille d'Ikaros au cours de l'hématopoïèse ?

Références
Bibliographiques

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., et al. (2005). Identification of Flt3+ Lympho-Myeloid Stem Cells Lacking Erythro-Megakaryocytic Potential. *Cell* *121*, 295–306.

Ali, M.A.E., Fuse, K., Tadokoro, Y., Hoshii, T., Ueno, M., Kobayashi, M., Nomura, N., Vu, H.T., Peng, H., Hegazy, A.M., et al. (2017). Functional dissection of hematopoietic stem cell populations with a stemness-monitoring system based on NS-GFP transgene expression. *Sci Rep* *7*, 11442.

Alomairi, J., Molitor, A.M., Sadouni, N., Hussain, S., Torres, M., Saadi, W., Dao, L.T.M., Charbonnier, G., Santiago-Algarra, D., Andrau, J.C., et al. (2020). Integration of high-throughput reporter assays identify a critical enhancer of the *Ikzf1* gene. *PLOS ONE* *15*, e0233191.

Amin, R.H., and Schlissel, M.S. (2008). Foxo1 directly regulates the transcription of recombination-activating genes during B cell development. *Nature Immunology* *9*, 613–622.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biology* *11*, R106.

Apostolov, A., Litim-Mecheri, I., Oravec, A., Goepf, M., Kirstetter, P., Marchal, P., Ittel, A., Mauvieux, L., Chan, S., and Kastner, P. (2016). Sumoylation Inhibits the Growth Suppressive Properties of Ikaros. *PLoS One* *11*.

Arkatkar, T., Du, S.W., Jacobs, H.M., Dam, E.M., Hou, B., Buckner, J.H., Rawlings, D.J., and Jackson, S.W. (2017). B cell-derived IL-6 initiates spontaneous germinal center formation during systemic autoimmunity. *Journal of Experimental Medicine* *214*, 3207–3217.

Asanuma, S., Yamagishi, M., Kawanami, K., Nakano, K., Sato-Otsubo, A., Muto, S., Sanada, M., Yamochi, T., Kobayashi, S., Utsunomiya, A., et al. (2013). Adult T-cell leukemia cells are characterized by abnormalities of Helios expression that promote T cell growth. *Cancer Science* *104*, 1097–1106.

Avitahl, N., Winandy, S., Friedrich, C., Jones, B., Ge, Y., and Georgopoulos, K. (1999). Ikaros Sets Thresholds for T Cell Activation and Regulates Chromosome Propagation. *Immunity* *10*, 333–343.

Basta, J., and Rauchman, M. (2015). The Nucleosome Remodeling and Deacetylase (NuRD) Complex in Development and Disease. *Transl Res* *165*, 36–47.

Bertolino, E., Reddy, K., Medina, K.L., Parganas, E., Ihle, J., and Singh, H. (2005). Regulation of interleukin 7-dependent immunoglobulin heavy-chain variable gene rearrangements by transcription factor STAT5. *Nat Immunol* 6, 836–843.

Billot, K., Soeur, J., Chereau, F., Arrouss, I., Merle-Béral, H., Huang, M.-E., Mazier, D., Baud, V., and Rebollo, A. (2011). Deregulation of Aiolos expression in chronic lymphocytic leukemia is associated with epigenetic modifications. *Blood* 117, 1917–1927.

Bottardi, S., Mavoungou, L., Bourgoïn, V., Mashtalir, N., Affar, E.B., and Milot, E. (2013). Direct Protein Interactions Are Responsible for Ikaros-GATA and Ikaros-Cdk9 Cooperativeness in Hematopoietic Cells. *Mol Cell Biol* 33, 3064–3076.

Bottardi, S., Mavoungou, L., Pak, H., Daou, S., Bourgoïn, V., Lakehal, Y.A., Affar, E.B., and Milot, E. (2014). The IKAROS Interaction with a Complex Including Chromatin Remodeling and Transcription Elongation Activities Is Required for Hematopoiesis. *PLoS Genet* 10.

Brayer, K.J., and Segal, D.J. (2008). Keep Your Fingers Off My DNA: Protein–Protein Interactions Mediated by C2H2 Zinc Finger Domains. *Cell Biochemistry and Biophysics* 50, 111–131.

Brodsky, S., Jana, T., Mittelman, K., Chapal, M., Kumar, D.K., Carmi, M., and Barkai, N. (2020). Intrinsically Disordered Regions Direct Transcription Factor In Vivo Binding Specificity. *Molecular Cell* 79, 459-471.e4.

Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merkschlager, M., and Fisher, A.G. (1997). Association of Transcriptionally Silent Genes with Ikaros Complexes at Centromeric Heterochromatin. *Cell* 91, 845–854.

Bruno, L., Ramlall, V., Studer, R.A., Sauer, S., Bradley, D., Dharmalingam, G., Carroll, T., Ghoneim, M., Chopin, M., Nutt, S.L., et al. (2019). Selective deployment of transcription factor paralogs with submaximal strength facilitates gene regulation in the immune system. *Nature Immunology* 20, 1372–1380.

Buenrostro, J., Wu, B., Chang, H., and Greenleaf, W. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* 109, 21.29.1-21.29.9.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* 10, 1213–1218.

Caballero, R., Setien, F., Lopez-Serra, L., Boix-Chornet, M., Fraga, M.F., Ropero, S., Megias, D., Alaminos, M., Sanchez-Tapia, E.M., Montoya, M.C., et al. (2007). Combinatorial effects of splice variants modulate function of Aiolos. *J Cell Sci* 120, 2619–2630.

Cariappa, A., Tang, M., Parng, C., Nebelitskiy, E., Carroll, M., Georgopoulos, K., and Pillai, S. (2001). The Follicular versus Marginal Zone B Lymphocyte Cell Fate Decision Is Regulated by Aiolos, Btk, and CD21. *Immunity* 14, 603–615.

Carrère, S., Verger, A., Flourens, A., Stehelin, D., and Duterque-Coquillaud, M. (1998). Erg proteins, transcription factors of the Ets family, form homo, heterodimers and ternary complexes via two distinct domains. *Oncogene* 16, 3261–3268.

Cheng, H., Zheng, Z., and Cheng, T. (2020). New paradigms on hematopoietic stem cell differentiation. *Protein Cell* 11, 34–44.

Cheng, Y., Chikwava, K., Wu, C., Zhang, H., Bhagat, A., Pei, D., Choi, J.K., and Tong, W. (2016). LNK/SH2B3 regulates IL-7 receptor signaling in normal and malignant B-progenitors. *J Clin Invest* 126, 1267–1281.

Choo, Y., and Klug, A. (1993). A role in DNA binding for the linker sequences of the first three zinc fingers of TFIIIA. *Nucleic Acids Res* 21, 3341–3346.

Chou, W.-C., Levy, D.E., and Lee, C.-K. (2006). STAT3 positively regulates an early step in B-cell development. *Blood* 108, 3005–3011.

Clark, M.R., Mandal, M., Ochiai, K., and Singh, H. (2014). Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nat Rev Immunol* 14, 69–80.

Cobaleda, C., Schebesta, A., Delogu, A., and Busslinger, M. (2007). Pax5: the guardian of B cell identity and function. *Nature Immunology* 8, 463–470.

Cobb, B.S., Morales-Alcelay, S., Kleiger, G., Brown, K.E., Fisher, A.G., and Smale, S.T. (2000). Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. *Genes Dev* 14, 2146–2160.

Corcoran, A.E., Riddell, A., Krooshoop, D., and Venkitaraman, A.R. (1998). Impaired immunoglobulin gene rearrangement in mice lacking the IL-7 receptor. *Nature* 391, 904–907.

Cortés, M., and Georgopoulos, K. (2004). Aiolos Is Required for the Generation of High Affinity Bone Marrow Plasma Cells Responsible for Long-Term Immunity. *J Exp Med* 199, 209–219.

Crocker, J., Preger-Ben Noon, E., and Stern, D.L. (2016). Chapter Twenty-Seven - The Soft Touch: Low-Affinity Transcription Factor Binding Sites in Development and Evolution. In *Current Topics in Developmental Biology*, P.M. Wassarman, ed. (Academic Press), pp. 455–469.

De, S., and Barnes, B.J. (2014). B cell transcription factors: Potential new therapeutic targets for SLE. *Clinical Immunology* 152, 140–151.

Decker, T., Pasca di Magliano, M., McManus, S., Sun, Q., Bonifer, C., Tagoh, H., and Busslinger, M. (2009). Stepwise Activation of Enhancer and Promoter Regions of the B Cell Commitment Gene Pax5 in Early Lymphopoiesis. *Immunity* 30, 508–520.

DeKoter, R.P., and Singh, H. (2000). Regulation of B Lymphocyte and Macrophage Development by Graded Expression of PU.1. 288, 3.

DeKoter, R.P., Lee, H.-J., and Singh, H. (2002). PU.1 Regulates Expression of the Interleukin-7 Receptor in Lymphoid Progenitors. *Immunity* 16, 297–309.

Delogu, A., Schebesta, A., Sun, Q., Aschenbrenner, K., Perlot, T., and Busslinger, M. (2006). Gene Repression by Pax5 in B Cells Is Essential for Blood Cell Homeostasis and Is Reversed in Plasma Cells. *Immunity* 24, 269–281.

Dengler, H.S., Baracho, G.V., Omori, S.A., Bruckner, S., Arden, K.C., Castrillon, D.H., DePinho, R.A., and Rickert, R.C. (2008). Distinct functions for the transcription factor Foxo1 at various stages of B cell differentiation. *Nature Immunology* 9, 1388–1398.

Dias, S., Silva, H., Cumano, A., and Vieira, P. (2005). Interleukin-7 is necessary to maintain the B cell potential in common lymphoid progenitors. *Journal of Experimental Medicine* 201, 971–979.

Ding, Y., Zhang, B., Payne, J.L., Song, C., Ge, Z., Gowda, C., Iyer, S., Dhanyamraju, P.K., Dorsam, G., Reeves, M.E., et al. (2019). Ikaros tumor suppressor function includes induction of active enhancers and super-enhancers along with pioneering activity. *Leukemia*.

Dogan, B., Kailasam, S., Corchado, A.H., Nikpoor, N., and Najafabadi, H.S. (2019). A DNA recognition code for probing the *in vivo* functions of zinc finger transcription factors at domain resolution (Bioinformatics).

Dovat, S., Ronni, T., Russell, D., Ferrini, R., Cobb, B.S., and Smale, S.T. (2002). A common mechanism for mitotic inactivation of C2H2 zinc finger DNA-binding domains. *Genes Dev* 16, 2985–2990.

Dovat, S., Montecino-Rodriguez, E., Schuman, V., Teitell, M.A., Dorshkind, K., and Smale, S.T. (2005). Transgenic Expression of Helios in B Lineage Cells Alters B Cell

Properties and Promotes Lymphomagenesis. *The Journal of Immunology* 175, 3508–3515.

Duhamel, M., Arrouss, I., Merle-Béral, H., and Rebollo, A. (2008). The Aiolos transcription factor is up-regulated in chronic lymphocytic leukemia. *Blood* 111, 3225–3228.

Dumortier, A., Jeannet, R., Kirstetter, P., Kleinmann, E., Sellars, M., dos Santos, N.R., Thibault, C., Barths, J., Ghysdael, J., Punt, J.A., et al. (2006). Notch Activation Is an Early and Critical Event during T-Cell Leukemogenesis in Ikaros-Deficient Mice. *Molecular and Cellular Biology* 26, 209–220.

Elrod-Erickson, M., Rould, M.A., Nekludova, L., and Pabo, C.O. (1996). Zif268 protein–DNA complex refined at 1.6Å: a model system for understanding zinc finger–DNA interactions. *Structure* 4, 1171–1180.

Eskandarian, Z., Fliegau, M., Bulashevskaya, A., Proietti, M., Hague, R., Smulski, C.R., Schubert, D., Warnatz, K., and Grimbacher, B. (2019). Assessing the Functional Relevance of Variants in the IKAROS Family Zinc Finger Protein 1 (IKZF1) in a Cohort of Patients With Primary Immunodeficiency. *Front Immunol* 10.

Fernandez-Rodriguez, J., and Marlovits, T.C. (2012). Induced heterodimerization and purification of two target proteins by a synthetic coiled-coil tag: Induced Heterodimerization and Purification of Proteins. *Protein Science* 21, 511–519.

Ferrando, A.A. (2009). The role of NOTCH1 signaling in T-ALL. *Hematology Am Soc Hematol Educ Program* 353–361.

Fistonich, C., Zehentmeier, S., Bednarski, J.J., Miao, R., Schjerven, H., Sleckman, B.P., and Pereira, J.P. (2018). Cell circuits between B cell progenitors and IL-7⁺ mesenchymal progenitor cells control B cell development. *The Journal of Experimental Medicine* 215, 2586–2599.

Fleming, H.E., and Paige, C.J. (2002). Cooperation between IL-7 and the pre-B cell receptor: a key to B cell selection. *Seminars in Immunology* 14, 423–430.

Fornes, O., Castro-Mondragon, J.A., Khan, A., van der Lee, R., Zhang, X., Richmond, P.A., Modi, B.P., Correard, S., Gheorghe, M., Baranašić, D., et al. (2020). JASPAR 2020: update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 48, D87–D92.

Funnell, A.P.W., and Crossley, M. (2012). Homo- and Heterodimerization in Transcriptional Regulation. In *Protein Dimerization and Oligomerization in Biology*, J.M. Matthews, ed. (New York, NY: Springer New York), pp. 105–121.

Galloway, J.L., and Zon, L.I. (2003). Ontogeny of hematopoiesis: examining the emergence of hematopoietic cells in the vertebrate embryo. *Curr Top Dev Biol* 53, 139–158.

Gandhi, A.K., Kang, J., Havens, C.G., Conklin, T., Ning, Y., Wu, L., Ito, T., Ando, H., Waldman, M.F., Thakurta, A., et al. (2014). Immunomodulatory agents lenalidomide and pomalidomide co-stimulate T cells by inducing degradation of T cell repressors Ikaros and Aiolos via modulation of the E3 ubiquitin ligase complex CRL4CRBN. *British Journal of Haematology* 164, 811–821.

Gay, D., Saunders, T., Camper, S., and Weigert, M. (1993). Receptor editing: an approach by autoreactive B cells to escape tolerance. *Journal of Experimental Medicine* 177, 999–1008.

Ge, Z., Gu, Y., Xiao, L., Han, Q., Li, J., Chen, B., Yu, J., Kawasaki, Y.I., Payne, K.J., Dovat, S., et al. (2016). Co-existence of IL7R high and SH2B3 low expression distinguishes a novel high-risk acute lymphoblastic leukemia with Ikaros dysfunction. *Oncotarget* 7, 46014–46027.

Georgopoulos, K., Bigby, M., Wang, J.-H., Molnar, A., Wu, P., Winandy, S., and Sharpe, A. (1994). The ikaros gene is required for the development of all lymphoid lineages. *Cell* 79, 143–156.

Georgopoulos, K., Winandy, S., and Avitahl, N. (1997). The Role of the Ikaros Gene in Lymphocyte Development and Homeostasis. *Annual Review of Immunology* 15, 155–176.

Gilfillan, S., Dierich, A., Lemeur, M., Benoist, C., and Mathis, D. (1993). Mice lacking TdT: mature animals with an immature lymphocyte repertoire. *Science* 261, 1175–1178.

Gilfillan, S., Benoist, C., and Mathis, D. (1995). Mice Lacking Terminal Deoxynucleotidyl Transferase: Adult Mice with a Fetal Antigen Receptor Repertoire. *Immunological Reviews* 148, 201–219.

Glick, B., Chang, T.S., and Jaap, R.G. (1956). The Bursa of Fabricius and Antibody Production. *Poultry Science* 35, 224–225.

Gómez-del Arco, P., Maki, K., and Georgopoulos, K. (2004). Phosphorylation Controls Ikaros's Ability To Negatively Regulate the G1-S Transition. *Mol Cell Biol* 24, 2797–2807.

Gómez-del Arco, P., Koipally, J., and Georgopoulos, K. (2005). Ikaros SUMOylation: Switching Out of Repression. *Mol Cell Biol* 25, 2688–2697.

Goodnow, C.C., Crosbie, J., Adelstein, S., Lavoie, T.B., Smith-Gill, S.J., Brink, R.A., Pritchard-Briscoe, H., Wotherspoon, J.S., Loblay, R.H., Raphael, K., et al. (1988). Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature* 334, 676–682.

Graves, B.J., and Petersen, J.M. (1998). Specificity within the ets Family of Transcription Factors. In *Advances in Cancer Research*, (Elsevier), pp. 1–57.

Gurel, Z., Ronni, T., Ho, S., Kuchar, J., Payne, K.J., Turk, C.W., and Dovati, S. (2008). Recruitment of Ikaros to Pericentromeric Heterochromatin Is Regulated by Phosphorylation. *J Biol Chem* 283, 8291–8300.

Hahm, K., Ernst, P., Lo, K., Kim, G.S., Turck, C., and Smale, S.T. (1994). The lymphoid transcription factor LyF-1 is encoded by specific, alternatively spliced mRNAs derived from the Ikaros gene. *Mol Cell Biol* 14, 7111–7123.

Hahm, K., Cobb, B.S., McCarty, A.S., Brown, K.E., Klug, C.A., Lee, R., Akashi, K., Weissman, I.L., Fisher, A.G., and Smale, S.T. (1998). Helios, a T cell-restricted Ikaros family member that quantitatively associates with Ikaros at centromeric heterochromatin. *Genes Dev* 12, 782–796.

Hardy, R.R., Carmack, C.E., Shinton, S.A., Kemp, J.D., and Hayakawa, K. (1991). Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *J Exp Med* 173, 1213–1225.

Heerema, N.A., Nachman, J.B., Sather, H.N., La, M.K., Hutchinson, R., Lange, B.J., Bostrom, B., Steinherz, P.G., Gaynon, P.S., Uckun, F.M., et al. (2004). Deletion of 7p or monosomy 7 in pediatric acute lymphoblastic leukemia is an adverse prognostic factor: a report from the Children’s Cancer Group. *Leukemia* 18, 939–947.

Heizmann, B., Kastner, P., and Chan, S. (2013). Ikaros is absolutely required for pre-B cell differentiation by attenuating IL-7 signals. *J Exp Med* 210, 2823–2832.

Heizmann, B., Sellars, M., Macias-Garcia, A., Chan, S., and Kastner, P. (2016). Ikaros limits follicular B cell activation by regulating B cell receptor signaling pathways. *Biochemical and Biophysical Research Communications* 470, 714–720.

Heizmann, B., Kastner, P., and Chan, S. (2018). The Ikaros family in lymphocyte development. *Current Opinion in Immunology* 51, 14–23.

Heizmann, B., Le Gras, S., Simand, C., Marchal, P., Chan, S., and Kastner, P. (2020). Ikaros antagonizes DNA binding by STAT5 in pre-B cells. *PLoS One* 15.

Heltemes-Harris, L.M., Willette, M.J.L., Vang, K.B., and Farrar, M.A. (2011). The role of STAT5 in the development, function, and transformation of B and T lymphocytes. *Annals of the New York Academy of Sciences* 1217, 18–31.

Herzog, S., Hug, E., Meixlsperger, S., Paik, J.-H., DePinho, R.A., Reth, M., and Jumaa, H. (2008). SLP-65 regulates immunoglobulin light chain gene recombination through the PI(3)K-PKB-Foxo pathway. *Nature Immunology* 9, 623–631.

Hieter, P.A., Korsmeyer, S.J., Waldmann, T.A., and Leder, P. (1981). Human immunoglobulin κ light-chain genes are deleted or rearranged in λ -producing B cells. *Nature* 290, 368–372.

Hossain, M.A., Barrow, J.J., Shen, Y., Haq, M.I., and Bungert, J. (2015). Artificial Zinc Finger DNA Binding Domains: Versatile Tools for Genome Engineering and Modulation of Gene Expression. *Journal of Cellular Biochemistry* 116, 2435–2444.

Hu, Y., Zhang, Z., Kashiwagi, M., Yoshida, T., Joshi, I., Jena, N., Somasundaram, R., Emmanuel, A.O., Sigvardsson, M., Fitamant, J., et al. (2016). Superenhancer reprogramming drives a B-cell-epithelial transition and high-risk leukemia. *Genes Dev.* 30, 1971–1990.

Hung, K.-H., Su, S.-T., Chen, C.-Y., Hsu, P.-H., Huang, S.-Y., Wu, W.-J., Chen, M.-J.M., Chen, H.-Y., Wu, P.-C., Lin, F.-R., et al. (2016). Aiolos collaborates with Blimp-1 to regulate the survival of multiple myeloma cells. *Cell Death Differ* 23, 1175–1184.

Iacobucci, I., and Mullighan, C.G. (2017). Genetic Basis of Acute Lymphoblastic Leukemia. *Journal of Clinical Oncology* 35, 975–983.

Inlay, M.A., Bhattacharya, D., Sahoo, D., Serwold, T., Seita, J., Karsunky, H., Plevritis, S.K., Dill, D.L., and 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, 2376–2381.

Isakova, A., Berset, Y., Hatzimanikatis, V., and Deplancke, B. (2016). Quantification of Cooperativity in Heterodimer-DNA Binding Improves the Accuracy of Binding Specificity Models. *J Biol Chem* 291, 10293–10306.

Jantz, D., and Berg, J.M. (2004). Reduction in DNA-binding affinity of Cys2His2 zinc finger proteins by linker phosphorylation. *Proc Natl Acad Sci U S A* 101, 7589–7593.

Jeannet, R., Mastio, J., Macias-Garcia, A., Oravec, A., Ashworth, T., Geimer Le Lay, A.-S., Jost, B., Le Gras, S., Ghysdael, J., Gridley, T., et al. (2010). Oncogenic activation of the Notch1 gene by deletion of its promoter in Ikaros-deficient T-ALL. *Blood* 116, 5443–5454.

John, L.B., and Ward, A.C. (2011). The Ikaros gene family: Transcriptional regulators of hematopoiesis and immunity. *Molecular Immunology* 48, 1272–1278.

Jolma, A., Yan, J., Whittington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-Binding Specificities of Human Transcription Factors. *Cell* 152, 327–339.

Joshi, I., Yoshida, T., Jena, N., Qi, X., Zhang, J., Van Etten, R.A., and Georgopoulos, K. (2014). Loss of Ikaros DNA-binding function confers integrin-dependent survival on pre-B cells and progression to acute lymphoblastic leukemia. *Nature Immunology* 15, 294–304.

Kaplan, T., Friedman, N., and Margalit, H. (2005). Ab Initio Prediction of Transcription Factor Targets Using Structural Knowledge. *PLoS Computational Biology* 1, e1.

Karnowski, A., Cao, C., Matthias, G., Carotta, S., Corcoran, L.M., Martensson, I.-L., Skok, J.A., and Matthias, P. (2008). Silencing and Nuclear Repositioning of the $\lambda 5$ Gene Locus at the Pre-B Cell Stage Requires Aiolos and OBF-1. *PLOS ONE* 3, e3568.

Katerndahl, C.D.S., Heltemes-Harris, L.M., Willette, M.J.L., Henzler, C.M., Fietze, S., Yang, R., Schjerven, H., Silverstein, K.A.T., Ramsey, L.B., Hubbard, G., et al. (2017). STAT5 antagonism of B cell enhancer networks drives leukemia and poor patient survival. *Nat Immunol* 18, 694–704.

Kelley, C.M., Ikeda, T., Koipally, J., Avitahl, N., Wu, L., Georgopoulos, K., and Morgan, B.A. (1998). Helios, a novel dimerization partner of Ikaros expressed in the earliest hematopoietic progenitors. *Current Biology* 8, 508-S1.

Kim, B., and Little, J.W. (1992). Dimerization of a Specific DNA-Binding. *Science* 255, 203–206.

Kim, H.-J., Barnitz, R.A., Kreslavsky, T., Brown, F.D., Moffett, H., Lemieux, M.E., Kaygusuz, Y., Meissner, T., Holderried, T.A.W., Chan, S., et al. (2015). Stable inhibitory activity of regulatory T cells requires the transcription factor Helios. *Science* 350, 334–339.

Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T., et al. (1999). Ikaros DNA-Binding Proteins Direct Formation of Chromatin Remodeling Complexes in Lymphocytes. *Immunity* 10, 345–355.

Kirstetter, P., Thomas, M., Dierich, A., Kastner, P., and Chan, S. (2002). Ikaros is critical for B cell differentiation and function. *European Journal of Immunology* 32, 720–730.

Kohler, J.J., and Schepartz, A. (2001). Kinetic Studies of Fos·Jun·DNA Complex Formation: DNA Binding Prior to Dimerization †. *Biochemistry* 40, 130–142.

Koipally, J., and Georgopoulos, K. (2000). Ikaros Interactions with CtBP Reveal a Repression Mechanism That Is Independent of Histone Deacetylase Activity. *J. Biol. Chem.* 275, 19594–19602.

Koipally, J., and Georgopoulos, K. (2002a). Ikaros-CtIP Interactions Do Not Require C-terminal Binding Protein and Participate in a Deacetylase-independent Mode of Repression. *J. Biol. Chem.* 277, 23143–23149.

Koipally, J., and Georgopoulos, K. (2002b). A Molecular Dissection of the Repression Circuitry of Ikaros. *J. Biol. Chem.* 277, 27697–27705.

Koipally, J., Renold, A., Kim, J., and Georgopoulos, K. (1999). Repression by Ikaros and Aiolos is mediated through histone deacetylase complexes. *EMBO J* 18, 3090–3100.

Koipally, J., Heller, E.J., Seavitt, J.R., and Georgopoulos, K. (2002). Unconventional Potentiation of Gene Expression by Ikaros. *J. Biol. Chem.* 277, 13007–13015.

Kondo, M., Weissman, I.L., and Akashi, K. (1997). Identification of Clonogenic Common Lymphoid Progenitors in Mouse Bone Marrow. *Cell* 91, 661–672.

Kudo, A., and Melchers, F. (1987). A second gene, VpreB in the lambda 5 locus of the mouse, which appears to be selectively expressed in pre-B lymphocytes. *The EMBO Journal* 6, 2267–2272.

Kuehn, H.S., Boisson, B., Cunningham-Rundles, C., Reichenbach, J., Stray-Pedersen, A., Gelfand, E.W., Maffucci, P., Pierce, K.R., Abbott, J.K., Voelkerding, K.V., et al. (2016). Loss of B Cells in Patients with Heterozygous Mutations in IKAROS. *N Engl J Med* 374, 1032–1043.

Kuehn, H.S., Niemela, J.E., Stoddard, J., Mannurita, S.C., Shahin, T., Goel, S., Hintermeyer, M., Heredia, R.J., Garofalo, M., Lucas, L., et al. (2021). Germline IKAROS dimerization haploinsufficiency causes hematologic cytopenias and malignancies. *Blood* 137, 349–363.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9, 357–359.

Li, Y., Liu, Y., Liu, C., Liu, F., Dou, D., Zheng, W., Liu, W., and Liu, F. (2018). Role of a non-canonical splice variant of the Helios gene in the differentiation of acute lymphoblastic leukemic T cells. *Oncology Letters* 15, 6957–6966.

Li, Z., Perez-Casellas, L.A., Savic, A., Song, C., and Dovati, S. (2011). Ikaros isoforms: The saga continues. *World J Biol Chem* 2, 140–145.

Liang, Z., Brown, K.E., Carroll, T., Taylor, B., Vidal, I.F., Hendrich, B., Rueda, D., Fisher, A.G., and Merckenschlager, M. (2017). A high-resolution map of transcriptional repression. *ELife* 6, e22767.

Lin, Y.C., Jhunjhunwala, S., Benner, C., Heinz, S., Welinder, E., Mansson, R., Sigvardsson, M., Hagman, J., Espinoza, C.A., Dutkowski, J., et al. (2010). A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nature Immunology* 11, 635–643.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15, 550.

Lu, R. (2003). IRF-4,8 orchestrate the pre-B-to-B transition in lymphocyte development. *Genes & Development* 17, 1703–1708.

Ma, H., Qazi, S., Ozer, Z., Zhang, J., Ishkhanian, R., and Uckun, F.M. (2013). Regulatory Phosphorylation of Ikaros by Bruton's Tyrosine Kinase. *PLOS ONE* 8, e71302.

Ma, S., Pathak, S., Trinh, L., and Lu, R. (2008). Interferon regulatory factors 4 and 8 induce the expression of Ikaros and Aiolos to down-regulate pre-B-cell receptor and promote cell-cycle withdrawal in pre-B-cell development. *Blood* 111, 1396–1403.

Ma, S., Pathak, S., Mandal, M., Trinh, L., Clark, M.R., and Lu, R. (2010). Ikaros and Aiolos Inhibit Pre-B-Cell Proliferation by Directly Suppressing c-Myc Expression. *Mol. Cell. Biol.* 30, 4149–4158.

Mandal, M., Powers, S.E., Maienschein-Cline, M., Bartom, E.T., Hamel, K.M., Kee, B.L., Dinner, A.R., and Clark, M.R. (2011). Epigenetic repression of the Igk locus by STAT5-mediated recruitment of the histone methyltransferase Ezh2. *Nat Immunol* 12, 1212–1220.

Mansson, R., Welinder, E., Ahsberg, J., Lin, Y.C., Benner, C., Glass, C.K., Lucas, J.S., Sigvardsson, M., and Murre, C. (2012). Positive intergenic feedback circuitry, involving EBF1 and FOXO1, orchestrates B-cell fate. *Proceedings of the National Academy of Sciences* 109, 21028–21033.

Mantha, S., Ward, M., McCafferty, J., Herron, A., Palomero, T., Ferrando, A., Bank, A., and Richardson, C. (2007). Activating Notch1 mutations are an early event in T-cell malignancy of Ikaros point mutant Plastic/+ mice. *Leukemia Research* 31, 321–327.

Marçais, A., Jeannet, R., Hernandez, L., Soulier, J., Sigaux, F., Chan, S., and Kastner, P. (2010). Genetic inactivation of Ikaros is a rare event in human T-ALL. *Leuk Res* 34, 426–429.

Marke, R., van Leeuwen, F.N., and Scheijen, B. (2018). The many faces of IKZF1 in B-cell precursor acute lymphoblastic leukemia. *Haematologica* 103, 565–574.

Martín-Ibáñez, R., Pardo, M., Giralt, A., Miguez, A., Guardia, I., Marion-Poll, L., Herranz, C., Esgleas, M., Garcia-Díaz Barriga, G., Edel, M.J., et al. (2017). Helios expression coordinates the development of a subset of striatopallidal medium spiny neurons. *Development* 144, 1566–1577.

McCarty, A.S., Kleiger, G., Eisenberg, D., and Smale, S.T. (2003). Selective Dimerization of a C2H2 Zinc Finger Subfamily. *Molecular Cell* 11, 459–470.

McGuire, K.L., and Vitetta, E.S. (1981). kappa/lambda Shifts do not occur during maturation of murine B cells. *The Journal of Immunology* 127, 1670–1673.

Melchers, F., Rolink, A., Grawunder, U., Winkler, T.H., Karasuyama, H., Ghia, P., and Andersson, J. (1995). Positive and negative selection events during B lymphopoiesis. *Current Opinion in Immunology* 7, 214–227.

Molnár, A., and Georgopoulos, K. (1994). The Ikaros gene encodes a family of functionally diverse zinc finger DNA-binding proteins. *Mol Cell Biol* 14, 8292–8303.

Molnár, A., Wu, P., Largespada, D.A., Vortkamp, A., Scherer, S., Copeland, N.G., Jenkins, N.A., Bruns, G., and Georgopoulos, K. (1996). The Ikaros gene encodes a family of lymphocyte-restricted zinc finger DNA binding proteins, highly conserved in human and mouse. *The Journal of Immunology* 156, 585–592.

Mombaerts, P., Iacomini, J., Johnson, R.S., Herrup, K., Tonegawa, S., and Papaioannou, V.E. (1992). RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* 68, 869–877.

Morel, G., Deau, M.-C., Simand, C., Caye-Eude, A., Arfeuille, C., Ittel, A., Miguet, L., Mauvieux, L., Herbrecht, R., Paillard, C., et al. (2019). Large deletions of the 5' region of IKZF1 lead to haploinsufficiency in B-cell precursor acute lymphoblastic leukaemia. *British Journal of Haematology* 186, e155–e159.

Morgan, B., Sun, L., Avitahl, N., Andrikopoulos, K., Ikeda, T., Gonzales, E., Wu, P., Neben, S., and Georgopoulos, K. (1997). Aiolos, a lymphoid restricted transcription factor that interacts with Ikaros to regulate lymphocyte differentiation. *EMBO J* 16, 2004–2013.

Nakase, K., Ishimaru, F., Fujii, K., Tabayashi, T., Kozuka, T., Sezaki, N., Matsuo, Y., and Harada, M. (2002). Overexpression of novel short isoforms of Helios in a patient with T-cell acute lymphoblastic leukemia. *Experimental Hematology* 30, 313–317.

Nakayama, Y., Kosek, J., Capone, L., Hur, E.M., Schafer, P.H., and Ringheim, G.E. (2017). Aiolos Overexpression in Systemic Lupus Erythematosus B Cell Subtypes and BAFF-Induced Memory B Cell Differentiation Are Reduced by CC-220 Modulation of Cereblon Activity. *J Immunol* 199, 2388–2407.

Nemazee, D.A., and Bürki, K. (1989). Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature* 337, 562–566.

Ng, S.Y.-M., Yoshida, T., Zhang, J., and Georgopoulos, K. (2009). Genome-wide Lineage-Specific Transcriptional Networks Underscore Ikaros-Dependent Lymphoid Priming in Hematopoietic Stem Cells. *Immunity* 30, 493–507.

Nielsen, P.J., Georgiev, O., Lorenz, B., and Schaffner, W. (1996). B lymphocytes are impaired in mice lacking the transcriptional co-activator Bob1/OCA-B/OBF1. *European Journal of Immunology* 26, 3214–3218.

Nurmemmedov, E., Yengo, R.K., Uysal, H., Karlsson, R., and Thunnissen, M.M.G.M. (2009). New insights into DNA-binding behavior of Wilms Tumor Protein (WT1) — A dual study. *Biophysical Chemistry* 145, 116–125.

Nye, J.A., Petersen, J.M., Gunther, C.V., Jonsen, M.D., and Graves, B.J. (1992). Interaction of murine ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. *Genes Dev.* 6, 975–990.

Oravec, A., Apostolov, A., Polak, K., Jost, B., Le Gras, S., Chan, S., and Kastner, P. (2015). Ikaros mediates gene silencing in T cells through Polycomb repressive complex 2. *Nat Commun* 6.

Papathanasiou, P., Perkins, A.C., Cobb, B.S., Ferrini, R., Sridharan, R., Hoyne, G.F., Nelms, K.A., Smale, S.T., and Goodnow, C.C. (2003). Widespread Failure of Hematolymphoid Differentiation Caused by a Recessive Niche-Filling Allele of the Ikaros Transcription Factor. *Immunity* 19, 131–144.

Park, S.-M., Cho, H., Thornton, A.M., Barlowe, T.S., Chou, T., Chhangawala, S., Fairchild, L., Taggart, J., Chow, A., Schurer, A., et al. (2018). IKZF2 Drives Leukemia Stem Cell Self-Renewal and Inhibits Myeloid Differentiation. *Cell Stem Cell*.

Parker, M., Licence, S., Erlandsson, L., Galler, G., Chakalova, L., Osborne, C.S., Morgan, G., Fraser, P., Jumaa, H., Winkler, T., et al. (2005). The pre-B-cell receptor induces silencing of VpreB and lambda5 transcription. *EMBO J* 24, 3895–3905.

Persikov, A.V., Wetzel, J.L., Rowland, E.F., Oakes, B.L., Xu, D.J., Singh, M., and Noyes, M.B. (2015). A systematic survey of the Cys2His2 zinc finger DNA-binding landscape. *Nucleic Acids Research* 43, 1965–1984.

Pietras, E.M., Reynaud, D., Kang, Y.-A., Carlin, D., Calero-Nieto, F.J., Leavitt, A.D., Stuart, J.M., Göttgens, B., and Passegué, E. (2015). Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem Cell* 17, 35–46.

Rebollo, A., and Schmitt, C. (2003). Ikaros, Aiolos and Helios: Transcription regulators and lymphoid malignancies. *Immunol Cell Biol* 81, 171–175.

Reynaud, D., Demarco, I., Reddy, K., Schjerven, H., Bertolino, E., Chen, Z., Smale, S.T., Winandy, S., and Singh, H. (2008). Regulation of B cell fate commitment and immunoglobulin VH gene rearrangements by Ikaros. *Nat Immunol* 9, 927–936.

Rickert, R.C. (2013). New insights into pre-BCR and BCR signalling with relevance to B cell malignancies. *Nature Reviews Immunology* 13, 578–591.

Roberts, C.W.M., and Orkin, S.H. (2004). The SWI/SNF complex — chromatin and cancer. *Nature Reviews Cancer* 4, 133–142.

Rolink, A., Grawunder, U., Haasner, D., Strasser, A., and Mdchers, F. (1993). Immature Surface Ig δ B Cells Can Continue to Rearrange Ic and L Chain Gene Loci. 8.

Sabbattini, P., Lundgren, M., Georgiou, A., Chow, C., Warnes, G., and Dillon, N. (2001). Binding of Ikaros to the $\lambda 5$ promoter silences transcription through a mechanism that does not require heterochromatin formation. *The EMBO Journal* 20, 2812–2822.

Sakaguchi, N., and Melchers, F. (1986). Lambda 5, a new light-chain-related locus selectively expressed in pre-B lymphocytes. *Nature* 324, 579–582.

Schebesta, M., Pfeffer, P.L., and Busslinger, M. (2002). Control of Pre-BCR Signaling by Pax5-Dependent Activation of the BLNK Gene. *Immunity* 17, 473–485.

Schjerven, H., McLaughlin, J., Arenzana, T.L., Fietze, S., Cheng, D., Wadsworth, S., Lawson, G.W., Bensinger, S.J., Farnham, P.J., Witte, O.N., et al. (2013). Differential regulation of lymphopoiesis and leukemogenesis by individual zinc fingers of Ikaros. *Nat Immunol* 14, 1073–1083.

Schjerven, H., Ayongaba, E.F., Aghajani-refah, A., McLaughlin, J., Cheng, D., Geng, H., Boyd, J.R., Eggesbo, L.M., Lindeman, I., Heath, J.L., et al. (2017). Genetic analysis

of Ikaros target genes and tumor suppressor function in BCR-ABL1(+) pre-B ALL. *JOURNAL OF EXPERIMENTAL MEDICINE* 214.

Schwickert, T.A., Tagoh, H., Gültekin, S., Dakic, A., Axelsson, E., Minnich, M., Ebert, A., Werner, B., Roth, M., Cimmino, L., et al. (2014). Stage-specific control of early B cell development by the transcription factor Ikaros. *Nat Immunol* 15, 283–293.

Schwickert, T.A., Tagoh, H., Schindler, K., Fischer, M., Jaritz, M., and Busslinger, M. (2019). Ikaros prevents autoimmunity by controlling anergy and Toll-like receptor signaling in B cells. *Nat Immunol* 20, 1517–1529.

Sellars, M., Reina-San-Martin, B., Kastner, P., and Chan, S. (2009). Ikaros controls isotype selection during immunoglobulin class switch recombination. *Journal of Experimental Medicine* 206, 1073–1087.

Sellars, M., Kastner, P., and Chan, S. (2011). Ikaros in B cell development and function. *World J Biol Chem* 2, 132–139.

Shinkai, Y., Rathbun, Gary, Lam, K.-P., Oltz, E.M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A.M., et al. (1992). RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* 68, 855–867.

Sridharan, R., and Smale, S.T. (2007). Predominant Interaction of Both Ikaros and Helios with the NuRD Complex in Immature Thymocytes. *J. Biol. Chem.* 282, 30227–30238.

Sun, J., Matthias, G., Mihatsch, M.J., Georgopoulos, K., and Matthias, P. (2003). Lack of the Transcriptional Coactivator OBF-1 Prevents the Development of Systemic Lupus Erythematosus-Like Phenotypes in Aiolos Mutant Mice. *The Journal of Immunology* 170, 1699–1706.

Sun, L., Liu, A., and Georgopoulos, K. (1996). Zinc finger-mediated protein interactions modulate Ikaros activity, a molecular control of lymphocyte development. *The EMBO Journal* 15, 5358–5369.

Takanashi, M., Yagi, T., Imamura, T., Tabata, Y., Morimoto, A., Hibi, S., Ishii, E., and Imashuku, S. (2002). Expression of the Ikaros gene family in childhood acute lymphoblastic leukaemia. *British Journal of Haematology* 117, 525–530.

Takatori, H., Kawashima, H., Matsuki, A., Meguro, K., Tanaka, S., Iwamoto, T., Sanayama, Y., Nishikawa, N., Tamachi, T., Ikeda, K., et al. (2015). Helios Enhances Treg Cell Function in Cooperation With FoxP3: ROLE OF HELIOS IN INDUCED TREG CELL FUNCTION. *Arthritis & Rheumatology* 67, 1491–1502.

Thompson, E.C., Cobb, B.S., Sabbattini, P., Meixlsperger, S., Parelho, V., Liberg, D., Taylor, B., Dillon, N., Georgopoulos, K., Jumaa, H., et al. (2007). Ikaros DNA-Binding Proteins as Integral Components of B Cell Developmental-Stage-Specific Regulatory Circuits. *Immunity* 26, 335–344.

Ubieta, K., Garcia, M., Grötsch, B., Uebe, S., Weber, G.F., Stein, M., Ekici, A., Schett, G., Mielenz, D., and Bozec, A. (2017). Fra-2 regulates B cell development by enhancing IRF4 and Foxo1 transcription. *Journal of Experimental Medicine* 214, 2059–2071.

Uckun, F.M., Ma, H., Zhang, J., Ozer, Z., Dovat, S., Mao, C., Ishkhanian, R., Goodman, P., and Qazi, S. (2012). Serine phosphorylation by SYK is critical for nuclear localization and transcription factor function of Ikaros. *Proc Natl Acad Sci U S A* 109, 18072–18077.

Van Nieuwenhove, E., Garcia-Perez, J.E., Helsen, C., Rodriguez, P.D., van Schouwenburg, P.A., Dooley, J., Schlenner, S., van der Burg, M., Verhoeyen, E., Gijssbers, R., et al. (2018). A kindred with Mutant IKAROS and autoimmunity. *J Allergy Clin Immunol* 142, 699-702.e12.

Virely, C., Moulin, S., Cobaleda, C., Lasgi, C., Alberdi, A., Soulier, J., Sigaux, F., Chan, S., Kastner, P., and Ghysdael, J. (2010). Haploinsufficiency of the IKZF1 (IKAROS) tumor suppressor gene cooperates with BCR-ABL in a transgenic model of acute lymphoblastic leukemia. *Leukemia* 24, 1200–1204.

Wagih, O. (2017). ggseqlogo: a versatile R package for drawing sequence logos. *Bioinformatics* 33, 3645–3647.

Wang, H., Xu, Z., Du, W., Lin, Z., and Liu, Z. (2019). N160 of Aiolos Determines its DNA-Binding Activity. *The Anatomical Record* 302, 2014–2019.

Wang, J., Malecka, A., Trøen, G., and Delabie, J. (2015). Comprehensive genome-wide transcription factor analysis reveals that a combination of high affinity and low affinity DNA binding is needed for human gene regulation. *BMC Genomics* 16, S12.

Wang, J.-H., Nichogiannopoulou, A., Wu, L., Sun, L., Sharpe, A.H., Bigby, M., and Georgopoulos, K. (1996). Selective Defects in the Development of the Fetal and Adult Lymphoid System in Mice with an Ikaros Null Mutation. *Immunity* 5, 537–549.

Wang, J.-H., Avitahl, N., Cariappa, A., Friedrich, C., Ikeda, T., Renold, A., Andrikopoulos, K., Liang, L., Pillai, S., Morgan, B.A., et al. (1998). Aiolos Regulates B Cell Activation and Maturation to Effector State. *Immunity* 9, 543–553.

Weller, S., and Descatoire, M. (2015). Les lymphocytes B IgM⁺ IgD⁺ CD27⁺ chez l'homme: Un rôle essentiel dans la défense contre les bactéries encapsulées. *médecine/sciences* 31, 647–653.

Westman, B.J., Perdomo, J., Matthews, J.M., Crossley, M., and Mackay, J.P. (2004). Structural Studies on a Protein-Binding Zinc-Finger Domain of Eos Reveal Both Similarities and Differences to Classical Zinc Fingers. *Biochemistry* 43, 13318–13327.

Wilson, N.K., Kent, D.G., Buettner, F., Shehata, M., Macaulay, I.C., Calero-Nieto, F.J., Sánchez Castillo, M., Oedekoven, C.A., Diamanti, E., Schulte, R., et al. (2015). Combined Single-Cell Functional and Gene Expression Analysis Resolves Heterogeneity within Stem Cell Populations. *Cell Stem Cell* 16, 712–724.

Winandy, S., Wu, P., and Georgopoulos, K. (1995). A dominant mutation in the Ikaros gene leads to rapid development of leukemia and lymphoma. *Cell* 83, 289–299.

Winandy, S., Wu, L., Wang, J.-H., and Georgopoulos, K. (1999). Pre-T Cell Receptor (Tcr) and Tcr-Controlled Checkpoints in T Cell Differentiation Are Set by Ikaros. *Journal of Experimental Medicine* 190, 1039–1048.

Winkler, T.H., Melchers, F., and Rolink, A.G. (1995). Interleukin-3 and Interleukin-7 Are Alternative Growth Factors for the Same B-Cell Precursors in the Mouse. *Blood* 85, 2045–2051.

Wojcik, H., Griffiths, E., Staggs, S., Hagman, J., and Winandy, S. (2007). Expression of a non-DNA-binding Ikaros isoform exclusively in B cells leads to autoimmunity but not leukemogenesis. *Eur J Immunol* 37, 1022–1032.

Yan, M., Brady, J.R., Chan, B., Lee, W.P., Hsu, B., Harless, S., Cancro, M., Grewal, I.S., and Dixit, V.M. (2001). Identification of a novel receptor for B lymphocyte stimulator that is mutated in a mouse strain with severe B cell deficiency. *Current Biology* 11, 1547–1552.

Yoshida, T., Ng, S.Y.-M., Zuniga-Pflucker, J.C., and Georgopoulos, K. (2006). Early hemopoietic lineage restrictions directed by Ikaros. *Nat Immunol* 7.

Yoshida, T., Landhuis, E., Dose, M., Hazan, I., Zhang, J., Naito, T., Jackson, A.F., Wu, J., Perotti, E.A., Kaufmann, C., et al. (2013). Transcriptional regulation of the Ikaros locus. *Blood* 122, 3149–3159.

Zandi, S., Mansson, R., Tsapogas, P., Zetterblad, J., Bryder, D., and Sigvardsson, M. (2008). EBF1 Is Essential for B-Lineage Priming and Establishment of a Transcription Factor Network in Common Lymphoid Progenitors. *The Journal of Immunology* 181, 3364–3372.

Zarnegar, M.A., and Rothenberg, E.V. (2012). Ikaros represses and activates PU.1 cell-type-specifically through the multifunctional Sfpi1 URE and a myeloid specific enhancer. *Oncogene* 31, 4647–4654.

Zhuang, Y., Li, D., Fu, J., Shi, Q., Lu, Y., and Ju, X. (2014). Overexpression of AIOLOS inhibits cell proliferation and suppresses apoptosis in Nalm-6 cells. *Oncology Reports* 31, 1183–1190.

Annexes

Large deletions of the 5' region of *IKZF1* lead to haploinsufficiency in B-cell precursor acute lymphoblastic leukaemia

Deletions or mutations of *IKZF1*, encoding the IKAROS transcription factor, are associated with poor prognosis in some subtypes of B-cell precursor acute lymphoblastic leukaemia (BCP-ALL) (Mullighan *et al*, 2009; Clappier *et al*, 2015). *IKZF1* status is therefore important for risk stratification and defining therapies (Stanulla *et al*, 2018; Yeoh *et al*, 2018). Most *IKZF1* abnormalities are intragenic deletions, or deletions of the entire locus (Mullighan *et al*, 2008; Iacobucci *et al*, 2009; Marke *et al*, 2018), which lead to loss of IKAROS expression and/or function. However, deletions affecting only the sequences 5' of the coding region, including the non-coding exon 1, have also been detected by comparative genomic hybridisation-arrays (Iacobucci *et al*, 2009, 2012; Dupuis *et al*, 2013) and multiplex ligation-dependent probe amplification (MLPA), which is routinely used to assess *IKZF1* status in patients (Fig 1A) (Caye *et al*, 2013). Because the sequences surrounding exon 1 are rich in G/C nucleotides, their detection by MLPA may sometimes be artefactual. Furthermore, the breakpoints of the 5' deletions, as well as their consequence on mRNA and protein expression, are uncertain. There are currently no clear guidelines for labelling "exon 1 deletions" as true *IKZF1* mutations. Understanding the molecular nature and impact of *IKZF1* 5' deletions is therefore important.

To determine if the exon 1 deletions (Δ ex1) designated by MLPA are bona-fide deletions, exon 1 genomic (g)DNA level was assessed for 13 MLPA-detected Δ ex1 BCP-ALL samples by quantitative polymerase chain reaction (qPCR) (Table S1). See Appendix S1 for Material and Methods. The 5' and 3' ends of exon 1 were analysed (Fig 1B). DNA from healthy donors (2 copies of *IKZF1*) and BCP-ALL samples with monosomy 7 or an *IKZF1* monoallelic deletion (both with one copy of *IKZF1*), were used as controls (Fig 1C). Nine of 13 (70%) MLPA-detected Δ ex1 samples displayed exon 1 levels at half those of healthy donors, similar to samples with single *IKZF1* copies. Similar results were obtained using two reference genes (*GAPDH*, *HBB*), confirming that exon 1 was deleted in the majority, though not all, MLPA-detected samples.

To map the genomic breakpoints of the 5' deletions, we screened gDNA from the 9 PCR-verified samples, from -500 kb to +14 kb of exon 1 by qPCR (Fig 1D and Table SII). Eight patients had 3' breakpoints near +1.5 kb, and 1 near +5 kb, suggesting a shared mechanism. In contrast, the 5' breakpoints were variable and mapped near

-15 kb, -29 kb, -79 kb, -110 kb, -115 kb, -120 kb and -165 kb. One deletion extended beyond -500 kb. We sequenced the breakpoint region of 3 patients, and found VDJ recombination signal sequences at both the 5' and 3' breakpoints, indicating a RAG enzyme-mediated mechanism (Fig 1E). Indeed, Patients 2 and 4 had nearly identical 3' breakpoints. These results suggested that large 5' deletions occur upstream of the *IKZF1* coding region, probably during VDJ recombination of *IGH* in B-cell precursors.

To investigate their impact of 5' deletions on protein expression, IKAROS levels were assessed by intracellular flow cytometry in 5 Δ ex1 samples (Fig 2A). Patient cells with normal *IKZF1* status were used as controls. IKAROS levels were reduced in all samples with a monoallelic *IKZF1* 5' deletion. Both alleles were mutated (a complete deletion and a 5' deletion) in Patient 6; here, IKAROS expression was completely lost in most cells, though some cells retained intermediate IKAROS levels, suggesting clonal heterogeneity. We also evaluated IKAROS expression by Western blot (Fig 2B). Full-length IK1 and the smaller IK2 isoforms were both reduced in the tested samples compared with control, and no novel isoforms were detected, indicating that 5' deletions result in loss of IKAROS protein expression.

To determine if the 5' deletions affect mRNA expression, *IKZF1* mRNA levels from 7 Δ ex1-only patients were compared with those of patients with different *IKZF1* status (wildtype [WT], mono- and biallelic deletions) by reverse transcription-qPCR, using primers to amplify exons 4-5 (Fig 2C). The 5' deletions were associated with significant reduction in *IKZF1* mRNA, similar to that observed in other types of monoallelic *IKZF1* deletions. Conversely, 2 MLPA-detected Δ ex1 samples not validated by PCR (non- Δ ex1) showed WT *IKZF1* mRNA levels. To confirm that the 5' deletions lead to a loss of transcription from the affected allele, we took advantage of polymorphisms in the *IKZF1* gene. One common single nucleotide polymorphism (SNP) (rs61731355; C→A at position Chr7:50,400,069 within exon 8) was detected in the gDNA and mRNA (cDNA) of 2 control patients, showing that *IKZF1* is biallelically expressed in BCP-ALL cells (Fig 2D). Patient 4 showed a novel SNP in their gDNA (C→A at position Chr7:50,400,071, also within exon 8) (Fig 2D). This polymorphism was not detected in the patient mRNA, indicating that only one allele was transcriptionally active (Fig 2D). These results suggested that

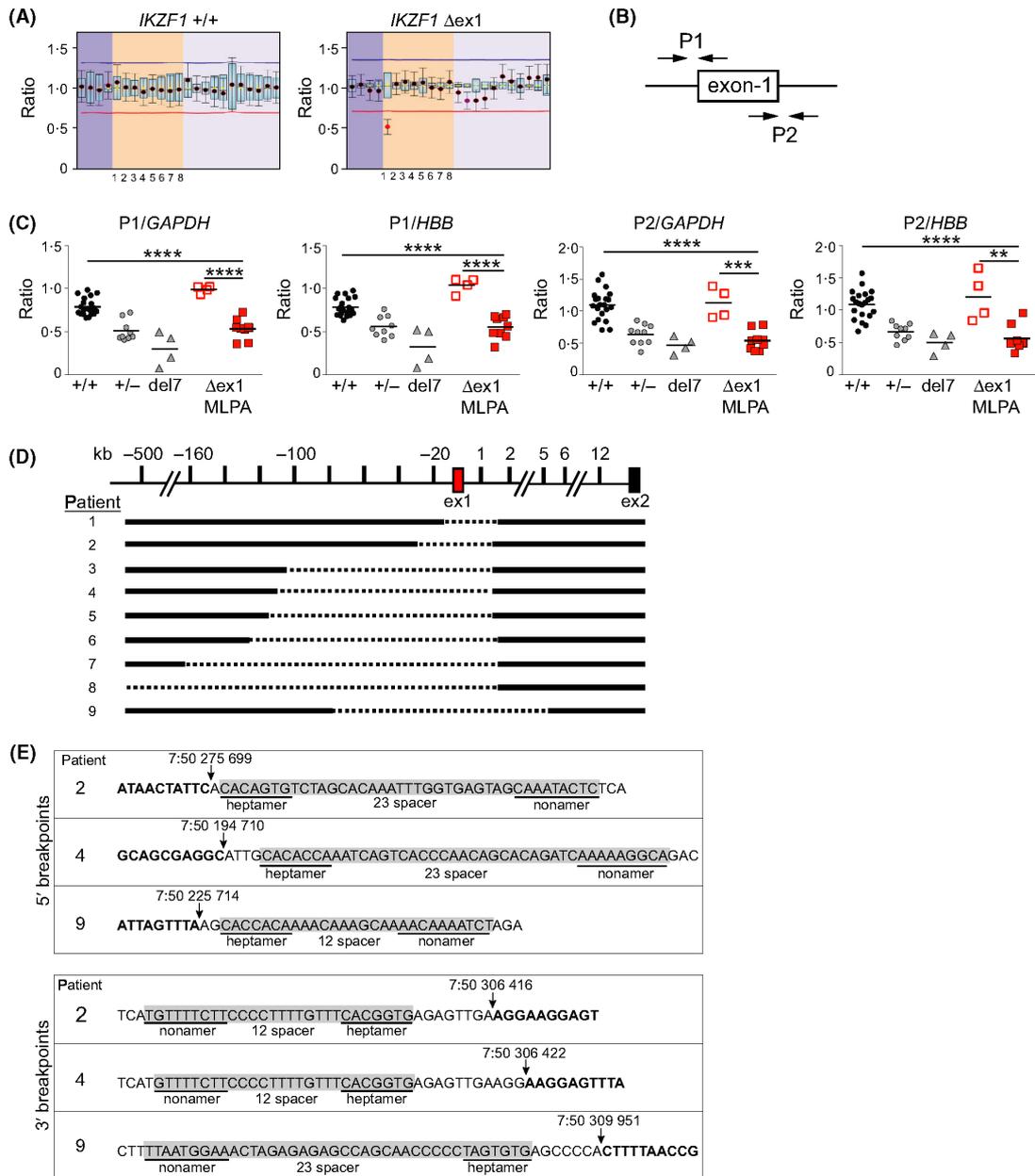


Fig 1. Analysis of *IKZF1* exon-1 deletions in BCP-ALL. (A) Example of multiplex ligation-dependent probe amplification (MLPA) analysis of *IKZF1* on the indicated samples. *IKZF1* exons are framed in orange; the bottom numbers denote the exons. Dots and bars represent the mean and standard deviation of relative copy numbers for each exon. Blue boxes represent the normal variation in a pool of 16 patient samples $+/+$ for the indicated exon. The copy number of each exon was calculated after normalization with internal reference genes. In the $\Delta ex1$ sample, the mean copy number of exon 1 is indicated as 0.5. (B) Schematic presentation of primer pairs used to analyse *IKZF1* exon by quantitative polymerase chain reaction (qPCR). (C) qPCR analysis of exon 1 in BCP-ALL samples. Ratios were defined between the indicated PCR primer pairs and the reference genes. $+/+$: blood DNA of healthy donors; $+/-$: monoallelic complete deletion of *IKZF1*; $del7$: monosomy 7; $\Delta ex1$ MLPA: patients found to be haploinsufficient for *IKZF1* exon 1 by MLPA. Statistical significance was analysed by unpaired Student's *t*-test. $**P = 0.0028$; $***P = 0.0005$; $****P < 0.0001$. (D) Mapping of *IKZF1* 5' deletions in patients with PCR-validated exon-1 deletion by qPCR (see Table SII for the qPCR data). The dotted lines represent deleted regions. (E) The 5' (top) and 3' (bottom) breakpoints of the indicated patients were analysed for the presence of recombination signal sequences (highlighted in grey). Arrows denote the position of the breakpoints. [Colour figure can be viewed at wileyonlinelibrary.com]

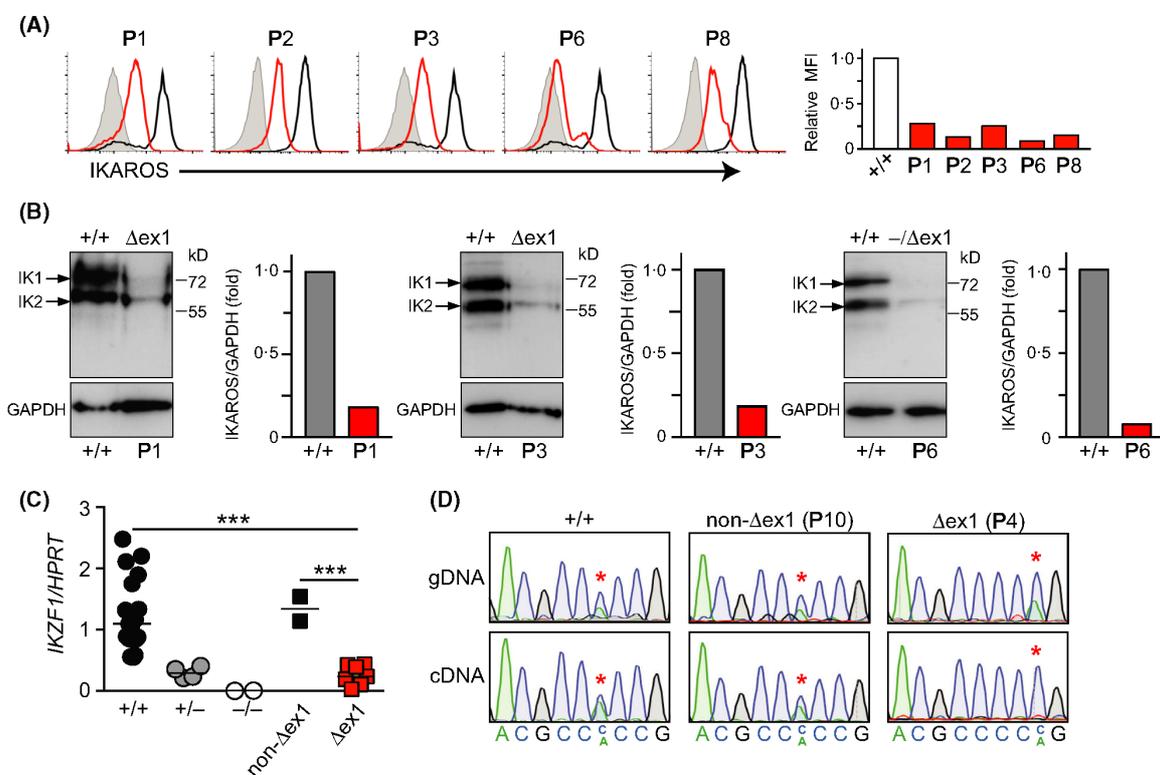


Fig 2. *IKZF1* 5' deletions result in loss of mRNA and protein expression. (A) IKAROS expression in cells of indicated patients with 5' deletions (red lines). As control, cells from a BCP-ALL patient with normal *IKZF1* status were used (black lines). Grey-filled histogram: secondary antibody staining only. Patient 6 (P6) had a complete deletion on one allele, and a 5' deletion on the other. Note that some leukaemic samples contained a minor population of non-leukaemic cells. The bar graph shows the quantification of the intracellular staining in terms of relative mean fluorescence intensity (MFI) of IKAROS expression. (B) Western blot analysis of IKAROS in cells from three patients with a 5' deletion. As control, cells from a BCP-ALL patient with normal *IKZF1* status were used. Arrows mark the major IKAROS isoforms IK1 and IK2. GAPDH was used as a loading control. Bar graphs show IKAROS levels normalized to *GAPDH* for each patient, where the normal *IKZF1* status was defined as 1.0. (C) Reverse transcription quantitative polymerase chain reaction (RT-qPCR) analysis of *IKZF1* relative to *HPRT* mRNA. +/+ : normal *IKZF1* status; +/- : monoallelic complete deletion of *IKZF1*; -/- : biallelic complete deletion of *IKZF1*; non- Δ ex1 : patients tested as haploinsufficient for *IKZF1* exon 1 by multiplex ligation-dependent probe amplification (MLPA), but revealed as +/+ by qPCR (see Fig 1C); Δ ex1 : patients tested as haploinsufficient for *IKZF1* exon 1 by MLPA and qPCR. *** $P < 0.0008$ (unpaired Student's *t*-test). (D) single nucleotide polymorphism (SNP) analysis of *IKZF1* gDNA (upper panels) and cDNA (lower panels) from cells of indicated BCP-ALL patients. Polymorphisms are indicated by the red asterisks. [Colour figure can be viewed at wileyonlinelibrary.com]

deletion of the 5' sequences transcriptionally inactivates the *IKZF1* gene in Δ ex1 patients.

Ascertaining the frequency of exon 1 deletions in BCP-ALL is complicated by the high G/C content in this region, which results in false-positive results using routine techniques. Interestingly, by improving the MLPA protocol (i.e., using a denaturation step of 20 min instead of 5, and controlling dubious cases with MLPA kits that contain more 5' probes), we reduced the frequency of exon 1 deletions from 1.5% to 0.5% [data from patients from the EORTC 58951 (1189 patients) and 58081 (388 patients) studies, respectively]. A higher frequency was reported with SNP microarrays (3%; data from the P9906 cohort (221 patients) (Mullighan *et al*, 2009), although it is unclear how the 5' deletions were defined in this study. A thorough assessment

of the frequency of 5' deletions in adult and paediatric BCP-ALL is therefore warranted.

In conclusion, our results indicate that *IKZF1* 5' deletions are recurrent alterations. They comprise large stretches of DNA that vary in length by hundreds of kb, and end at their 3' end near Chr7:50,306,420 in most cases. We further show that the mutant alleles lead to loss of *IKZF1* mRNA and protein. *IKZF1* mutations and deletions are often associated with poor prognosis in BCP-ALL (Mullighan *et al*, 2009; Clappier *et al*, 2015), and it will be important to determine if the present 5' deletions negatively impact prognosis. Our study, however, aimed at characterizing exon 1 deletions in molecular terms, and our cohort was too small and heterogeneous (Table SI) to correlate with clinical outcome. Nonetheless, since the 5' deletions are null alleles, like Δ ex2-8 and entire

Correspondence

IKZF1 deletions, we suggest that they be considered as such during risk stratification (Yeoh *et al*, 2018).

Acknowledgements

This work was supported by grants from the Conférence de Coordination Inter-Régionale du Grand-Est of the Ligue Contre le Cancer to B. Heizmann (01Q-2018), the Institut National du Cancer to P. Kastner and H. Cavé (2015-114), and by funds from ANR-10-LABX-0030-INRT. G. Morel received an “Année Recherche” fellowship from the Faculté de Médecine of the Université de Strasbourg. M.-C. Deau received an IGBMC International PhD Programme fellowship (from ANR-10-LABX-0030-INRT). C. Simand received a PhD fellowship from the Ligue Contre le Cancer.

Author contribution

G.M. and M.C.D. performed experiments, interpreted the data and revised the manuscript; C.S. collected and organized the patients’ data; A.CE., C.A., A.I. and L.Mi. prepared and analyzed patients’ samples at diagnosis; L.Ma., R.H., C.P. and M.S. provided patients’ samples and clinical data; H.C. provided patients’ samples and clinical data and contributed to the study with fruitful discussions. S.C. and P.K. designed the research, interpreted the data and wrote the manuscript; B.H. performed experiments, designed and supervised the research, interpreted the data and wrote the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

Guillaume Morel^{1,2,3,4,5,*} 
Marie-Céline Deau^{1,2,3,4,*} 
Célestine Simand^{1,2,3,4,5} 
Auréli Caye-Eude^{6,7}
Chloé Arfeuille^{6,7}
Antoine Ittel^{8,†}
Laurent Miguet^{8,9,10}
Laurent Mauvieux^{8,9,10}

References

Caye, A., Beldjord, K., Mass-Malo, K., Drunat, S., Soulier, J., Gandemer, V., Baruchel, A., Bertrand, Y., Cavé, H. & Clappier, E. (2013) Breakpoint-specific multiplex polymerase chain reaction allows the detection of *IKZF1* intragenic deletions and minimal residual disease monitoring in B-cell precursor acute lymphoblastic leukemia. *Haematologica*, **98**, 597–601.

Clappier, E., Grardel, N., Baikus, M., Rapon, J., De Moerloose, B., Kastner, P., Caye, A., Vivent, J., Costa, V., Ferster, A., Lutz, P., Mazingue, F., Millot, F., Plantaz, D., Plat, G., Plouvier, E.,

Poirée, M., Sirvent, N., Uytbroeck, A., Yakouben, K., Yakouben, K., Girard, S., Dastugue, N., Suciu, S., Benoit, Y., Bertrand, Y. & Cavé, H. (2015) *IKZF1* deletion is an independent prognostic marker in childhood B-cell precursor acute lymphoblastic leukemia, and distinguishes patients benefiting from pulses during maintenance therapy: results of the EORTC Children’s Leukemia Group study 58951. *Leukemia*, **29**, 2154–2161.

Dupuis, A., Gaub, M.P., Legrain, M., Drenou, B., Mauvieux, L., Lutz, P., Herbrecht, R., Chan, S. & Kastner, P. (2013) Biclinal and biallelic deletions occur in 29% of B-ALL cases with *IKZF1* mutations. *Leukemia*, **27**, 503–507.

Iacobucci, I., Storlazzi, C.T., Cilloni, D., Lonetti, A., Ottaviani, E., Soverini, S., Astolfi, A., Chiaretti, S., Vitale, A., Messa, F., Impera, L., Baldazzi, C., Addabbo, P.D., Papayannidis, C., Lonoce, A., Colarossi, S., Vignetti, M., Piccaluga, P.P., Paolini, S., Russo, D., Pane, F., Saglio, G., Bacarani, M., Foà, R. & Martinelli, G. (2009) Identification and molecular characterization of recurrent genomic deletions on 7p12 in the *IKZF1* gene in a large cohort of BCR-ABL1 – positive acute lymphoblastic leukemia patients. *Hematology*, **114**, 2159–2167.

Iacobucci, I., Iraci, N., Messina, M., Lonetti, A., Chiaretti, S., Valli, E., Ferrari, A., Papayannidis, C., Paolini, F., Vitale, A., Storlazzi, C.T.,

Raoul Herbrecht^{5,9,10}
Catherine Paillard^{10,11,12}
Marion Strullu^{7,13}
Hélène Cavé^{6,7}
Susan Chan^{1,2,3,4}
Philippe Kastner^{1,2,3,4,10}
Beate Heizmann^{1,2,3,4} 

¹Institut de Génétique et de Biologie Moléculaire et Cellulaire, ²Centre National de la Recherche Scientifique, UMR7104, ³Institut National de la Santé et de la Recherche Médicale, U964, ⁴Université de Strasbourg, Illkirch, ⁵Département d’Oncologie et d’Hématologie, CHU Strasbourg, Hôpital de Hautepierre, Strasbourg, ⁶Département de Génétique, Assistance Publique des Hôpitaux de Paris (AP-HP), Hôpital Robert Debré, ⁷INSERM UMR_S1131, Institut Universitaire d’Hématologie, Université Paris Diderot, Paris-Sorbonne-Cité, Paris, ⁸Laboratoire d’Hématologie, CHU Strasbourg, Hôpital de Hautepierre, ⁹Université de Strasbourg, INSERM, UMR-S1113/IRFAC, ¹⁰Faculté de Médecine, Université de Strasbourg, ¹¹Service d’Hémo-Oncologie Pédiatrique, CHU Strasbourg, Hôpital de Hautepierre, ¹²Université de Strasbourg, INSERM UMR_S1109, Strasbourg and ¹³Assistance Publique des Hôpitaux de Paris AP-HP, Hôpital Robert Debré, Service d’Hématologie Pédiatrique, Paris, France

[†]Present address: Institut Paoli-Calmettes, Marseille, France.

E-mail: heizmann@igbmc.fr

*These authors contributed equally.

Keywords: acute leukaemia, IKAROS, haploinsufficiency, 5’ deletions

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Patients information and *IKZF1* exon-1 status.

Table S2. qPCR analysis of *IKZF1* Dex1-patients to map the 5’ and 3’ region deletions.

Table S3. Primers.

Appendix S1. Materials and methods.

- Ottaviani, E., Guadagnuolo, V., Durante, S., Vignetti, M., Soverini, S., Pane, F., Foà, R., Baccharani, M., Mischen, M., Perini, G. & Martinelli, G. (2012) IKAROS deletions dictate a unique gene expression signature in patients with adult B-cell acute lymphoblastic Leukemia. *PLoS One*, **7**, e40934.
- Marke, R., Van Leeuwen, F.N. & Scheijen, B. (2018) The many faces of IKZF1 in B-cell precursor acute lymphoblastic leukemia. *Haematologica*, **103**, 565–574.
- Mullighan, C.G., Miller, C.B., Radtke, I., Phillips, L.A., Dalton, J., Ma, J., White, D., Hughes, T.P., Le Beau, M.M., Pui, C.H., Relling, M.V., Shurtleff, S.A. & Downing, J.R. (2008) BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature*, **453**, 110–114.
- Mullighan, C., Su, X., Zhang, J., Radtke, I., Letha, P., Christopher, M., Ma, J., Liu, W., Devidas, M., Gerhard, D.S., Yang, W. & Mary, V. (2009) Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *New England Journal of Medicine*, **360**, 470–480.
- Stanulla, M., Dagdan, E., Zaliova, M., Anja, M., Palmi, C., Cazzaniga, G., Kronnie, G., Bourquin, J., Bornhauser, B., Kochler, R., Bartram, C.R., Bleckmann, K., Groeneveld-krentz, S., Schewe, D., Junk, S.V., Hinze, L., Kratz, C.P., Biondi, A., Borkhardt, A., Kulozik, A., Muckenthaler, M.U., Basso, G., Valsecchi, M.G., Izraeli, S., Petersen, B.S., Franke, A., Dörge, P., Steinemann, D., Haas, O.A., Panzer-Grümayer, R. & Cavé, H. (2018) IKZF1 plus defines a new minimal residual disease – dependent very-poor prognostic profile in pediatric B-cell precursor acute lymphoblastic. *Leukemia*, **36**, 1240–1249.
- Yeoh, A.E.J., Lu, Y., Chin, W.H.N., Chiew, E.K.H., Lim, E.H., Li, Z., Kham, S.K.Y., Chan, Y.H., Abdullah, W.A., Lin, H.P., Chan, L.L., Lam, J.C.M., Tan, P.L., Quah, T.C., Tan, A.M. & Ariffin, H. (2018) Intensifying treatment of childhood B-lymphoblastic leukemia with *IKZF1* deletion reduces relapse and improves overall survival: results of Malaysia-Singapore ALL 2010 study. *Journal of Clinical Oncology*, **36**, 2726–2735.

Helios inhibits megakaryopoiesis and supports lymphoid priming via chromatin compaction in hematopoietic stem and progenitor cells

Giovanni Cova^{1,2,3,4*}, Marie-Céline Deau^{1,2,3,4}, Qi Cai^{1,2,3,4}, Vincent Mittelheisser^{1,2,3,4}, Matthieu Jung^{1,2,3,4,5}, Marie Cerciat^{1,2,3,4,5}, Christelle Thibault-Carpentier^{1,2,3,4,5}, Bernard Jost^{1,2,3,4,5}, Leif Carlsson⁶, Angela M. Thornton⁷, Ethan M. Shevach⁷, Peggy Kirstetter^{1,2,3,4*#}, Philippe Kastner^{1,2,3,4,8#} and Susan Chan^{1,2,3,4,9#}

* Equal contribution

Senior author

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

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

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

⁴ Université de Strasbourg, ILLKIRCH, France

⁵ Plateforme GenomEast, Infrastructure France Génomique, ILLKIRCH, France

⁶ Umeå Center for Molecular Medicine, Umeå University, Umeå, Sweden

⁷ Laboratory of Immune System Biology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA

⁸ Faculté de Médecine, Université de Strasbourg, Strasbourg, France

⁹ Lead contact

Correspondence:

Peggy Kirstetter (kirstett@igbmc.fr)

Philippe Kastner and Susan Chan (scpk@igbmc.fr)

ABSTRACT

Our understanding of cell fate decisions in hematopoiesis remains incomplete. Here we show that the transcription factor Helios (*Irf2*) is required to control megakaryopoiesis in the murine bone marrow, where it is highly expressed in all hematopoietic stem and progenitor cells (HSPCs). Helios null HSPCs are phenotypically and functionally biased towards the megakaryocyte lineage at the expense of the lymphoid, and resemble HSPCs of aging animals. Helios acts mainly in quiescent HSPCs, where it represses the megakaryocyte gene expression program in cells as early as the hematopoietic stem cell stage. Helios promotes chromatin compaction at target genes involved in multiple aspects of HSPC biology, including cell adhesion and migration. Notably, Helios limits chromatin accessibility at regulatory regions of platelet-specific genes also targeted by the Gata2 and Runx1 transcriptional activators. Thus, Helios is a critical negative regulator of the megakaryocyte fate and allows lymphoid priming in HSPCs.

INTRODUCTION

Hematopoietic stem and progenitor cells (HSPCs) sit atop the hierarchy that gives rise to all blood cells (Wang and Ema, 2016). HSPCs are mostly quiescent or slow-proliferating, and currently comprise pluripotent long- (LT) and short-term (ST) hematopoietic stem cells (HSC) and multipotent progenitor (MPP) cells biased towards the megakaryocyte/erythroid (MPP2), myeloid (MPP3) and lymphoid (MPP4) fates (Oguro et al., 2013; Pietras et al., 2015). LT-HSCs are the only subset with lifelong repopulation potential (Oguro et al., 2013; Bernitz et al., 2016). ST-HSCs and MPPs have a lower capacity for self-renewal even though they contribute extensively to steady state hematopoiesis (Rodriguez-Fraticelli et al., 2018). HSPCs are heterogeneous and our understanding of their relationships, and the factors controlling their development, are constantly evolving.

Indeed, only about half of the LT-HSCs are now thought to be truly pluripotent (Carrelha et al., 2018). Among the HSCs with more restricted potential, a significant number express the platelet glycoprotein heterodimeric complex CD41 (encoded by the *Itga2b* gene) and CD61 (encoded by *Itgb3*), indicative of megakaryocyte precursors (Gekas and Graf, 2013; Sanjuan-Pla et al., 2013; Grover et al., 2016). The centrality of the megakaryocyte lineage was uncovered by cell tracking experiments, which showed that HSCs generate more megakaryocytes than myeloid or lymphoid cells, which rise from MPP cells under homeostatic conditions (Busch et al., 2015; Rodriguez-Fraticelli et al., 2018). Further, genome-wide studies revealed that megakaryocytes and HSCs have overlapping active chromatin signatures, suggesting that targeted chromatin remodeling may initiate the differentiation from HSC to megakaryocyte (Heuston et al., 2018). The molecular events that govern megakaryocyte lineage priming and cell fate restriction remain partially understood.

Key transcription factors driving lineage commitment in HSPCs have been identified. The Gata factors (1 and 2) and Runx1 are well-characterized activators of the megakaryocyte/erythroid fate, while Pu.1 is required for differentiation towards the myeloid and lymphoid lineages (Huang et al., 2009; Ikonomi et al., 2000; Kawaguchi et al., 2016; Pencovich et al., 2011; Zang et al., 2016; Draper et al., 2017; Pang et al., 2018). In contrast, negative regulators of cell fate decisions are less well understood, in particular those involved in megakaryopoiesis. Helios, encoded by the *Ikzf2* gene, is a member of the Ikaros family of transcription factors. It has been documented to be a major player in regulatory T cell development and function (Getnet et al., 2010; Thornton et al., 2010; Kim et al., 2015; Sebastian et al., 2016; Skadow et al., 2019; Thornton and Shevach, 2019). Yet Helios is also highly expressed in HSPCs at the mRNA level, suggesting that it may be important for HSPC function. Here we show that Helios is a negative regulator of megakaryocyte priming

in hematopoietic stem and progenitor cells, and biases the lymphoid to platelet fate through gene repression.

RESULTS

Helios is highly expressed in hematopoietic stem and progenitor cells

To identify the cells expressing Helios in the murine bone marrow (BM), we evaluated the HSPC populations (CD41⁺ and CD41⁻ LT-HSC, ST-HSC, MPP2-4), committed progenitor cells (GMP, MkP, EryP, CLP) and mature cell lineages (lymphoid, myeloid, erythroid) for Helios proteins, by flow cytometry and intracellular staining (**Fig. 1**). Our experiments showed that >99% of HSPCs and ~80% of committed progenitor cells expressed high levels of Helios (**Fig. 1, a, b and e**), suggesting that Helios is active in early hematopoietic cell stages. These results were somewhat unexpected, given that GMPs, MEPs and CLPs express low *Ikzf2* mRNA levels (Drissen et al., 2016; ImmGen Consortium, 2016). Helios proteins were not observed in immature B cells, myeloid cells and erythrocytes (**Fig. 1, c and e**), though they were easily detected in a portion of T cells, especially CD4⁺Foxp3⁺ regulatory T (Treg) cells (**Fig. 1, d and e**), as published (Sebastian et al., 2016). Indeed, Treg cells expressed the highest levels of Helios in the BM.

To determine if Helios may also play a role in non-hematopoietic BM cells, we analyzed its expression pattern in endothelial cells (EC) and mesenchymal stem cells (MSC), the two more abundant populations implicated in HSPC maintenance and differentiation (**Fig. 1f**). ECs and MSCs did not express detectable Helios proteins. We concluded that Helios is selectively and abundantly expressed in the HSPCs, committed progenitor cells and Treg cells of the BM.

Helios skews hematopoiesis towards the lymphoid and away from the megakaryocyte/myeloid lineages

To investigate if Helios influences hematopoietic cell homeostasis, we analyzed the proportions of HSPCs, committed progenitors and mature lineages in the BM of adult WT and germline Helios null (KO) mice at different ages, by flow cytometry (**Figs. 2 and S1**) (Cai et al., 2009). Our results showed that, while the BM cellularity was relatively similar between genotypes (**Fig. S1a**), multiple populations were affected in the absence of Helios. The most significant decreases in the KO BM were observed among the lymphoid-biased MPP4 (CD150⁻CD48⁺Flt3⁺ LSK) and CLP (IL7R⁺Flt3⁺ Lin⁻cKit^{med}Sca1^{med}) cells (**Fig. 2, a and b**), at all ages tested (6, 10 and 20 weeks). Immature B cells (CD19⁺B220⁺) were also decreased compared with WT (**Fig. S1b**). In contrast, the most significant increases were seen in the megakaryocyte-biased CD41⁺ LT-HSC (CD150⁺CD48⁻ LSK), CD41⁺ MPP2 (CD150⁺CD48⁺ LSK) and megakaryocyte progenitor MkP (CD150⁺CD41⁺ LK) cells (**Fig. 2, d-f**), at all ages evaluated. Myeloid-biased populations (MMP3, GMP), as well as BM

neutrophils were increased (**Figs. 2, a and c, and S1c**), though the results were more variable and less significant. The proportions of erythroid lineage (MPP2, EryP, erythrocytes) and T cells were not overtly affected by Helios loss (**Figs. 2, e and g, and S1, d and e**). These results suggested that hematopoiesis is deregulated in the absence of Helios.

To determine if the surface marker phenotype correlated with functional alterations in the KO progenitor cells, we evaluated the ability of WT and KO BM cells to give rise to lymphoid, myeloid and megakaryocyte-containing colonies in vitro (**Fig. 3**). Colony forming unit (CFU) assays were performed with total BM cells cultured in semi-solid medium supplemented with IL7 (for CFU-preB) or SCF, TPO, EPO, IL3 and IL11 (for CFU-Myeloid-Mk) for 7 days (**Fig. 3a**). Our results showed that KO BM cells gave rise to significantly less CFU-preB, and a higher ratio of CFU-Mk, compared with WT (**Fig. 3, b and c**). To dissect the cellular source of this differentiation bias, we cultured single LT-HSCs (CD150⁺CD48⁻LSK) in the presence of SCF, TPO and IL3 for 10 days to induce the formation of Mk⁺ (megakaryocyte-only or Mk-myeloid mix) and myeloid-only colonies (**Fig. 3d**). These experiments showed that KO LT-HSCs gave rise to more Mk⁺ colonies, and less myeloid-only colonies, than WT HSCs. We also evaluated the potential of single MPP (MPP3+4; CD150⁻CD48⁺ LSK) cells to form myeloid colonies after 7 days in culture with SCF, EPO, IL3 and IL6. In these cultures, KO cells gave rise to significantly more myeloid colonies (**Fig. 3e**).

Collectively, our results revealed that Helios is required to promote lymphopoiesis and inhibit megakaryopoiesis in HSPCs.

Helios regulates HSPC differentiation in a hematopoietic cell-intrinsic manner

To determine if Helios affects hematopoiesis in a cell-intrinsic or -extrinsic manner, we tested the ability of BM MPPs and LT-HSCs to repopulate the hematopoietic system after adoptive transfer into irradiated recipient mice. To evaluate MPP function, 5×10^3 MPP3 or MPP4 cells from WT or Helios KO mice (CD45.2⁺) were intravenously injected into sublethally-irradiated recipients (CD45.1⁺CD45.2⁺), and blood was analyzed after 2 weeks for CD45.2⁺ lymphoid (B220⁺) and myeloid (CD11b⁺) cells, by flow cytometry (**Fig. 4a**). These results showed that, when MPP3 cells were injected, both WT and KO genotypes gave rise almost exclusively to myeloid cells. However, when MPP4 cells were injected, WT cells were biased towards the lymphoid lineage, as expected, but KO cells were similarly divided into lymphoid and myeloid cells, suggesting a loss of lymphoid and/or gain of myeloid potential without Helios.

To evaluate LT-HSC function, we injected 100 LT-HSCs from WT or Helios KO mice (CD45.2⁺), along with 5x10⁵ supporting BM cells (CD45.1⁺) into lethally-irradiated mice (CD45.1⁺CD45.2⁺) (**Fig. 4b**). Blood cells were analyzed after 2 months for CD45.2⁺ B220⁺ B, CD3⁺ T, and CD11b⁺ myeloid cells. We observed a significant decrease of B cells, and a corresponding increase of myeloid cells, in recipients that received KO HSCs compared with WT. The proportion of T cells among all recipients remained similar. These results showed that both LT-HSC and MPP cells give rise to proportionately less B cells and more myeloid cells in the absence of Helios, and that the effects are hematopoietic cell-intrinsic.

Helios regulates HSPC function independently of T cells and external factors

In the course of our experiments, we observed that the BM of Helios KO mice contained less CD4⁺CD25⁺Foxp3⁺ Treg cells, and more CD4⁺Foxp3⁻ T cells producing IFN γ , IL-2 and IL-10 than WT animals (**Fig. S2, a and b**), suggesting a chronic inflammation of the KO BM environment. Moreover, the LSK pool was increased in 20-week-old KO mice, a feature often observed during inflammatory responses (**Fig. S2c**) (Mirantes et al., 2014; Schürch et al., 2014; Haas et al., 2015). As inflammation and Treg deficiency may impact HSPC function, we studied hematopoiesis in *Ikzf2^{fl/fl}* x CD4-Cre⁺ (TKO) mice (**Fig. S3a**), where Helios is specifically deleted in T cells (Sebastian et al., 2016; Pierini et al., 2017; Hirata et al., 2018). We reasoned that if Helios-expressing T cells were responsible for the HSPC bias in lymphoid versus megakaryocyte/myeloid potential, then we would expect to see a similar bias when Helios is deleted only in T cells. On the other hand, if the hematopoietic defect is intrinsic to HSPCs, then we would expect to see no such defects in hematopoiesis. Indeed this was the case. Phenotypic analyses showed that WT and TKO mice had similar percentages of cells in all hematopoietic populations tested. Thus, Helios-proficient HSPCs are not affected by the loss of Helios in T cells.

To determine if HSPCs are influenced by external signals from other hematopoietic cells, we performed competitive repopulation experiments between WT and KO cells. Lethally-irradiated recipient mice (CD45.1⁺CD45.2⁺) were reconstituted with WT or KO BM cells (CD45.2⁺), and competitor WT BM cells (CD45.1⁺), at a 1:1 ratio (**Fig. S3b**). After 2 months, CD45.2⁺ and CD45.1⁺ blood cells were analyzed for B220⁺ B and CD11b⁺ myeloid cells. These results showed that KO donor cells (in red) exhibited the same loss of B cell potential, and gain of myeloid lineage potential, even in the presence of WT BM cells. On the other hand, the WT competitor HSPCs retained their normal potential whether they developed in the presence of KO or WT cells (compare the gray datapoints between the left and right graphs).

Altogether, our results provide strong evidence that Helios regulates the developmental potential of HSPCs in a cell-autonomous manner.

Helios influences gene expression in all HSPC populations

To determine how gene expression is regulated by Helios, we analyzed the global gene expression changes in WT and Helios KO LT-HSCs (CD150⁺CD48⁻ LSK), MPP3 (CD150⁻CD48⁺Flt3⁻ LSK) and MPP4 (CD150⁻CD48⁺Flt3⁺ LSK) cells by RNA-sequencing (**Fig. 5**). Differentially expressed genes were identified by DEseq2, and the genes with a p-value of <0.05 and a log2 fold-change (log2FC) of >0.56 (1.5-fold) between KO and WT samples were subjected to hierarchical clustering (**Fig. 5a and Table S1**). Our results showed that gene expression changes were detected in all populations analyzed, with the highest number of changes found between WT and KO LT-HSCs. Gene set enrichment analyses (GSEA) indicated that genes associated with a MkP signature (eg. *Selp*, *F2r12*, *Tbxa2r*, *Cd9*, *Vwf*, *Itgb3*, *Plek*) were highly enriched amongst the up-regulated genes in the KO LT-HSCs compared with WT, while those associated with a MPP4 signature (eg. *Itgax*, *Cd52*, *Itgam*, *Ccr7*, *Clec12a*) were markedly diminished (**Fig. 5, a and b**). Similar biases were observed in the MPP3 and MPP4 populations, where the MkP signature was enriched, and the MPP4 and CLP signatures were reduced (**Fig. 5c**). Interestingly, the KO LT-HSC profile resembled that of HSCs from old mice. Up-regulated genes included those related to megakaryopoiesis, but also non-Mk-associated genes like *Gprc5c*, *Fzd1*, *Il17rs*, *Clu* and *Plek2*, amongst others; down-regulated genes included those related to lymphopoiesis, but also non-MPP4 genes like *Ms4a4b*, *Ccl6*, *Dock10* and *Mmp9* (**Fig. 5b and Table S1**). Metascape analysis of the gene expression changes in the LT-HSC compartment suggested increased activity in platelet activation, and reduced activity in cellular component movement, microtubule-based processes, and migration (**Fig. 5d**).

Because HSPCs are also characterized by other cell surface markers, we analyzed the transcriptomes of CD34⁻Flt3⁻ LT-HSCs and lymphoid-primed multipotent progenitors (LMPP: CD34⁺Flt3⁺ LSK) by microarray (**Fig. S4, a and b**). Hierarchical clustering and GSEA showed that CD34⁻ LT-HSCs exhibited the highest number of gene expression changes between WT and KO, with a strong enrichment of the MkP and aging HSC signatures in the KO, similar to CD150⁺CD48⁻ HSCs. Further we observed a concerted decrease in the CLP signature amongst the LMPP KO samples. These results suggested that Helios expression is less required in the HSPCs of old mice. To address this question, we compared the levels of Helios proteins in the LT-HSC, ST-HSC and MPP populations of young (10-week-old) and old (>1-year-old) WT mice, by flow cytometry (**Fig. S4c**); this

Altogether, our results provide strong evidence that Helios regulates the developmental potential of HSPCs in a cell-autonomous manner.

Helios influences gene expression in all HSPC populations

To determine how gene expression is regulated by Helios, we analyzed the global gene expression changes in WT and Helios KO LT-HSCs (CD150⁺CD48⁻ LSK), MPP3 (CD150⁻CD48⁺Flt3⁻ LSK) and MPP4 (CD150⁻CD48⁺Flt3⁺ LSK) cells by RNA-sequencing (**Fig. 5**). Differentially expressed genes were identified by DEseq2, and the genes with a p-value of <0.05 and a log2 fold-change (log2FC) of >0.56 (1.5-fold) between KO and WT samples were subjected to hierarchical clustering (**Fig. 5a and Table S1**). Our results showed that gene expression changes were detected in all populations analyzed, with the highest number of changes found between WT and KO LT-HSCs. Gene set enrichment analyses (GSEA) indicated that genes associated with a MkP signature (eg. *Selp*, *F2r12*, *Tbxa2r*, *Cd9*, *Vwf*, *Itgb3*, *Plek*) were highly enriched amongst the up-regulated genes in the KO LT-HSCs compared with WT, while those associated with a MPP4 signature (eg. *Itgax*, *Cd52*, *Itgam*, *Ccr7*, *Clec12a*) were markedly diminished (**Fig. 5, a and b**). Similar biases were observed in the MPP3 and MPP4 populations, where the MkP signature was enriched, and the MPP4 and CLP signatures were reduced (**Fig. 5c**). Interestingly, the KO LT-HSC profile resembled that of HSCs from old mice. Up-regulated genes included those related to megakaryopoiesis, but also non-Mk-associated genes like *Gprc5c*, *Fzd1*, *Il17rs*, *Clu* and *Plek2*, amongst others; down-regulated genes included those related to lymphopoiesis, but also non-MPP4 genes like *Ms4a4b*, *Ccl6*, *Dock10* and *Mmp9* (**Fig. 5b and Table S1**). Metascape analysis of the gene expression changes in the LT-HSC compartment suggested increased activity in platelet activation, and reduced activity in cellular component movement, microtubule-based processes, and migration (**Fig. 5d**).

Because HSPCs are also characterized by other cell surface markers, we analyzed the transcriptomes of CD34⁻Flt3⁻ LT-HSCs and lymphoid-primed multipotent progenitors (LMPP: CD34⁺Flt3⁺ LSK) by microarray (**Fig. S4, a and b**). Hierarchical clustering and GSEA showed that CD34⁻ LT-HSCs exhibited the highest number of gene expression changes between WT and KO, with a strong enrichment of the MkP and aging HSC signatures in the KO, similar to CD150⁺CD48⁻ HSCs. Further we observed a concerted decrease in the CLP signature amongst the LMPP KO samples. These results suggested that Helios expression is less required in the HSPCs of old mice. To address this question, we compared the levels of Helios proteins in the LT-HSC, ST-HSC and MPP populations of young (10-week-old) and old (>1-year-old) WT mice, by flow cytometry (**Fig. S4c**); this

revealed that Helios was significantly reduced in the LT-HSCs of old mice compared with young, while the ST-HSCs and MPP cells from animals of both age groups retained similar levels of Helios. Importantly, we also evaluated the expression of Ikaros (encoded by the *Ikzf1* gene) and Eos (*Ikzf4*) in Helios KO cells, as Ikaros has been implicated in HSC maintenance and lymphoid priming (Nichogiannopoulou et al., 1999; Ng et al., 2009), and Eos, also expressed in HSCs, is most homologous to Helios (Papathanasiou et al., 2009). *Ikzf1* mRNA levels were not significantly altered between WT and Helios KO HSPC populations, while *Ikzf4* levels were similarly low in WT and KO HSCs, and undetectable in MPP cells (**Fig. S4d**). Thus, the phenotypes observed were not due to changes in Ikaros or Eos expression.

Collectively, our results indicate that Helios negatively regulates age-associated gene expression, represses megakaryocyte commitment, and promotes lymphoid lineage priming in HSPCs.

Helios promotes cell fate choice in HSPCs at the single cell level

To better understand how the gene expression changes lead to the pro-megakaryocyte HSPC phenotype in the absence of Helios, we analyzed the LSK cells from WT and Helios KO mice (2 of each genotype) by single cell RNA-seq. Purified LSK cells from 10-week-old animals were each stained with anti-CD150 and anti-CD48 oligo-tagged Abs to track HSCs and MPP cells (CITE-seq), and then labeled with Hash-Tag-Oligo conjugated anti-CD45 and anti-MHC class I Abs to identify the mouse of origin (Stoeckius et al., 2017, 2018). 2500 cells from every sample were pooled (cell hashing), and subjected to 10X Genomics library preparation, which allowed the mRNA from single cells and their associated Ab-oligos to be tagged with a unique index. After library sequencing and further quality control filtering, these experiments allowed us to analyze the mRNA profiles of 3470 unique cells out of the starting 1×10^4 (**Figs. 6, S5 and S6**).

To visualize the organization of the HSPC compartment, we generated a 2D map of the 3470 cells with the Uniform Manifold Approximation and Projection (UMAP) tool, and evaluated the characteristics that assigned each cell to a particular area. Principal component analysis was used to dissect the gene expression signatures, and components 1, 2 and 3 (PC1-3) captured the biggest differences between the cells (**Fig. S5a**). PC1 and -2 were mainly defined by genes related to the cell cycle (**Fig. S5, b-e**). The cells identified by negative PC1 coordinates expressed transcripts for genes involved in DNA replication, and they clustered at the bottom of the UMAP plot (**Fig. S5, b and c**); on the other hand, PC1-positive genes were expressed by almost all cells. The cells identified by negative PC2

coordinates expressed mitosis-related genes (**Fig. S5, d and e**); PC2-positive genes were linked to pathways generally associated with the cell cycle, such as ribosome biogenesis and DNA integrity checkpoint. Thus, the vertical axis of the UMAP plot separated cells according to their proliferative status: quiescent cells (blue) were found at the top, while the cells at early (orange) and late (green) stages of the cell cycle clustered at the bottom (**Fig. 6a**).

PC3 genes were associated with lineage priming (**Figs. 6b and S6, a and b**). The cells identified by negative PC3 values expressed high mRNA levels of genes associated with HSCs and MPP2 cells (e.g. *Mpl*, *Gata2* and *Cd9*), according to our RNA-seq data and the ImmGen datasets (MPP2). These cells clustered to the left of the UMAP plot (**Figs. 6b and S6b**). By contrast, the cells identified by positive PC3 values expressed abundant transcripts for genes characteristic of lymphoid-primed MPP4 cells, like *Flt3*, *Dntt* and *Cd53*, and they clustered to the right of the UMAP plot. The cells in the middle of the plot expressed PC3 genes at intermediate levels when compared with HSC/MPP2 and MPP4 cells, and probably corresponded to MPP3 cells (**Fig. S6b**). Similar results were obtained when we compared the location of cells expressing CD150 (encoded by the *Slamf1* gene) and CD48 at the mRNA and protein levels (**Fig. S6c**), when the barcode indexes pertaining to the CITE-seq antibodies were integrated. CD150⁺CD48⁻ cells were located to the left (as were *Slamf1*⁺*Cd48*⁻ and *Cd34*⁺*Cd48*⁻ cells), and CD150⁻CD48⁺ (and *Slamf1*⁻*Cd48*⁺) cells (which included MPP3 and MPP4 cells) were mapped to the center and the right.

Interestingly, our analyses revealed that the Helios KO LSK population contained more HSC/MPP2 (*Flt3*⁺*Mpl*⁺) than MPP4 (*Flt3*⁺*Mpl*⁻) cells (**Fig. 6b**), while the opposite was true for the WT LSK population, in agreement with the KO megakaryocyte bias and reduced lymphoid potential observed in the RNA-seq experiments. To determine if these alterations involved both quiescent and proliferating cells, we generated heatmaps of WT and KO quiescent and cycling cells, by evaluating the expression of the top 30 PC3-negative and -positive genes, and ranking the cells according to their PC3 coordinates (**Fig. 6, c and d, left**). To quantify the lineage priming bias in each cell, a "lineage priming score" was calculated, based on the total number of expressed HSC/MPP2 (counted as -1 per gene) vs MPP4 (+1 per gene) genes. Cells that expressed a predominantly HSC/MPP2 gene expression signature had an overall negative score, while those that expressed an MPP4 signature had an overall positive score. Strikingly, Helios KO quiescent HSPCs were highly enriched for cells expressing HSC/MPP2 genes (**Fig. 6c, right**). A smaller but still significant result was also observed among cycling cells (**Fig. 6d, right**).

These results indicated that, at single cell resolution, Helios is required to promote lymphopoiesis, and inhibit megakaryopoiesis, particularly among quiescent HSPCs.

Helios targets HSPC genes through gene repression

To identify Helios target genes in HSPCs, we investigated its genomic binding sites in hematopoietic precursor cells, by ChIP-sequencing. We studied the hematopoietic precursor cell line HPC7 (Pinto do O et al., 1998), as we could not purify enough primary LSK cells for these experiments. HPC7 cells express Helios and comprise a mix of cells with LSK and LK phenotypes (**Fig. S7a**), which differentiate into multiple hematopoietic lineages in the presence of appropriate cytokines. Our results revealed 16,522 Helios-bound peaks which mapped near 8884 genes; binding was observed at gene bodies, distal regions and proximal promoters (**Fig. S7b**). Motif enrichment analyses identified GGAA-containing motifs as highly enriched at Helios-bound peaks (**Fig. S7c**), as expected for an Ikaros family member. Helios binding was detected at 65% of the genes implicated in lineage priming (from the PC3 genes listed in Fig. 6), and 56% of those highly expressed in WT HSPCs (**Fig. S7d**). However, when the impact of binding on gene expression was investigated, we found that Helios bound only 37% of the genes whose transcripts were up- or down-regulated in the Helios KO HSPCs. This value was just slightly higher than that observed for a random set of genes. Thus, Helios binding is widespread in HSPCs, but binding may not be predictive of gene regulation.

To better understand Helios function, we analyzed the effect of Helios deficiency on chromatin accessibility, by ATAC-sequencing of WT and Helios KO LSK cells. These experiments revealed $\sim 132 \times 10^3$ accessible chromatin regions that were mostly shared between WT and KO samples. When integrated with the HPC7 ChIP-seq data, we found that nearly all (15,121; 91%) of the sites recognized by Helios ("Helios sites") coincided with open chromatin. Strikingly, the chromatin regions with Helios sites were significantly more accessible in the KO samples compared with WT, which was not true for chromatin regions with no Helios sites (**Fig. 7a, left**). In addition, we evaluated the importance of Helios sites in the chromatin regions that became more open (red) or closed (black) upon Helios deficiency, and studied the ATAC-seq peaks that were strongly ($FC > 0.3$) or weakly ($FC < 0.3$) affected (**Fig. 7b**). In both cases, Helios sites were frequently and selectively found in the chromatin regions that were opened in the KO cells. These results indicated that Helios deficiency is associated with increased chromatin accessibility at regions where Helios targets DNA. Thus, Helios binding is broadly correlated with chromatin compaction.

To elucidate the biological processes linked to Helios binding, we performed metascape analyses on the regions with the biggest changes in ATAC-seq signals (top 30% up or down, p-value <0.05) in the Helios KO cells (**Fig. 7c**). This revealed pathways related to cell migration, cell adhesion, signaling, cell quiescence and platelet function, many of which we previously identified by RNA-seq (**Fig. 5d**). The identification of platelet-related pathways suggested a critical role for Helios in the repression of genes important for megakaryopoiesis at the epigenetic level. Increased chromatin accessibility upon Helios loss was clearly visible at MkP-associated genes with Helios binding sites (**Fig. 7, a and d**). In contrast, MPP4-related genes were minimally affected by Helios deficiency: they exhibited similar ATAC-seq patterns in WT and KO cells (**Fig. S7e**), and lymphoid-related terms did not belong to the processes identified for genes with decreased chromatin accessibility (**Fig. S7f**).

Interestingly, the Helios binding regions identified above were also enriched for Gata2 and Runx1 motifs (**Fig. 7e**). Since the Gata2 and Runx1 binding profiles were previously mapped in HPC7 cells (Wilson et al., 2010), we integrated these data with our Helios ChIP-seq results. This showed that the Helios, Gata2 and Runx1 binding regions consistently co-localized with chromatin regions of higher accessibility in Helios KO LSK cells, and notably at genes associated with the megakaryocyte lineage (*Itgb3*, *Cd9*, *F2rl2*, *Gp1ba*, *Grb2*, *Itga2b*) (**Fig. 7, d and f**), suggesting transcriptional coordination among these factors.

Our results thus indicate that Helios promotes widespread chromatin compaction at target genes important for HSPC commitment and function, with potentially negative consequences for megakaryocyte lineage priming by Gata2 and Runx1.

DISCUSSION

Here we demonstrate that Helios is a critical regulator of cell fate choice in early hematopoiesis. We show that Helios is required in a cell-intrinsic manner in hematopoietic stem and progenitor cells to limit megakaryocyte commitment and promote lymphocyte priming. We further show that Helios negatively regulates the size of the megakaryocyte-biased HSPC pool, by restricting chromatin accessibility at key platelet-specific genes to repress their mRNA expression. Indeed, Helios is associated with closed chromatin across most its target gene repertoire where chromatin remodeling occurs. Interestingly, a high number of Helios target genes encode proteins important for cell-cell adhesion and interaction, like integrins and selectins, as well as proteins involved in cell movement and migration. Thus, Helios may contribute to essential developmental processes by modulating the interactions between HSPCs and the BM microenvironment.

Helios null HSPCs resemble both cells committed to the megakaryocyte lineage as well as HSPCs from old mice, suggesting a common mechanism. Given that HSC aging is regulated at least in part by BM inflammation (Pietras, 2017; Ho et al., 2019), and that Helios KO mice have more IFN γ -expressing T cells in the BM, we initially tested the possibility that the mutant T cells might be responsible for the HSPC phenotype. However, Helios-specific deletion in T cells did not increase megakaryopoiesis nor decrease lymphopoiesis *in vivo*, and competitive BM transplantation experiments clearly revealed a strong cell-intrinsic component to the KO HSPC bias. As Helios levels are reduced in LT-HSCs with age, we hypothesize that loss of Helios may also favor the switch from lymphopoiesis to megakaryo- and myelopoiesis in the HSPCs of older animals.

Interestingly, our single cell RNA-seq experiment revealed the presence of quiescent HSPCs with HSC/MMP2, MPP3 and MPP4 features in both the WT and KO samples. We speculate that these cells, with stem cell-like properties, might represent the apex of the hematopoietic system. Helios removal affects mainly this pool, where it biases lineage priming towards the megakaryocyte fate. The gene expression changes of these few quiescent KO MPP cells would not have been detected amongst the larger proliferating MPP populations, which may explain why we did not find a tremendous difference between WT and KO MPP cells in the population RNA-seq data.

Helios binding is associated with reduced chromatin accessibility, suggesting that Helios is a transcriptional repressor in HSPCs. Helios-mediated recruitment of the Nucleosome Remodeling and Deacetylase (NuRD) complex may be involved in this function (Sridharan and Smale, 2007), and the Helios homolog Ikaros has been shown to trigger nucleosome deposition at a target gene via NuRD recruitment (Liang et al., 2017). In

addition, Ikaros induces the deposition of the repressive epigenetic mark H3K27me3 via the recruitment of the Polycomb Repressive Complex 2 (PRC2) to certain target gene loci in thymocytes, including the megakaryocyte lineage genes *Cd9* and *Vwf* (Oravecz et al., 2015). Whether Helios interacts with PRC2 remains to be investigated.

At genes that gain accessibility in the Helios null HSPCs, there is significant overlap between Helios binding sites and those bound by Gata2 and Runx1, two major transcriptional activators of the megakaryocyte program. Helios may therefore antagonize Gata2 and Runx1 function at common target genes. Of interest, Gata2 and Runx1 binding sites were previously shown to be enriched for Ets factor motifs (Doré et al., 2012), and Ets factors themselves are implicated in megakaryopoiesis (Lulli et al., 2006; Pang et al., 2006; Kruse et al., 2009; Tijssen and Ghevaert, 2013; Dalby et al., 2018). Both Ets factors and the Ikaros family bind GGAA, and it is tempting to speculate that Helios may compete with Ets proteins for binding. Indeed, Ikaros competes with Elf1 to repress *Dntt* expression in thymocytes (Trinh et al., 2001). It will be important to investigate if Helios competes for binding to target sequences, or if Helios repression overrides Gata2 and Runx1 activity, possibly via recruitment of chromatin remodeling complexes.

Both Helios and Ikaros are highly expressed in HSPCs. Like Helios, Ikaros deficiency results in decreased lymphocyte commitment and increased myelopoiesis (Ng et al., 2009). Unlike Helios, Ikaros KO mice exhibit a striking reduction in HSC repopulation potential (Nichogiannopoulou et al., 1999). These findings indicate both overlapping and unique functions of Helios and Ikaros in HSPCs. It is worth noting that Ikaros and Pegasus, another Ikaros family member, have been reported to influence megakaryopoiesis (Yoshida et al., 2006; Malinge et al., 2013; Lentaigne et al., 2019), suggesting that the megakaryocyte lineage is particularly sensitive to the Ikaros family. Furthermore, recent studies of multiple myeloma and myelodysplastic syndrome patient samples have shown that lenalidomide treatment, which down-regulates Ikaros and Aiolos in human cells (Krönke et al., 2014), leads to differentiation along the megakaryocyte pathways via Gata and Runx1 activation (Liu et al., 2018; Martínez-Høyer et al., 2020). Since Ikaros function in HSPCs was evaluated before the discovery of megakaryocyte-biased CD41⁺ LT-HSCs, it remains unclear if Ikaros acts like Helios in this subset. That Helios KO HSPCs express WT levels of Ikaros in our experiments clearly indicates that Ikaros alone is not sufficient to negatively regulate megakaryopoiesis.

Recently, Helios was reported to play a role in acute myeloid leukemias (AML) by promoting leukemia stem cell renewal and inhibiting their differentiation (Park et al., 2019). It remains unclear how the physiological function of Helios in HSPCs relates to its oncogenic

activity in AML, as genes or pathways important for AML leukemic stem cell biology did not emerge from our study. Curiously, the authors did not find a role for Helios in normal early hematopoiesis. However, their progenitor analyses as described were focused on the LSK, LK, GMP, CMP and MEP populations of young adult mice, cell types where we also did not detect differences at that age, and the effects of Helios on HSCs and MPPs were not closely studied. In addition, hematopoietic reconstitution was followed in a non-competitive setting and donor cell contribution was evaluated as a function of the total cell pool (donor and host), which differed from our competitive transplantation assays and may not reveal the lineage-priming potential of the mutant cells. Lastly, the mouse models studied were different (germline vs. a tissue-specific KO using a Vav-Cre deleter strain), though the mutations are similar. Thus it is currently not possible to directly compare the results from our two studies. These questions will be important to elucidate in the future.

ACKNOWLEDGEMENTS

We thank members of the Chan-Kastner lab for scientific discussion and help, and P. Marchal for technical support. We thank the IGBMC GenomEast platform for help with microarrays and sequencing, the IGBMC flow cytometry facility (C. Ebel, M. Philipps), the IGBMC animal facility (M. Gendron, S. Falcone, W. Magnant, A. Vincent), and the IGBMC cell culture facility.

S. Chan and P. Kastner received funding from the Agence Nationale de la Recherche (ANR-11-BSV3-018-01, ANR-17-CE15-0023-01), the Institut National du Cancer (INCa, PLBIO-2015-114), La Ligue Contre le Cancer (LNCC, Equipe Labellisée 2015-2017), Fondation pour la Recherche Médicale (Equipe FRM 2019), Fondation ARC, LNCC Grand Est/Bourgogne Franche Comté, and institute funds from INSERM, CNRS, Université de Strasbourg and the ANR-10-LABX-0030-INRT grant. G.C. and M-C.D. were funded by predoctoral fellowships from the ANR-10-LABX-0030-INRT grant and the FRM. G.C. received postdoctoral funding from the ANR-17-CE15-0023-01 grant. Q.C. received a predoctoral fellowship from Fondation ARC and postdoctoral funding from the INCa PLBIO-2015-114 grant.

Author contributions

Performed experiments: G.C., M-C.D., Q.C., V.M., P. Kirstetter.

Analyzed data: G.C., M-C.D., Q.C., V.M., P. Kirstetter, S.C.

Performed and analyzed high throughput sequencing experiments: M.J., M.C., C.T-C., B.J.

Contributed essential reagents: L.C., A.M.T., E.M.S.

Supervised the work: P. Kirstetter, P. Kastner, S.C.

Designed the study and wrote the manuscript: G.C., P. Kirstetter, P. Kastner, S.C.

Declaration of Interest

The authors declare no conflicting interests.

Accession Number

The sequencing data from this study are available at the GEO database (GSE149354).

REFERENCES

- Akashi, K., Traver, D., Miyamoto, T., and Weissman, I.L. (2000). A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* *404*, 193–197.
- Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* *31*, 166–169.
- Bernitz, J.M., Kim, H.S., MacArthur, B., Sieburg, H., and Moore, K. (2016). Hematopoietic Stem Cells Count and Remember Self-Renewal Divisions. *Cell* *167*, 1296–1309.e10.
- Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* *10*, 1213–1218.
- Busch, K., Klapproth, K., Barile, M., Flossdorf, M., Holland-Letz, T., Schlenner, S.M., Reth, M., Höfer, T., and Rodewald, H.-R. (2015). Fundamental properties of unperturbed haematopoiesis from stem cells in vivo. *Nature* *518*, 542–546.
- Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* *36*, 411–420.
- Cai, Q., Dierich, A., Oulad-Abdelghani, M., Chan, S., and Kastner, P. (2009). Helios deficiency has minimal impact on T cell development and function. *J. Immunol.* *183*, 2303–2311.
- Carrelha, J., Meng, Y., Kettle, L.M., Luis, T.C., Norfo, R., Alcolea, V., Boukarabila, H., Grasso, F., Gambardella, A., Grover, A., et al. (2018). Hierarchically related lineage-restricted fates of multipotent haematopoietic stem cells. *Nature* *554*, 106–111.
- Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* *14*, 959–962.
- Dalby, A., Ballester-Beltrán, J., Lincetto, C., Mueller, A., Foad, N., Evans, A., Baye, J., Turro, E., Moreau, T., Tijssen, M.R., et al. (2018). Transcription Factor Levels after Forward Programming of Human Pluripotent Stem Cells with GATA1, FLI1, and TAL1 Determine Megakaryocyte versus Erythroid Cell Fate Decision. *Stem Cell Reports* *11*, 1462–1478.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* *29*, 15–21.
- Doré, L.C., Chlon, T.M., Brown, C.D., White, K.P., and Crispino, J.D. (2012). Chromatin occupancy analysis reveals genome-wide GATA factor switching during hematopoiesis. *Blood* *119*, 3724–3733.
- Draper, J.E., Sroczyńska, P., Leong, H.S., Fadlullah, M.Z.H., Miller, C., Kouskoff, V., and Lacaud, G. (2017). Mouse RUNX1C regulates pre-megakaryocytic/erythroid output and maintains survival of megakaryocyte progenitors. *Blood* *130*, 271–284.

- Drissen, R., Buza-Vidas, N., Woll, P., Thongjuea, S., Gambardella, A., Giustacchini, A., Mancini, E., Zriwil, A., Lutteropp, M., Grover, A., et al. (2016). Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nat. Immunol.* *17*, 666–676.
- Gekas, C., and Graf, T. (2013). CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age. *Blood* *121*, 4463–4472.
- Getnet, D., Grosso, J.F., Goldberg, M.V., Harris, T.J., Yen, H.-R., Bruno, T.C., Durham, N.M., Hipkiss, E.L., Pyle, K.J., Wada, S., et al. (2010). A role for the transcription factor Helios in human CD4(+)CD25(+) regulatory T cells. *Mol. Immunol.* *47*, 1595–1600.
- Grover, A., Sanjuan-Pla, A., Thongjuea, S., Carrelha, J., Giustacchini, A., Gambardella, A., Macaulay, I., Mancini, E., Luis, T.C., Mead, A., et al. (2016). Single-cell RNA sequencing reveals molecular and functional platelet bias of aged haematopoietic stem cells. *Nat Commun* *7*, 11075.
- Haas, S., Hansson, J., Klimmeck, D., Loeffler, D., Velten, L., Uckelmann, H., Wurzer, S., Prendergast, Á.M., Schnell, A., Hexel, K., et al. (2015). Inflammation-Induced Emergency Megakaryopoiesis Driven by Hematopoietic Stem Cell-like Megakaryocyte Progenitors. *Cell Stem Cell* *17*, 422–434.
- Heuston, E.F., Keller, C.A., Lichtenberg, J., Giardine, B., Anderson, S.M., NIH Intramural Sequencing Center, Hardison, R.C., and Bodine, D.M. (2018). Establishment of regulatory elements during erythro-megakaryopoiesis identifies hematopoietic lineage-commitment points. *Epigenetics Chromatin* *11*, 22.
- Hirata, Y., Furuhashi, K., Ishii, H., Li, H.W., Pinho, S., Ding, L., Robson, S.C., Frenette, P.S., and Fujisaki, J. (2018). CD150^{high} Bone Marrow Tregs Maintain Hematopoietic Stem Cell Quiescence and Immune Privilege via Adenosine. *Cell Stem Cell* *22*, 445-453.e5.
- Ho, Y.-H., Del Toro, R., Rivera-Torres, J., Rak, J., Korn, C., García-García, A., Macías, D., González-Gómez, C., Del Monte, A., Wittner, M., et al. (2019). Remodeling of Bone Marrow Hematopoietic Stem Cell Niches Promotes Myeloid Cell Expansion during Premature or Physiological Aging. *Cell Stem Cell* *25*, 407-418.e6.
- Huang, Z., Dore, L.C., Li, Z., Orkin, S.H., Feng, G., Lin, S., and Crispino, J.D. (2009). GATA-2 reinforces megakaryocyte development in the absence of GATA-1. *Mol. Cell. Biol.* *29*, 5168–5180.
- Ikonomi, P., Rivera, C.E., Riordan, M., Washington, G., Schechter, A.N., and Noguchi, C.T. (2000). Overexpression of GATA-2 inhibits erythroid and promotes megakaryocyte differentiation. *Exp. Hematol.* *28*, 1423–1431.
- ImmGen Consortium (2016). Open-source ImmGen: mononuclear phagocytes. *Nat. Immunol.* *17*, 741.
- Kawaguchi, M., Kitajima, K., Kanokoda, M., Suzuki, H., Miyashita, K., Nakajima, M., Nuriya, H., Kasahara, K., and Hara, T. (2016). Efficient production of platelets from mouse embryonic stem cells by enforced expression of Gata2 in late hemogenic endothelial cells. *Biochemical and Biophysical Research Communications* *474*, 462–468.

Kim, H.-J., Barnitz, R.A., Kreslavsky, T., Brown, F.D., Moffett, H., Lemieux, M.E., Kaygusuz, Y., Meissner, T., Holderried, T.A.W., Chan, S., et al. (2015). Stable inhibitory activity of regulatory T cells requires the transcription factor Helios. *Science* 350, 334–339.

Krönke, J., Udeshi, N.D., Narla, A., Grauman, P., Hurst, S.N., McConkey, M., Svinkina, T., Heckl, D., Comer, E., Li, X., et al. (2014). Lenalidomide causes selective degradation of IKZF1 and IKZF3 in multiple myeloma cells. *Science* 343, 301–305.

Kruse, E.A., Loughran, S.J., Baldwin, T.M., Josefsson, E.C., Ellis, S., Watson, D.K., Nurden, P., Metcalf, D., Hilton, D.J., Alexander, W.S., et al. (2009). Dual requirement for the ETS transcription factors Fli-1 and Erg in hematopoietic stem cells and the megakaryocyte lineage. *Proc Natl Acad Sci U S A* 106, 13814–13819.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.

Lentaigne, C., Greene, D., Sivapalaratnam, S., Favier, R., Seyres, D., Thys, C., Grassi, L., Mangles, S., Sibson, K., Stubbs, M., et al. (2019). Germline mutations in the transcription factor IKZF5 cause thrombocytopenia. *Blood* 134, 2070–2081.

Liang, Z., Brown, K.E., Carroll, T., Taylor, B., Vidal, I.F., Hendrich, B., Rueda, D., Fisher, A.G., and Merkenschlager, M. (2017). A high-resolution map of transcriptional repression. *ELife* 6, e22767.

Liu, A., Li, S., Donnenberg, V., Fu, J., Gollin, S.M., Ma, H., Lu, C., Stolz, D.B., Mapara, M.Y., Monaghan, S.A., et al. (2018). Immunomodulatory drugs downregulate IKZF1 leading to expansion of hematopoietic progenitors with concomitant block of megakaryocytic maturation. *Haematologica* 103, 1688–1697.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550.

Lulli, V., Romania, P., Morsilli, O., Gabbianelli, M., Pagliuca, A., Mazzeo, S., Testa, U., Peschle, C., and Marziali, G. (2006). Overexpression of Ets-1 in human hematopoietic progenitor cells blocks erythroid and promotes megakaryocytic differentiation. *Cell Death Differ.* 13, 1064–1074.

Malinge, S., Thiollier, C., Chlon, T.M., Doré, L.C., Diebold, L., Bluteau, O., Mabialah, V., Vainchenker, W., Dessen, P., Winandy, S., et al. (2013). Ikaros inhibits megakaryopoiesis through functional interaction with GATA-1 and NOTCH signaling. *Blood* 121, 2440–2451.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal* 17, 10–12.

Martinez-Hoyer, S., Deng, Y., Parker, J., Jiang, J., Mo, A., Docking, T.R., Gharaee, N., Li, J., Umlandt, P., Fuller, M., et al. (2020). Loss of lenalidomide-induced megakaryocytic differentiation leads to therapy resistance in del(5q) myelodysplastic syndrome. *Nat. Cell Biol.* 22, 526–533.

- Mastio, J., Simand, C., Cova, G., Kastner, P., Chan, S., and Kirstetter, P. (2018). Ikaros cooperates with Notch activation and antagonizes TGF β signaling to promote pDC development. *PLoS Genet.* *14*, e1007485.
- Mirantes, C., Passegué, E., and Pietras, E.M. (2014). Pro-inflammatory cytokines: emerging players regulating HSC function in normal and diseased hematopoiesis. *Exp. Cell Res.* *329*, 248–254.
- Ng, S.Y.-M., Yoshida, T., Zhang, J., and Georgopoulos, K. (2009). Genome-wide lineage-specific transcriptional networks underscore Ikaros-dependent lymphoid priming in hematopoietic stem cells. *Immunity* *30*, 493–507.
- Nichogiannopoulou, A., Trevisan, M., Neben, S., Friedrich, C., and Georgopoulos, K. (1999). Defects in hemopoietic stem cell activity in Ikaros mutant mice. *J. Exp. Med.* *190*, 1201–1214.
- Oguro, H., Ding, L., and Morrison, S.J. (2013). SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* *13*, 102–116.
- Oravec, A., Apostolov, A., Polak, K., Jost, B., Le Gras, S., Chan, S., and Kastner, P. (2015). Ikaros mediates gene silencing in T cells through Polycomb repressive complex 2. *Nat Commun* *6*, 8823.
- Pang, L., Xue, H.-H., Szalai, G., Wang, X., Wang, Y., Watson, D.K., Leonard, W.J., Blobel, G.A., and Poncz, M. (2006). Maturation stage-specific regulation of megakaryopoiesis by pointed-domain Ets proteins. *Blood* *108*, 2198–2206.
- Pang, S.H.M., de Graaf, C.A., Hilton, D.J., Huntington, N.D., Carotta, S., Wu, L., and Nutt, S.L. (2018). PU.1 Is Required for the Developmental Progression of Multipotent Progenitors to Common Lymphoid Progenitors. *Front Immunol* *9*, 1264.
- Papathanasiou, P., Attema, J.L., Karsunky, H., Hosen, N., Sontani, Y., Hoyne, G.F., Tunngley, R., Smale, S.T., and Weissman, I.L. (2009). Self-renewal of the long-term reconstituting subset of hematopoietic stem cells is regulated by Ikaros. *Stem Cells* *27*, 3082-3092.
- Park, S.-M., Cho, H., Thornton, A.M., Barlowe, T.S., Chou, T., Chhangawala, S., Fairchild, L., Taggart, J., Chow, A., Schurer, A., et al. (2019). IKZF2 Drives Leukemia Stem Cell Self-Renewal and Inhibits Myeloid Differentiation. *Cell Stem Cell* *24*, 153-165.e7.
- Pencovich, N., Jaschek, R., Tanay, A., and Groner, Y. (2011). Dynamic combinatorial interactions of RUNX1 and cooperating partners regulates megakaryocytic differentiation in cell line models. *Blood* *117*, e1-14.
- Pierini, A., Nishikii, H., Baker, J., Kimura, T., Kwon, H.-S., Pan, Y., Chen, Y., Alvarez, M., Strober, W., Velardi, A., et al. (2017). Foxp3+ regulatory T cells maintain the bone marrow microenvironment for B cell lymphopoiesis. *Nat Commun* *8*, 15068.
- Pietras, E.M. (2017). Inflammation: a key regulator of hematopoietic stem cell fate in health and disease. *Blood* *130*, 1693–1698.

Pietras, E.M., Reynaud, D., Kang, Y.-A., Carlin, D., Calero-Nieto, F.J., Leavitt, A.D., Stuart, J.M., Göttgens, B., and Passegué, E. (2015). Functionally Distinct Subsets of Lineage-Biased Multipotent Progenitors Control Blood Production in Normal and Regenerative Conditions. *Cell Stem Cell* 17, 35–46.

Pinto do O, P., Kolterud, A., and Carlsson, L. (1998). Expression of the LIM-homeobox gene LH2 generates immortalized steel factor-dependent multipotent hematopoietic precursors. *EMBO J.* 17, 5744–5756.

Rodriguez-Fraticelli, A.E., Wolock, S.L., Weinreb, C.S., Panero, R., Patel, S.H., Jankovic, M., Sun, J., Calogero, R.A., Klein, A.M., and Camargo, F.D. (2018). Clonal analysis of lineage fate in native haematopoiesis. *Nature* 553, 212–216.

Sanjuan-Pla, A., Macaulay, I.C., Jensen, C.T., Woll, P.S., Luis, T.C., Mead, A., Moore, S., Carella, C., Matsuoka, S., Bouriez Jones, T., et al. (2013). Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature* 502, 232–236.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* 33, 495–502.

Schürch, C.M., Riether, C., and Ochsenbein, A.F. (2014). Cytotoxic CD8+ T cells stimulate hematopoietic progenitors by promoting cytokine release from bone marrow mesenchymal stromal cells. *Cell Stem Cell* 14, 460–472.

Sebastian, M., Lopez-Ocasio, M., Metidji, A., Rieder, S.A., Shevach, E.M., and Thornton, A.M. (2016). Helios Controls a Limited Subset of Regulatory T Cell Functions. *J. Immunol.* 196, 144–155.

Skadow, M., Penna, V.R., Galant-Swofford, J., Shevach, E.M., and Thornton, A.M. (2019). Helios Deficiency Predisposes the Differentiation of CD4+Foxp3- T Cells into Peripherally Derived Regulatory T Cells. *J. Immunol.* 203, 370–378.

Sridharan, R., and Smale, S.T. (2007). Predominant Interaction of Both Ikaros and Helios with the NuRD Complex in Immature Thymocytes. *J. Biol. Chem.* 282, 30227–30238.

Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P.K., Swerdlow, H., Satija, R., and Smibert, P. (2017). Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* 14, 865–868.

Stoeckius, M., Zheng, S., Houck-Loomis, B., Hao, S., Yeung, B.Z., Mauck, W.M., Smibert, P., and Satija, R. (2018). Cell Hashing with barcoded antibodies enables multiplexing and doublet detection for single cell genomics. *Genome Biol.* 19, 224.

Stuart, T., and Satija, R. (2019). Integrative single-cell analysis. *Nat. Rev. Genet.* 20, 257–272.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550.

Sun, D., Luo, M., Jeong, M., Rodriguez, B., Xia, Z., Hannah, R., Wang, H., Le, T., Faull, K.F., Chen, R., et al. (2014). Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell Stem Cell* *14*, 673–688.

Thornton, A.M., and Shevach, E.M. (2019). Helios: still behind the clouds. *Immunology* *158*, 161–170.

Thornton, A.M., Korty, P.E., Tran, D.Q., Wohlfert, E.A., Murray, P.E., Belkaid, Y., and Shevach, E.M. (2010). Expression of Helios, an Ikaros transcription factor family member, differentiates thymic-derived from peripherally induced Foxp3⁺ T regulatory cells. *J. Immunol.* *184*, 3433–3441.

Tijssen, M.R., and Ghevaert, C. (2013). Transcription factors in late megakaryopoiesis and related platelet disorders. *J Thromb Haemost* *11*, 593–604.

Trinh, L.A., Ferrini, R., Cobb, B.S., Weinmann, A.S., Hahm, K., Ernst, P., Garraway, I.P., Merkschlager, M., and Smale, S.T. (2001). Down-regulation of TDT transcription in CD4(+)CD8(+) thymocytes by Ikaros proteins in direct competition with an Ets activator. *Genes Dev.* *15*, 1817–1832.

Wang, Z., and Ema, H. (2016). Mechanisms of self-renewal in hematopoietic stem cells. *Int J Hematol* *103*, 498–509.

Wilson, N.K., Foster, S.D., Wang, X., Knezevic, K., Schütte, J., Kaimakis, P., Chilarska, P.M., Kinston, S., Ouwehand, W.H., Dzierzak, E., et al. (2010). Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* *7*, 532–544.

Wolfer, A., Bakker, T., Wilson, A., Nicolas, M., Ioannidis, V., Littman, D.R., Lee, P.P., Wilson, C.B., Held, W., MacDonald, H.R., et al. (2001). Inactivation of Notch 1 in immature thymocytes does not perturb CD4 or CD8T cell development. *Nat. Immunol.* *2*, 235–241.

Yoshida, T., Ng, S.Y.-M., Zuniga-Pflucker, J.C., and Georgopoulos, K. (2006). Early hematopoietic lineage restrictions directed by Ikaros. *Nat Immunol* *7*, 382–391.

Zang, C., Luyten, A., Chen, J., Liu, X.S., and Shivdasani, R.A. (2016). NF-E2, FLI1 and RUNX1 collaborate at areas of dynamic chromatin to activate transcription in mature mouse megakaryocytes. *Sci Rep* *6*, 30255.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* *10*, 1523.

MATERIALS AND METHODS

Mice

Helios germline (*Ikzf2*^{-/-}, KO), Helios floxed (*Ikzf2*^{fl/fl}) and CD4-Cre tg mice were previously described (Cai et al., 2009; Sebastian et al., 2016; Wolfer et al., 2001). Helios KO mice were backcrossed 6 generations onto the C57BL/6 background, while the *Ikzf2*^{fl/fl} and CD4-Cre mice were backcrossed >10 generations on C57BL/6. All mice were bred and maintained in a SPF facility and used according to IGBMC Ethical Committee (Com'Eth) guidelines. Both female and male mice were investigated and evenly distributed in the different experimental groups; aged-matched animals were studied.

Flow cytometry and cell sorting

Cells were stained with the antibody (Ab) combinations listed in Table S2 in PBS/0.5% BSA/2mM EDTA. Samples were acquired on a LSR II or Fortessa analyzer (BD Biosciences) and analyzed with FlowJo10 software. For intracellular staining, cells were fixed overnight (ON) at 4°C, then permeabilized and stained with the primary Ab for 1h at RT and, when necessary, with a secondary Ab for 1h on ice using the eBiosciences Foxp3/Transcription Factor Staining Kit (Thermofisher).

For cytokine staining, CD4⁺ T cells were first magnetically enriched using the Dynabeads Untouched™ Mouse CD4 Cells Kit (Thermofisher), then stimulated for 2h at 37°C with PMA and ionomycin (0.5ug/ml each) in the presence of BD GolgiPlug (BD Biosciences) in IMDM supplemented with 10% FCS, MEM, HEPES (10mM), GlutaMAX (Gibco), sodium pyruvate (1mM), 2-mercaptoethanol (50uM). Cells were stained and analyzed as above.

For sorting HSPC subsets, cells were stained for lineage marker expression (Table S2) before depletion with Dynabeads sheep anti-rat IgG beads (Thermofisher). Lineage-negative cells were stained as above and sorted in the presence of DAPI (1uM/ml) using a FACS ARIA II or Fusion (BD Biosciences). Sort purity was >98%.

Colony forming unit (CFU) assays

For Mk and myeloid CFUs, 50x10³ cells were cultured in methylcellulose media (MethoCult 3231; Stem Cell Technologies) in the presence of recombinant mSCF (50ng/ml), mTPO (50ng/ml), hEPO (5U/ml), mIL3 (20ng/ml) and mIL11 (40ng/ml) (cytokines from PeproTech), as described (Akashi et al., 2000). For preB CFUs, 15x10⁴ cells were cultured in methylcellulose media containing hIL7 (MethoCult 3630; Stem Cell Technologies). All assays were performed in duplicate and CFUs were counted after 7d. Single colonies were

cytospun onto slides, subjected to May-Grünwald Giemsa staining, and analyzed with an upright microscope (Leica DM4000 B 40X). Images were captured with the Coolsnap software.

HSPC single cell cultures

For LT-HSCs (CD48⁻CD150⁺ LSK), single cells were sorted and cultured in U-shaped 96-well plates containing 50ul StemSpan SFEM (StemCell Technologies) supplemented with 10% FCS, 2-mercaptoethanol (50uM), penicillin/streptomycin, mIL3 (20ng/ml), mSCF (50ng/ml) and mTPO (50ng/ml) (cytokines from PeproTech) (Sanjuan-Pla et al., 2013). Colonies was evaluated at d10. For MPPs (CD48⁺CD150⁻ LSK), single cells were sorted and cultured in flat-bottom 96-well plates in 50ul complete methylcellulose media (MethoCult 3434; Stem Cell Technologies). Colony number and morphology were evaluated after 7d. Single colonies were cytospun and stained as above.

BM competitive transplantation assays

15x10⁴ WT or KO BM cells from 10- to 15-week-old CD45.2⁺ mice were i.v. injected with 15x10⁴ CD45.1⁺ WT BM cells into 9Gy lethally-irradiated CD45.1⁺CD45.2⁺ hosts. Peripheral blood was analyzed 2 and 4 months after reconstitution, as described (Mastio et al., 2018). The contributions to myeloid (CD11b⁺B220⁻), B (CD11b⁺B220⁺) and T cells (CD11b⁻B220⁻CD4⁺CD8⁺) from donor or competitor origin were defined respectively as % cells within the CD45.2⁺Ter119⁻ or CD45.1⁺Ter119⁻ populations.

LT-HSC repopulation assays

100 WT or KO LT-HSCs (CD48⁻CD150⁺ LSK) from 6-week-old CD45.2⁺ mice were i.v. injected with 5x10⁵ supporting CD45.1⁺ WT BM cells into 9Gy lethally-irradiated CD45.1⁺CD45.2⁺ hosts. Peripheral blood was analyzed 2 and 4 months after transplantation. The contributions to donor myeloid (CD11b⁺B220⁻), B (CD11b⁺B220⁺) and T cells (CD11b⁻B220⁻CD4⁺CD8⁺) were defined as % cells within the CD45.2⁺Ter119⁻ population.

MPP3 and MPP4 potential in vivo assays

5x10³ WT or KO MPP3 (CD48⁺CD150⁻Flt3⁻ LSK) or MPP4 (CD48⁺CD150⁻Flt3⁺ LSK) cells from 15-week-old CD45.2⁺ mice were i.v. injected into 6Gy sublethally-irradiated CD45.1⁺CD45.2⁺ hosts. Peripheral blood was analyzed 2 weeks after transplantation. Mice with <0.5% CD45.2⁺ cells were excluded from the analysis (in total 4/60 mice were

excluded). The % CD45.2⁺ myeloid (CD11b⁺B220⁻) and B (CD11b⁺B220⁺) cells were analyzed.

Transcriptome analysis by microarray and GSEA

The transcriptomes of 5x10³ LT-HSCs (CD34⁻Flt3⁻ LSK) and 5x10⁴ LMPPs (CD34⁺Flt3⁺ LSK) from WT and KO 6-week-old mice (3 each) were analyzed with Affymetrix 430.2.0 arrays. RNA was extracted with the RNeasy Micro kit (Qiagen), and biotinylated cRNA targets were prepared, starting from 2- to 20ng of total RNA, using the Message Amp II aRNA Amplification Kit (Ambion) with two rounds of amplification, according to manufacturer's recommendations. cRNAs were hybridized to the microarrays using standard conditions. Hierarchical clustering were performed using the Cluster 3 software. GSEA was performed using the GSEA 2.0 software (Subramanian et al., 2005). In these experiments only, the mice came from a mixed C57BL/6 and 129/Sv genetic background.

Transcriptome analysis by RNA-seq

Total RNA was extracted from 1-4x10⁴ LT-HSC (CD48⁻CD150⁺ LSK), MPP3 (CD48⁺ CD150⁻ Flt3⁻ LSK) and MPP4 (CD48⁺CD150⁻Flt3⁻ LSK) cells from WT and KO 10-week-old mice (3 each), using the RNeasy Plus Micro Kit (Qiagen). Libraries were prepared with the Clontech SMART-seq v4 Ultra Low Input RNA Kit for Sequencing, followed by the Illumina Nextera XT DNA library preparation kit, and sequenced with a HiSeq 4000 (Illumina) with single-end 50bp read length. Reads were pre-processed to remove adapters using Cutadapt version 1.10 (Martin, 2011), poly(A) and low-quality sequences (Phred quality score below 20). Reads shorter than 40 bases were discarded, and then aligned to the *M. musculus* genome mm10 using STAR v2.5.3a (Dobin et al., 2013). Gene expression quantification was performed from uniquely aligned reads using HTSeq (v0.6.1p1), with annotations from ENSEMBL GRCm38 and "union" mode (Anders et al., 2015). Differential gene expression analyses were performed using the Bioconductor package DESeq2 (v1.16.1) on R (v3.3.2) (Love et al., 2014). Wald statistical test was used to identify genes differentially expressed between WT and KO samples. Heatmaps of differentially expressed genes were created using Cluster and Java TreeView softwares. For the GSEA, up- and down-regulated genes identified in the LT-HSC, MPP3 or MPP4 KO populations (p-value <0.05) were used to create the ranked lists. The MPP4 signature was obtained using our transcriptome data. MMP4 signature genes were identified as the genes with significant differential expression (p-adjusted value <0.05, log2FC >1) in both the WT LT-HSC vs WT MPP4, and WT MPP3 vs WT MPP4, comparisons. The CLP, MkP and old HSC signatures were taken from the

literature (Drissen et al., 2016; Sun et al., 2014). Pathway analyses were performed at the metascape website (<http://metascape.org>) (Zhou et al., 2019).

Single cell mRNA-sequencing

3.5x10⁴ LSK cells from WT and KO 10-week-old mice (2 each) were individually labeled with unique oligonucleotide-tagged Abs against CD45 and MHC class I (TotalSeq Hashtag 1 #A0301, Hashtag 2 #A0302, Hashtag 3 #A0303, Hashtag 4 #A0304; BioLegend) in order to multiplex the 4 samples (Anders et al., 2015). Each sample was then labeled with anti-CD150 and anti-CD48 Abs conjugated with an oligonucleotide containing a unique barcode for Ab identification (TotalSeq anti-CD150 #115945, anti-CD48 #103477; BioLegend) (Stoeckius et al., 2017, 2018). After HASH- and CITE-tagged cells were counted and checked for viability, 10⁴ cells (2500 LSK cells from each sample) were loaded into a Chromium Controller (10x Genomics), captured and library prepped with the Chromium Single Cell 3' v2 Kit (10x Genomics). Briefly, following cell lysis inside droplets, cellular mRNA and Ab-derived oligos were reverse-transcribed and indexed with a shared cellular barcode. Indexed cDNAs were pooled and amplified by PCR according to the Chromium Single Cell 3' v2 protocol with specific primers to amplify the Ab tags. SPRI bead size selection was performed to separate the mRNA-derived cDNA (>300bp) and the Ab-derived tagged cDNAs (180bp). For the mRNA-derived cDNA library preparation, we followed the manufacturer's instructions for Single Cell 3' v2 protocol. For Ab-derived tagged libraries, we used the KAPA HiFi HotStart Library Amplification Kit (Roche) with the following primers and amplification program.

CITE-seq (ADT) library:

Amplification was performed at 95°C for 3min followed by 10 cycles at 95°C for 20s, 60°C for 30s, 72°C for 20s, and elongation at 72°C for 5min with the 5'-CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCCTTGGCACCCGAG AATTCCA small RNA RPI1 primer as the i7 primer (Illumina).

HASH-seq (HTO) library:

Amplification was performed at 95°C for 3min followed by 10 cycles at 95°C for 20s 64°C for 30s, 72°C for 20s, and elongation at 72°C for 5min with the 5'-CAAGCAGAAGACGGCATACGAGATCGAGTAATGTGACTGGAGTTCAGACGTGTGCTCTTTC TruSeq D701_s primer as the i7 primer (Illumina).

The 5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTC SI-PCR primer was used as the i5 primer (Illumina) for both CITE- and HASH- seq library amplification.

Following final bead purification, all 3 libraries were pooled as 80% mRNA library, 10% ADT library and 10% HTO library before sequencing. cDNA libraries were sequenced on the HiSeq 4000 paired-end 100bp read length.

Gene expression and library analyses:

The 3' mRNA-seq library reads were processed using CellRanger pipeline v3.0.2 on mm10 and ENSEMBL GRCm38. Cell hashing- and CITE-seq-derived library reads were merged together and processed with 3' mRNA-seq library using the Feature Barcoding Analysis option. Tag identification from the HTO library was performed using the approach proposed by the Seurat R package v3.0.0 and R v3.5.1 on HTO counts (Butler et al., 2018; Stoekius et al., 2018; Stuart and Satija, 2019). Further, the CellRanger count was used to filter out low quality cells (doublets and/or mitochondrial genes). To detect and remove doublets, the distribution of the total number of unique molecular identifier (UMI) of each cell was computed for the 3 libraries. The filtering threshold used corresponded to the 99th percentile of the distribution. Hence, the threshold was 32,883 UMIs for the 3' mRNA library, 740 UMIs for the ADT library and 3,234 UMIs for the HTO library. Further, we filtered out cells with >5% mitochondrial gene contamination (indicating possibly poor quality or dying cells). Principal component analysis (PCA) of the datasets were performed with Seurat R package Satija Lab, v2.2 (Satija et al., 2015). The 3 first principal components, containing the highest standard deviation, were further used to run a Uniform Manifold Approximation and Projection (UMAP) non-linear dimensional reduction. UMAP coordinates were subsequently visualized with Loupe browser (10xGenomics) for further analysis.

The Loupe Browser was used to 'label' and export the cell population of interest. The Loupe manual filtering option, using the PC1-2 genes or K-means clustering (K=2 or 3), was used to identify proliferative and quiescent cells on the UMAP. Given that both approaches gave similar results, we decided to proceed further using K-means clustering, which allowed a more precise definition of the border of the proliferative and quiescent populations.

Single cell heatmaps were obtained by using R Seurat package. The 30 top negative and 30 top positive genes identified with PC3 were selected, and the cells were ranked based on their PC3 score. PC3-negative genes were defined as HSC-MPP2 genes as they were significantly more expressed in LT-HSC with respect to MPP4, while PC3-positive genes were considered as lymphoid-specific genes because they were significantly more expressed in MPP4 with respect to LT-HSC.

The lineage priming score was calculated as follows: for each cell, a gene from the HSC/MPP2 or MPP4 signature was considered expressed when its z-score value was >1. Each HSC/MPP2 gene was counted as -1; each MPP4 gene was counted as +1. Thus, cells

that expressed a predominantly HSC/MPP2 gene expression signature had an overall negative score, while those that expressed an MPP4 signature had an overall positive score.

Pathway analysis of PC1 and PC2 genes (30 top positive and 30 top negative genes) was performed at the metascape website (<http://metascape.org>) (Satija et al., 2015).

HPC7 cells

The HPC7 cell line was cultured in IMDM supplemented with 5% FCS, monothiolglycerol (150 μ M), penicillin/streptomycin, mSCF (50ng/ml) and mIL6 (10ng/ml) (Pinto do O et al., 1998). HPC7 cells were plated at 8×10^4 cells/ml and split every 2-3 days.

ChIP-sequencing

50×10^6 HPC7 cells were used for ChIP-seq. The proteins were crosslinked in 2mM disuccinimidyl glutarate (ChemCruz) for 45min at RT. Protein-DNA crosslinking was performed in 1% paraformaldehyde (PFA) for 10min at RT followed by addition of 2M glycine for 5min at RT to quench. Cell lysis was performed in the presence of a protease inhibitor cocktail (Roche #11836153001) as follows: 10min at 4°C with LB1 buffer (50mM tris-HCl pH8, 2mM EDTA, 0.1% NP-40, 10% glycerol) followed by 10min at 4°C with pre-chilled LB2 buffer (10mM tris-HCl pH8, 1mM EDTA, 0.5mM EGTA, 200mM NaCl). Nuclei were resuspended in LB2.3 (18mM tris-HCl pH8, 1mM EDTA, 0.5mM EGTA, 200mM NaCl, 0.033% Na-deoxycholate, 0.16% N-laurylsarcosine) and sonicated for 6 cycles (5min high, 15s on/15s off) with a Bioruptor (Diagenode). The sonicated chromatin was diluted 10x into a dilution buffer (50mM tris-HCl pH8, 200mM NaCl, 5mM EDTA, 0.5% NP-40) and pre-cleared for 1h at 4°C. The cleared chromatin was incubated ON at 4°C with a rabbit monoclonal anti-Helios Ab (Cell Signaling). Immunoprecipitation was performed with protein A-Sepharose beads (Sigma) in the presence of 0.1% SDS for 1h at 4°C. Immunoprecipitated complexes were washed 6x in washing buffer (20mM tris-HCl pH8, 500mM NaCl, 2mM EDTA, 1% NP-40, 0.1% SDS) and 3x in TE. The immunoprecipitated complexes were eluted in 2% SDS in TE for 1min at RT. De-crosslinking was performed in the elution buffer ON at 65°C whereas the input chromatin was de-crosslinked in 200mM NaCl in the same condition. DNA was purified using MinElute PCR Purification Kit (Qiagen). Libraries were prepared using the MicroPlex Library Preparation Kit v2 (Diagenode #C05010014) following the manufacturer's instructions. Single-end 50bp sequences were obtained with the Hiseq 4000. Image analysis and base calling were performed using RTA 2.7.7 and bcl2fastq 2.17.1.14. Adapter-dimer reads were removed using DimerRemover

(<https://sourceforge.net/projects/dimerremover/>). The reads were aligned to mm10 using Bowtie. Peak calling was performed with MACS v2.2.4 (Langmead and Salzberg, 2012).

Fast-ATAC-sequencing and data analysis

5×10^4 LSK cells from WT and KO 10-week-old mice (2 each) were used for fast-ATAC-seq (Corces et al., 2017). Briefly, cells were pelleted by centrifugation at 500g for 5min at 4°C, and resuspended in 50ul of transposase mixture (25ul 2x TD buffer, 2.5ul TDE1, 0.5ul 1% digitonin, and 22ul nuclease-free water) (FC-121-1030, Illumina; G9441, Promega). Transposition reactions were performed at 37°C for 30 min under agitation at 300rpm. Transposed DNA was purified using MinElute PCR Purification Kit (Qiagen) and eluted in 20ul elution buffer. Tagmented fragments were amplified for 12 cycles (Buenrostro et al., 2013). Paired-end 100bp reads were obtained with the Hiseq 4000. Data were analyzed using the Encode ATAC-seq pipeline v1.5.1. Reads were aligned to mm10 using Bowtie, and peak calling was performed using MACS2 v2.2.4 (Langmead and Salzberg, 2012). The peaks detected in each sample were combined using Bedtools merge v2.26.0. Reads were normalized across libraries (Anders et al., 2015). Comparisons of interest were performed using the method proposed by the DESeq2 Bioconductor library (DESeq2 v1.24.0) (Love et al., 2014). Visualization of normalized data was generated from post-processed files using Homer make UCSC file v4.11.1 with the "-norm 3e7 -style dnase" parameter (Anders et al., 2015).

Identification of the ATAC-seq peaks bound by Helios was performed by intersecting the intervals of the Bedtools (v2.29.0; usegalaxy.eu) with ATAC-seq regions (KO vs WT $p < 0.05$), and with Helios ChIP-seq peaks using a pileup > 15 . Motif research was performed at the meme website (<http://meme-suite.org/>) using Discriminative Regular Expression Motif Elicitation (DREME) v5.1.1. Data were visualized using Integrative Genomic Viewer (IGV_2.6.2) and seqMINER_1.3.3g. Pathway analyses were performed at the metascape website (<http://metascape.org>) (Zhou et al., 2019).

FIGURE LEGENDS**Figure 1 - Helios is highly expressed in HSPCs.**

(a-d) Representative gating strategies and Helios expression in BM hematopoietic cell populations. The gray histograms correspond to total BM cells stained with the secondary Ab alone. In (a) the LSK (Lin⁻Sca1⁺cKit⁺, red gate), LK (Lin⁻Sca1⁻cKit⁺, blue gate) and Lin⁻Sca1⁻ckit⁻ (orange gate) cells are indicated. Numbers correspond to percentages. The indicated HSPC populations are: LT-HSC: CD48⁻CD150⁺ LSK; MPP2: CD48⁺CD150⁺ LSK; MPP3: CD48⁺CD150⁻Flt3⁻ LSK; MPP4: CD48⁺CD150⁻Flt3⁺ LSK; MkP (megakaryocyte progenitor): CD150⁺CD41⁺ LK; GMP (granulocyte-monocyte progenitor): CD150⁻CD16/32⁺ LK; EryP (erythrocyte progenitor): CD150⁺CD71⁺ LK; CLP (common lymphoid progenitor): Lin⁻ckit^{med}sca1^{med}Flt3⁺IL7R⁺. The mature cell populations are: erythroid: Ter119⁺CD71⁺CD11b⁻; myeloid: CD11b⁺B220⁻Ter119⁻; B: B220⁺CD19⁺; CD4⁺ T: CD4⁺CD8⁻; CD8⁺ T: CD4⁻CD8⁺; Treg: CD4⁺Foxp3⁺, non-Treg T: CD4⁺Foxp3⁻. (e) Helios Median Fluorescence Intensity (MFI) in the indicated populations. Light blue bars indicate HSPC populations. Medium blue bars indicate committed progenitors. Dark blue bars indicate mature cells. (f) Helios MFI in BM non-hematopoietic cells: EC: CD45⁺NG2⁻CD31⁺; MSC: CD45⁻NG2⁺CD31⁻; total LT-HSCs as positive control. In (e) and (f), mean±SD of 2-3 independent experiments.

Figure 2 - Helios deficiency affects BM megakaryocyte and lymphoid progenitor frequencies.

(a-g) Relative abundance of the indicated BM populations in WT and KO mice at 6, 10 and 20 weeks of age. Representative contour plots and statistical significance are shown. In the histograms of (d) and (e), the control histogram shows the CD41 level of WT MPPs (CD48⁺CD150⁻ LSK). Numbers in the plots correspond to percentages. Each datapoint corresponds to one mouse; mean±SD (standard deviation) are shown per population for 4-8 mice from multiple independent experiments. Unpaired two-tailed t-test: *p<0.05, **p<0.01, ***p<0.001.

Figure 3 - In vitro differentiation of Helios KO HSPCs.

(a) Strategy of preB and Myeloid-Mk colony forming unit (CFU) assays. (b) Number of CFU-preB from 15x10⁴ BM cells seeded after 7d (left). Each datapoint represents the mean of duplicate cultures for each mouse. Mean±SD of 4 independent experiments. Representative May-Grünwald Giemsa (MGG) staining of preB cell colony (right). (c) Ratios of CFU-Myeloid (granulo-monocyte) and CFU-Mk from 5x10⁴ BM cells seeded after 7d (left). Mean±SD of 4

independent experiments. Representative MGG staining of different colonies (right). **(d)** Strategy of single cell LT-HSC cultures (left). 100 wells were seeded for each sample. % of wells containing Mk^+ or Myeloid-only colonies after 10d (middle) are indicated. Boxes represent the mean \pm SD of 6 independent experiments. Representative MGG staining of colonies (right). **(e)** Strategy of single cell MPP cultures (left). 120 wells were seeded for each sample. % of wells containing Myeloid colonies after 7d (middle) are indicated. Boxes represent the mean \pm SD of 4 independent experiments. For (c) and (d), Mk: megakaryocyte; Mo: monocyte/macrophage; PMN: polymorphonuclear granulocyte. Unpaired two-tailed t-test: * $p < 0.05$, ** $p < 0.01$; paired two-tailed t-test: ## $p < 0.01$.

Figure 4 - In vivo lymphoid potential of Helios KO HSPCs.

(a) Strategy of MPP transplantation (left). Representative contour plot of peripheral blood cell stainings (middle). Numbers correspond to percentages. Mean \pm SD of 4 independent experiments with 3-5 recipient mice per donor genotype (right). Each datapoint represents % donor cells per recipient. B: $B220^+$ cells; My: $CD11b^+$ cells. **(b)** Strategy of LT-HSC transplantation (top left). Representative contour plot of peripheral blood cell stainings (top right). Numbers correspond to percentages. Mean \pm SD of 2 (T cells) or 3 (B and Myeloid cells) independent experiments, with 6-8 recipient mice per donor genotype (bottom). Each datapoint represents % donor cells per recipient. Unpaired two-tailed t-test: ** $p < 0.01$, *** $p < 0.001$.

Figure 5 - Helios KO HSPCs are transcriptionally primed towards the megakaryocyte lineage.

(a) Heatmap of up- and down-regulated genes in LT-HSC, MPP3 and MPP4 (KO vs WT, $\log_2FC > 0.56$, p -value < 0.05). Highlighted genes in the LT-HSC samples correspond to those specific for megakaryocyte progenitors and up-regulated (in red), or those specific for MPP4s and down-regulated (in blue). **(b, c)** GSEA of the indicated populations. Ranked gene lists comprise genes with significant differential expression between WT and KO samples (p -value < 0.05), ranked according to their FC (KO vs WT). Signature lists of MkP and CLP (Grover et al., 2016), MPP4 (described in Materials and Methods), and up- and down-regulated genes in old LT-HSC (Sun et al., 2014) are indicated. The Normalized Enrichment Score (NES) and p -values are shown for each analysis. **(d)** Metascape analyses of up- (red) and down- (blue) regulated genes in the KO LT-HSCs ($\log_2FC > 0.56$, p -value < 0.05). Each cluster (color coded in the map on the left) is defined by its identity name

and is composed of nodes that share similar genes. Node size is proportional to the gene content number.

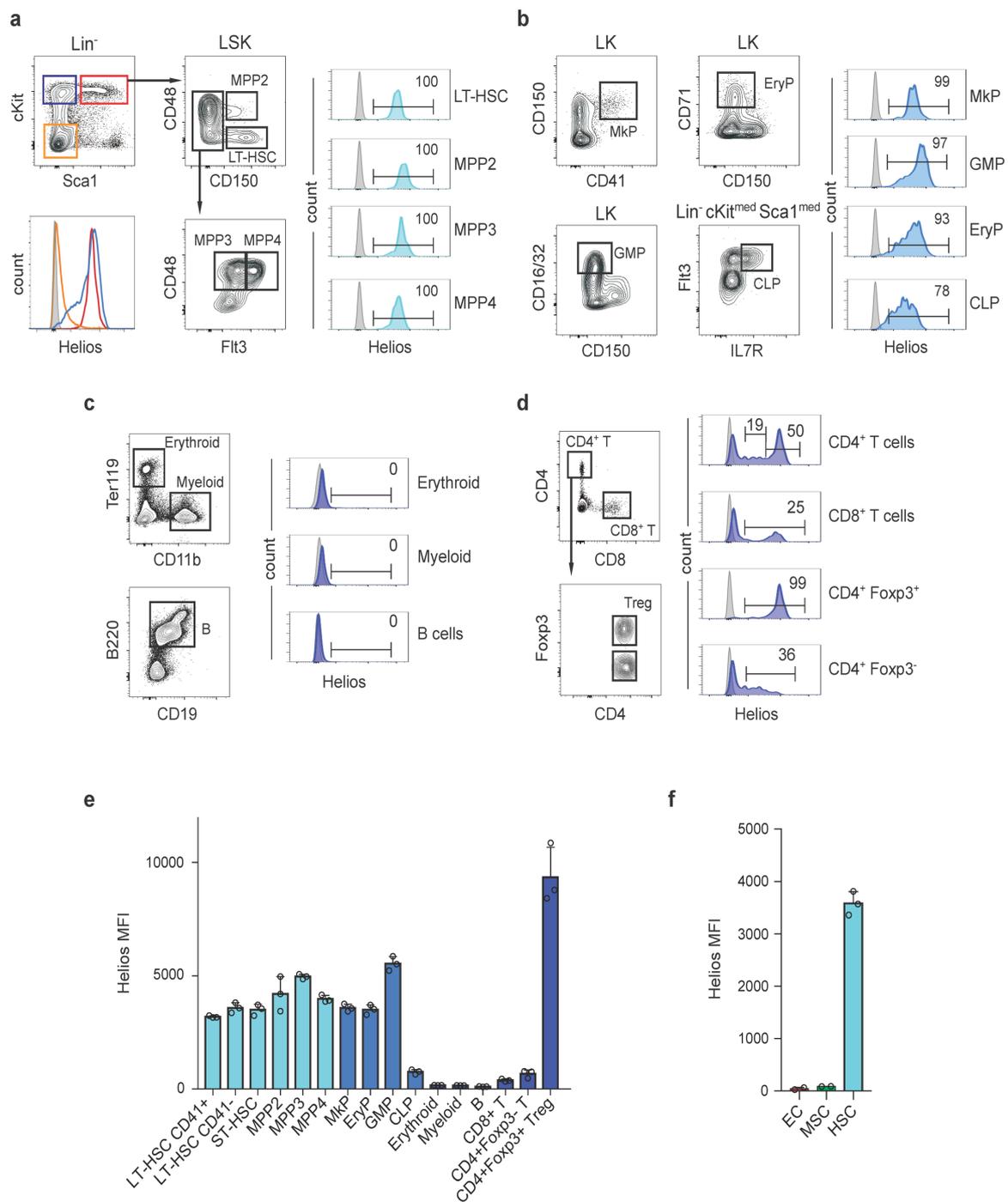
Figure 6 - Helios affects lineage priming at the single cell level.

(a-b) UMAP plots of WT and KO HSPCs, derived from one WT and one KO sample with similar numbers of sequenced cells. **(a)** Characterization and quantification of HSPCs based on their cell cycle feature, as indicated. **(b)** Characterization and quantification of HSPCs based on their lineage priming features. The quantification of the cells is shown below the UMAP plots. **(c)** Left: Heatmaps of single, quiescent WT and KO HSPCs. The order of the cells from left to right corresponds to their order on the PC3 axis. The y-axis indicates the 60 genes that contribute the most to the PC3 variance (top 30 negative genes in red, top 30 positive genes in blue). The white vertical lines correspond to the median positions of cells with a lineage priming score of 0. Right: Lineage priming score of WT and KO quiescent HSPCs. Mean \pm SD of 2 pooled independent experiments. **(d)** Left: Heatmaps of single, cycling WT and KO HSPCs, ordered as in **(c)**. Right: Lineage priming score of WT and KO cycling HSPCs. Mean \pm SD of 2 pooled independent experiments. Statistical significance calculated with Mann-Whitney test.

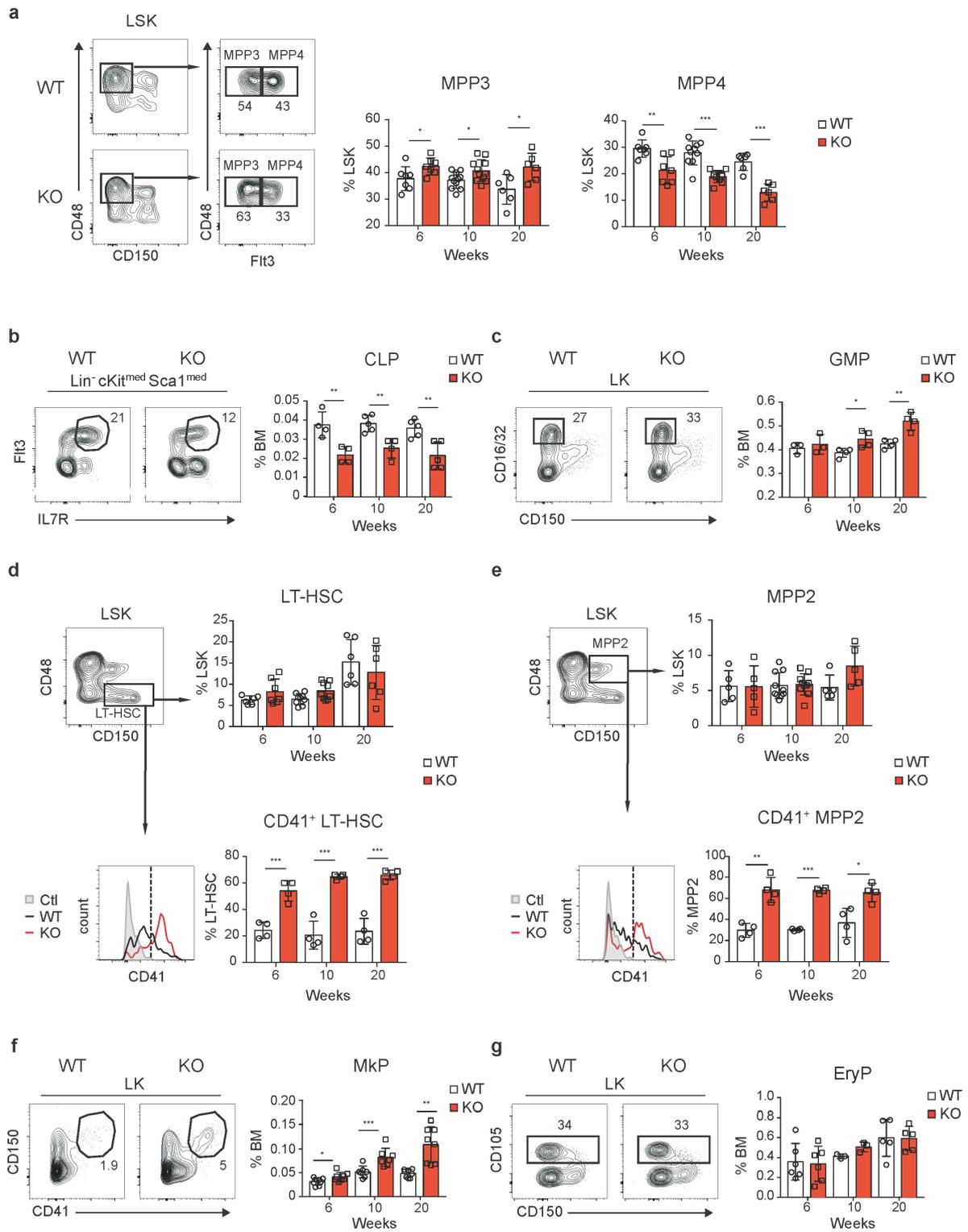
Figure 7 - Helios targets HSPC genes through gene repression

(a) Box plots showing the distribution of the log₂FC values of the ATAC-seq peaks in KO vs WT LSK samples, bound or not by Helios in HPC7 cells, either among all ATAC-seq peaks, or those associated with MkP signature genes. Each box plot was generated with a random selection of 300 regions. Mann-Whitney test: ***p<0.001. **(b)** % of chromatin regions containing "Helios sites" where the ATAC-seq peaks varied between WT and KO cells. ATAC-seq peaks were divided into 4 subsets, based on whether their intensity decreased or increased < or >0.3-fold (log₂FC) in KO cells. The numbers above each bar indicate the number of Helios-bound peaks (top) within the total number of peaks in the subset (bottom). Hypergeometric test: **p<0.01, ***p<0.001. **(c)** Metascape analysis of the enriched pathways associated with Helios-bound regions with significant increases in their chromatin opening (top 30% of peaks with a positive log₂FC and p<0.05; n=763 ATAC-seq peaks corresponding to 495 genes). Platelet-related pathways are highlighted in red. **(d)** Integrative genome viewer screenshots of representative platelet genes. Shown are ATAC-seq signals in WT and KO LSK cells, and Helios, Gata2 and Runx1 binding in HPC7 cells (Gata2 and Runx1 data are from the GEO dataset GSE22178). Boxed regions belong to the regions selected in **(c)** as having significant increases of ATAC-seq signals in KO cells. **(e)**

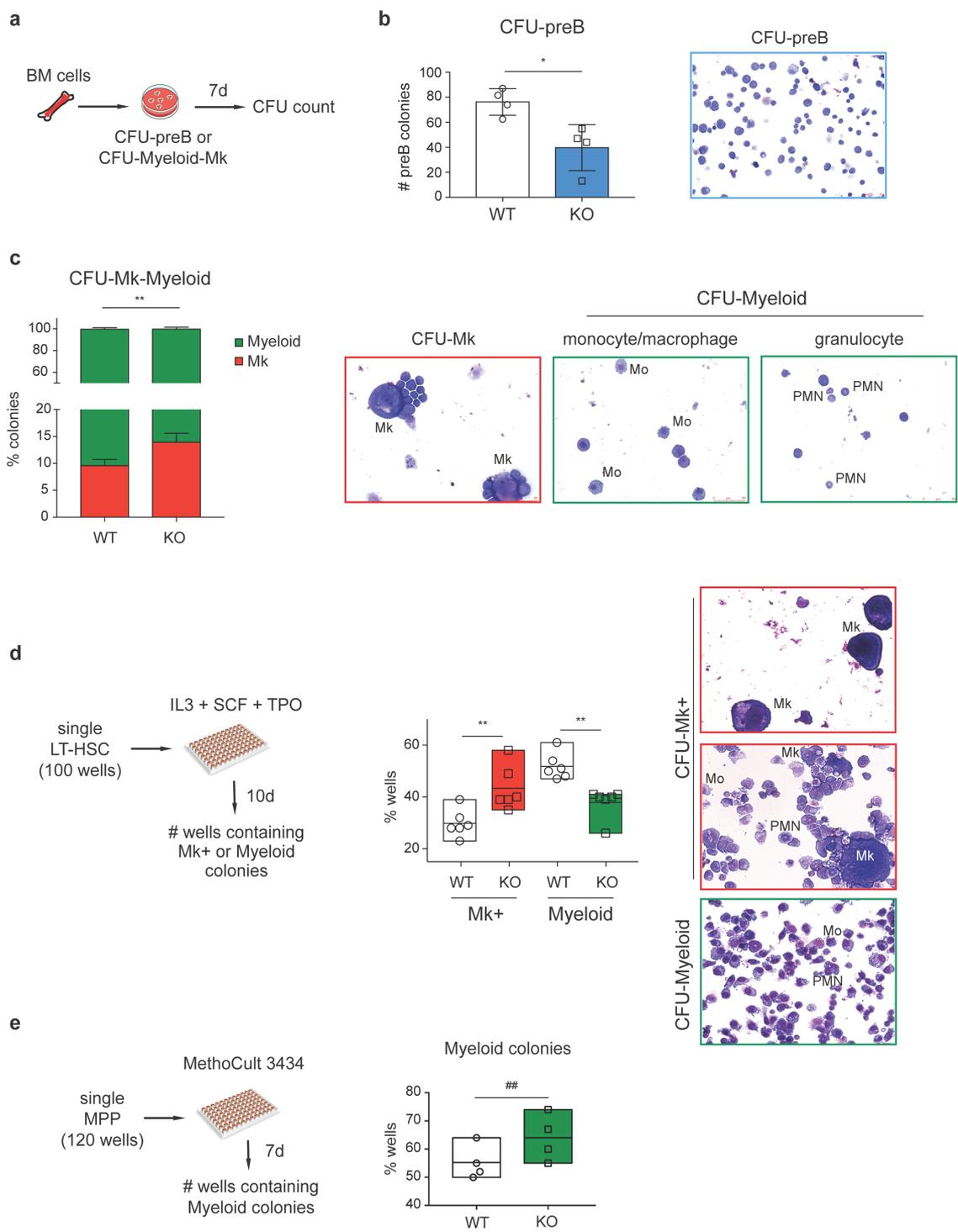
Motif enrichment within the 763 ATAC-seq peaks from (c). **(f)** Seq-miner heatmap showing ATAC-seq signals in WT and KO LSK cells, along with Helios, Gata2 and Runx1 binding in HPC7 cells. The 1576 regions without Helios binding were selected using the same criteria as those with Helios binding.



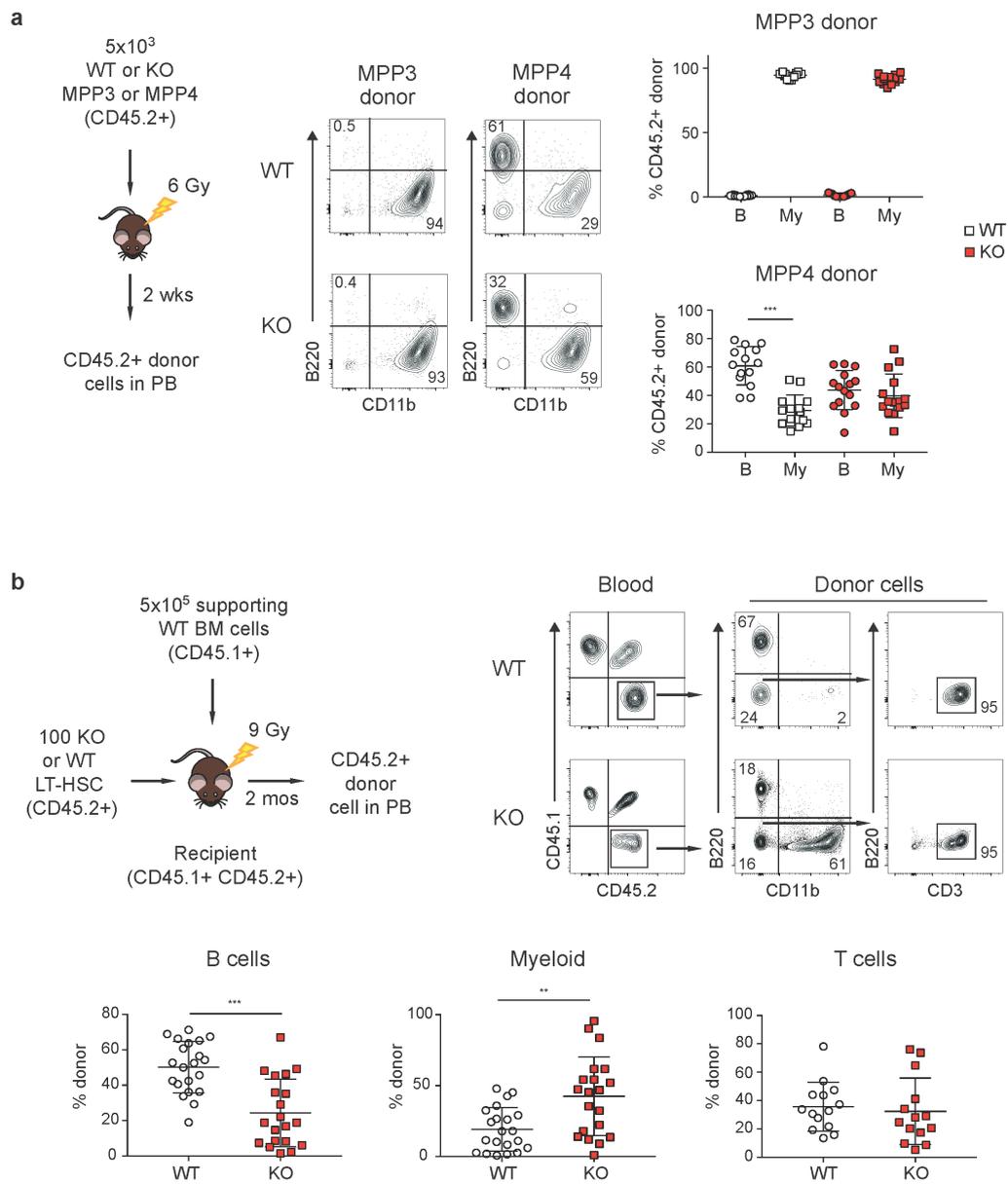
Cova et al. Figure 1



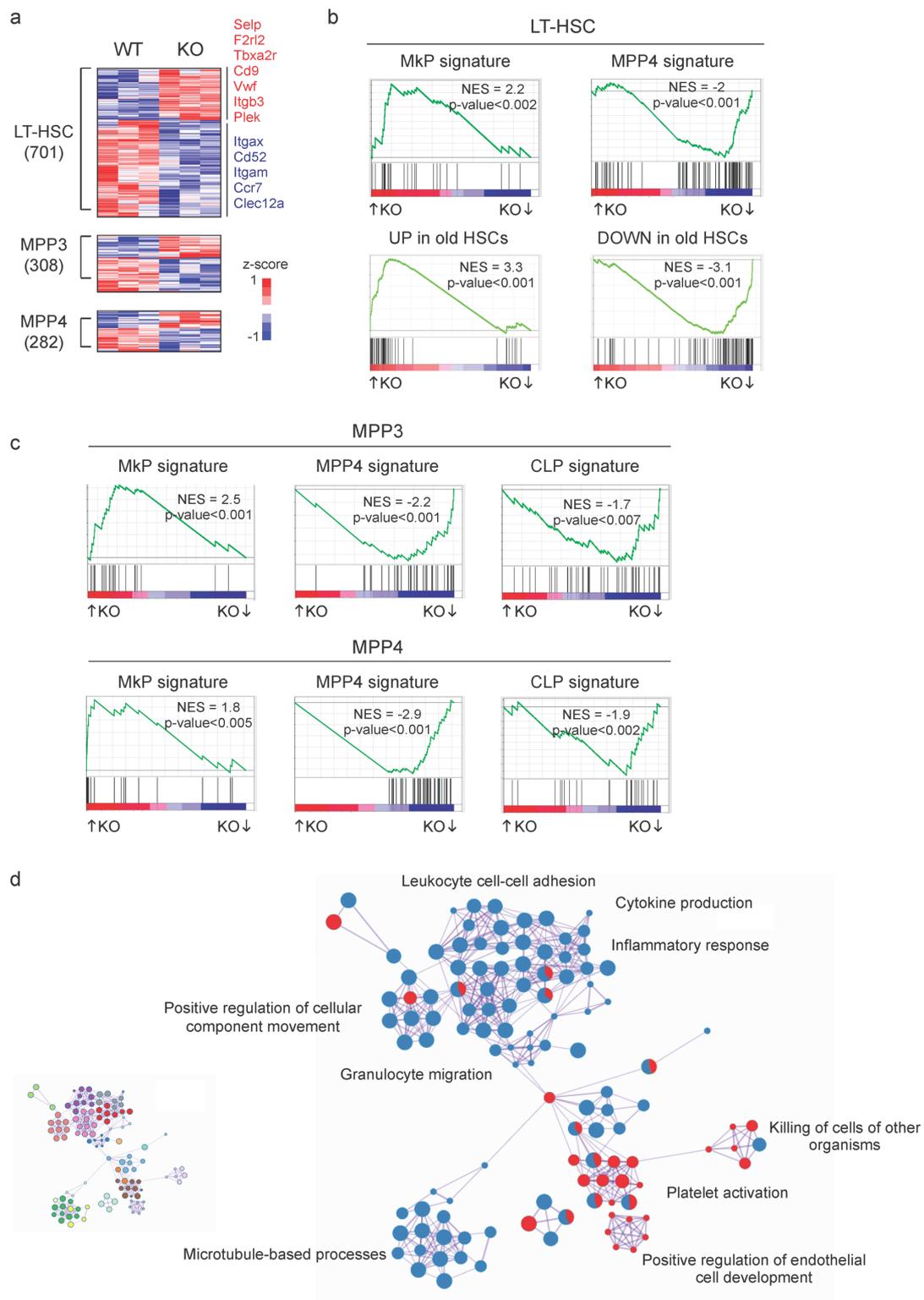
Cova et al. Figure 2



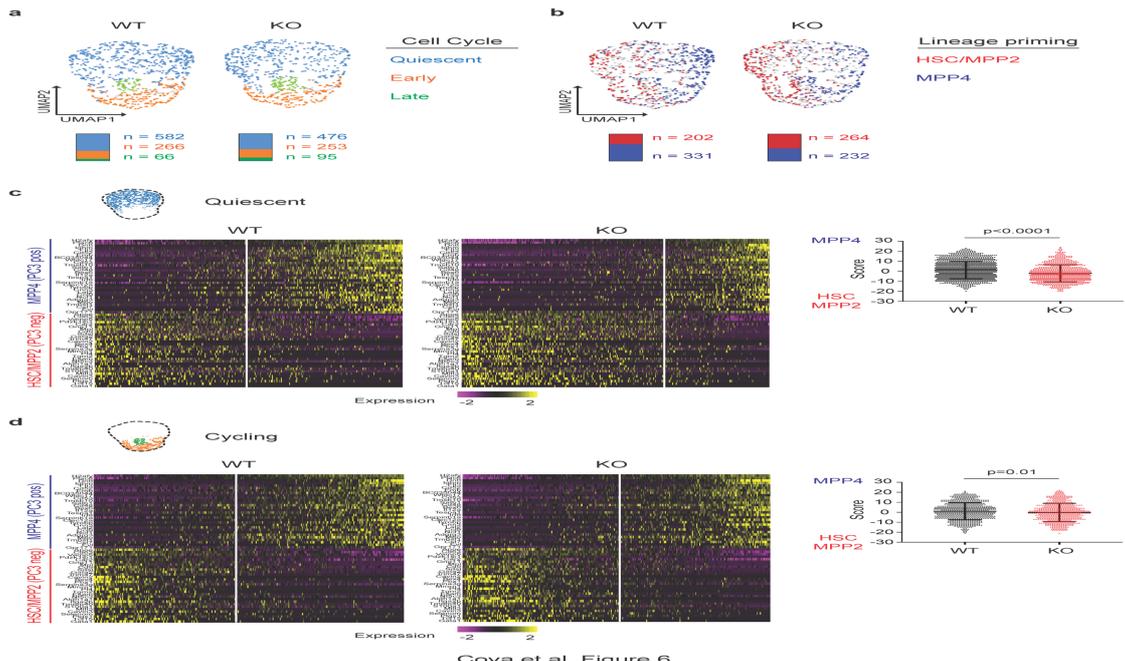
Cova et al. Figure 3



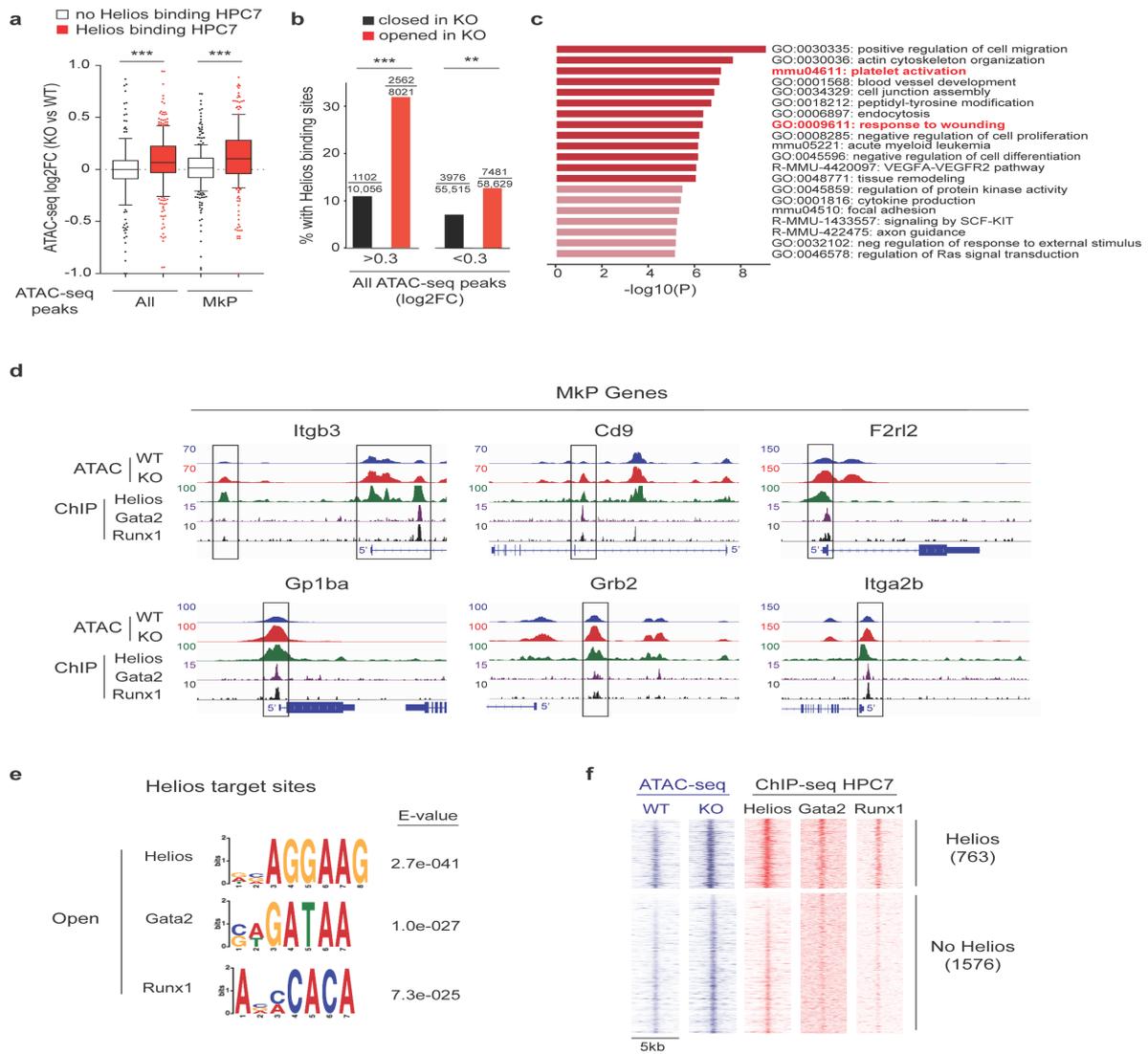
Cova et al. Figure 4



Cova et al. Figure 5



Cova et al. Figure 6



Cova et al. Figure 7

Dissection des spécificités fonctionnelles des facteurs de transcription de la famille d'Ikaros au cours du développement des lymphocytes B

Beaucoup de facteurs de transcription fonctionnent comme homo- ou hétérodimères. Cependant il n'est pas connu quelles sont les fonctions de ces différents dimères et leur spécificité de liaison à l'ADN. Nous avons adressé cette question pour les facteurs de transcription de la famille d'Ikaros exprimés au cours de la différenciation des lymphocytes B : Ikaros et Aiolos. Ces facteurs ont un domaine de liaison à l'ADN composé de quatre doigts de zinc et un domaine de dimérisation avec deux doigts de zinc permettant la formation d'homo- et d'hétérodimères. Pour définir le rôle des différents dimères dans la différenciation des lymphocytes pré-B, nous avons utilisé une lignée de lymphocytes pré-B murins déficients pour Ikaros, les cellules BH1 dans lesquelles des paires de ces protéines ont respectivement été fusionnées aux parties N- et C-terminale de la protéine Venus. L'association de V1 avec V2 crée un épitope pour un anticorps permettant l'immunoprécipitation spécifique des dimères. Tous les dimères V1/V2 sont capables de mener la différenciation des cellules BH1 comme cela avait été montré pour Ikaros. Nous avons analysé les profils de liaison à l'ADN des différents dimères et nous avons montré que l'homodimère d'Ikaros se lie à peu de régions sur le génome et que ces régions sont également liées par l'homodimère d'Aiolos et l'hétérodimère Ikaros/Aiolos. Il est possible qu'Ikaros se dimérise préférentiellement avec Aiolos dès que celui-ci est exprimé et fonctionne plus comme un hétérodimère que comme un homodimère. Cette étude montre pour la première fois des différences fonctionnelles majeures entre les membres de la famille d'Ikaros et suggère que les hétérodimères puissent être les médiateurs clés de leurs fonctions

Mots clés : Ikaros, Aiolos, dimérisation, lymphocytes B, différenciation

Many transcription factors function as homo- or heterodimers. However, it is still unknown how different patterns of dimerization shape their function, particularly with respect to DNA target site selection. We addressed this question for the Ikaros family members Ikaros and Aiolos in early B cells. These two factors are expressed during B lymphopoiesis. These proteins share a common structure that harbors a four zinc fingers (ZF) DNA binding domain and a two ZF dimerization domain and bind DNA as homo- or heterodimers. To define the role of the homo- and heterodimers between these factors, we used the Ikaros-deficient pre-B cell line BH1, in which we expressed pairs of these proteins that were respectively fused to the N- and C-terminal halves of Venus (V1 and V2). Dimerization brings V1 and V2 together and reconstitutes Venus fluorescence. Furthermore, the association of V1 and V2 generates an epitope for an antibody that can be used to immunoprecipitate the dimeric complex. All V1/V2-tagged dimers remained functional and were able to drive the differentiation of BH1 cells, as previously shown for Ikaros. We mapped the genome wide binding of each homo- and heterodimer, by ChIP-sequencing. Our data show differences in the binding patterns of the different dimers. Actually, it seems that Ikaros homodimers bind specifically a few numbers of regions and most of them are also bound by the homodimers of Aiolos and the heterodimers formed by Ikaros and Aiolos. It is believed that Ikaros preferentially dimerizes with Aiolos as soon as the latter is expressed and functions more efficiently as a heterodimer than as a homodimer. This study provides new insights into the specific and redundant functions of Ikaros family members.

Key words: Ikaros, Aiolos, dimerization, B cells, differentiation