



Université de Montréal

**Fonctions de l'oncoprotéine LMO2 déterminées par  
ses interactions protéiques**

par

Marie-Claude Sincennes

Programme de Biologie Moléculaire

Faculté de Médecine

Thèse présentée à la Faculté de Médecine  
en vue de l'obtention du grade de Docteur  
en biologie moléculaire

Octobre 2011

© Marie-Claude Sincennes, 2011

Université de Montréal

Faculté de Médecine

Cette thèse intitulée :

Fonctions de l'oncoprotéine LMO2 déterminées par ses interactions protéiques

présentée par :

Marie-Claude Sincennes

A été évaluée par un jury composé des personnes suivantes :

Muriel Aubry, Président-rapporteur

Trang Hoang, Directrice de recherche

Alain Verreault, Membre du jury

Frédéric Barabé, Examineur externe

Isabelle Royal, Représentant du Doyen

## Résumé

La leucémie lymphoïde représente environ 30% des cas de cancer chez l'enfant. Elle est souvent causée par des réarrangements chromosomiques impliquant des gènes encodant des facteurs de transcription, qui contrôlent des programmes génétiques complexes. Par exemple, LMO2 (*LIM-only 2*) est un facteur de transcription oncogénique fréquemment exprimé de façon aberrante dans les leucémies lymphoblastiques aiguës des cellules T (T-ALL). Dans l'hématopoïèse normale, LMO2 est essentiel à la génération des cellules souches hématopoïétiques à l'origine de toutes les cellules sanguines. D'ailleurs, certaines cellules leucémiques possèdent des propriétés normalement réservées aux cellules souches hématopoïétiques. Ainsi, l'étude de la fonction de LMO2 dans les cellules souches hématopoïétiques peut être pertinente autant dans le contexte hématopoïétique normal que leucémique.

Afin de mettre en évidence de nouvelles fonctions moléculaires pour LMO2, j'ai choisi d'identifier les protéines qui s'y associent. En plus de ses partenaires connus, j'ai identifié plusieurs protéines de transcription/remodelage de la chromatine, en accord avec son rôle transcriptionnel. Plusieurs nouvelles fonctions potentielles ont été révélées, indiquant que cette protéine adaptatrice pourrait faire partie de complexes non transcriptionnels, régulant d'autres processus cellulaires. Les oncogènes comme LMO2 pourraient être des régulateurs à large spectre.

Particulièrement, j'ai identifié des interactions entre LMO2 et des protéines de réplication de l'ADN. J'ai montré que LMO2 contrôle la réplication de l'ADN dans les cellules hématopoïétiques, et possiblement durant la leucémogénèse, indépendamment de son rôle transcriptionnel. Ensemble, ces études ont donc permis de révéler de nouvelles fonctions pour LMO2, et pourraient servir de paradigme pour d'autres facteurs de transcription oncogéniques, particulièrement aux autres protéines de la famille LMO, qui sont aussi des oncogènes puissants.

**Mots-clés** : Leucémie, hématopoïèse, cellule souche, LMO2, réplication de l'ADN, interactions protéiques

## Abstract

Lymphoid leukemia represents about 30% of childhood cancer cases. It is often caused by chromosomal rearrangements involving genes coding for transcription factors, controlling complex genetic programs. As an example, the oncogenic transcription factor LMO2 (LIM-only 2) is often aberrantly expressed in T cell acute lymphoblastic leukemia (T-ALL). In normal hematopoiesis, LMO2 is essential for the generation of hematopoietic stem cells that give rise to all blood cells. Moreover, some leukemic cells possess properties normally reserved to hematopoietic stem cells. Thus, studying the role of LMO2 in hematopoietic stem cells could be relevant to the contexts of normal hematopoiesis and leukemogenesis.

To reveal new molecular functions for LMO2, I chose to identify its associated proteins. In addition to its known protein partners, I identified many proteins involved in transcription/chromatin remodeling, in agreement with its transcriptional role. In addition, several new potential functions have been revealed, indicating that this scaffold protein could be part of non-transcriptional protein complexes, regulating different cell processes. Oncogenes like LMO2 could be master regulators in normal hematopoietic and leukemic cells.

Particularly, I identified protein-protein interactions between LMO2 and DNA replication proteins. I demonstrated that LMO2 controls S phase progression in hematopoietic cells, independently of its association in transcriptional complexes. LMO2 overexpression in mice induces T-ALL and affects specifically the cell cycle status of thymocyte progenitors, which are targets of transformation by LMO2. Thus, LMO2 promotes DNA replication in hematopoietic cells, and possibly in leukemogenesis. Together, these studies allowed to reveal new functions for LMO2, and could serve as a paradigm for other oncogenic transcription factors, especially for other LMO proteins which are all potent oncogenes.

**Keywords** : Leukemia, hematopoiesis, stem cell, LMO2, DNA replication, protein-protein interactions

## Table des matières

Résumé	iii
Abstract	iv
Liste des figures	x
Liste des tableaux	xii
Liste des abréviations	xiii
Remerciements	xv
Avant-propos	xvii
<b>1. INTRODUCTION</b>	<b>1</b>
1.1 Hématopoïèse	1
1.1.1 Établissement du système hématopoïétique	2
1.1.2 Homéostasie du système hématopoïétique (la cellule souche hématopoïétique)	3
1.1.2.1 Cycle cellulaire et division des cellules souches hématopoïétiques	3
1.1.2.2 La niche hématopoïétique (microenvironnement)	4
1.1.3 SCL et LMO2 comme régulateurs essentiels de l'hématopoïèse	5
1.1.3.1 Expression de SCL et LMO2 durant l'hématopoïèse	6
1.1.3.2 Le rôle essentiel de SCL et LMO2 dans l'hématopoïèse	6
1.2 Cancer et leucémie	12
1.2.1 Propriétés d'une cellule cancéreuse : le cancer en tant que processus multi-étapes	12
1.2.1.1 Intégrité génomique	13
1.2.1.2 Prolifération cellulaire	14
1.2.1.3 Survie cellulaire et apoptose	15
1.2.1.4 Potentiel répliatif illimité	15
1.2.2 La cellule souche cancéreuse	16
1.2.3 La leucémie de type T-ALL	17
1.2.3.1 Les principaux oncogènes associés à la leucémie T-ALL	18

1.2.3.2 LMO2 et SCL en tant qu'oncogènes _____	21
1.3 Le gène LMO2 _____	26
1.3.1 Régulation de l'expression de LMO2 _____	26
1.3.1.1 Régulation transcriptionnelle de LMO2 _____	26
1.3.1.2 Régulation des niveaux protéiques de LMO2 par l'action de micro-ARNs _____	29
1.3.1.3 Régulation des niveaux protéiques de LMO2 par l'ubiquitination _____	31
1.3.2 Les protéines à domaines LIM _____	31
1.3.3 Les protéines de la famille LMO _____	34
1.3.4 Le rôle transcriptionnel de LMO2 : interactions protéiques et gènes cibles _____	38
1.3.5 Possibilité d'un rôle non-transcriptionnel pour LMO2 _____	41
1.4 Problématique et choix de la méthodologie _____	43
1.5 Objectifs de la thèse _____	45
<b>2. CROSS-REGULATION OF TRANSCRIPTION AND REPLICATION, TWO DNA-DEPENDENT PROCESSES</b>	<b>47</b>
2.1 Contribution des co-auteurs _____	48
2.2 Summary _____	49
2.3 Introduction _____	49
2.4 Proximity between replication origins and gene promoters _____	50
2.5 Epigenetic factors _____	54
2.6 Origin selection: another task for transcription factors? _____	55
2.7 c-MYC _____	57
2.8 MYB _____	59
2.9 RB and E2F _____	60
2.10 HOX proteins _____	61
2.11 Concluding remarks and perspectives _____	62
2.11.1 Are transcription factors important for origin	

selection/activation? _____	63
2.11.2 Early versus late origins _____	64
2.11.3 Oncogenic transcription factors as regulators of DNA replication _____	64
2.12 Acknowledgements _____	66
<b>3. LMO2 REGULATES DNA REPLICATION IN HEMATOPOIETIC CELLS</b>	
3.1 Contribution des co-auteurs _____	68
3.2 Summary _____	69
3.3 Introduction _____	69
3.4 Results _____	71
3.4.1 LMO2 interacts with DNA replication proteins _____	71
3.4.2 LMO2 levels vary during the cell cycle _____	74
3.4.3 LMO2 colocalizes with replication proteins _____	74
3.4.4 LMO2 binds to DNA replication origins _____	77
3.4.5 S phase progression depends on LMO2 levels in hematopoietic cells _____	79
3.4.6 LMO2 overexpression increases the percentage of thymocyte progenitors in S phase _____	83
3.5 Discussion _____	86
3.5.1 A novel role for LMO2 in S phase regulation _____	86
3.5.2 LMO2 as a T-lineage oncogene _____	87
3.5.3 Cross-talk between transcription and DNA replication _____	88
3.6 Experimental procedures _____	90
3.7 Acknowledgements _____	93
3.8 Supplementary data _____	94
<b>4. SYSTEMATIC MAPPING OF PROTEIN-PROTEIN INTERACTIONS FOR THE ONCOPROTEIN LMO2</b>	<b>100</b>
4.1 Contribution des co-auteurs _____	101
4.2 Summary _____	102
4.3 Introduction _____	102
4.4 Results _____	104
4.5 Discussion _____	113



4.5.1 Comparison of SCL and LMO2-interacting partners _____	113
4.5.2 The LMO2 interactome: more than transcription _____	114
4.5.3 LMO2 protein interactions in normal and leukemic stem cells _____	115
4.6 Experimental procedures _____	116
4.7 Acknowledgements _____	117
4.8 Supplementary data _____	118
<b>5. DISCUSSION</b>	<b>122</b>
5.1 Librairie d'ADNc de progéniteurs hématopoïétiques : un outil unique et puissant _____	122
5.1.1 Cellules souches et progéniteurs hématopoïétiques comme modèle biologique _____	122
5.1.2 Bibliothèques de cDNA commercialement disponibles et leurs limitations	125
5.1.3 Utilisation de la librairie de cDNA dans l'étude des cellules souches hématopoïétiques _____	126
5.2 Nouveaux partenaires potentiels de LMO2 _____	128
5.2.1 Identification de partenaires connus de LMO2 par double hybride _	129
5.2.2 LMO2 : petite protéine, grand nombre de partenaires _____	129
5.2.3 Les protéines à domaines LIM pourraient réguler des processus cellulaires semblables _____	131
5.2.4 Possibilités de fonctions non-transcriptionnelles pour LMO2 _____	132
5.2.4.1 Lien potentiel entre LMO2 et les protéines liant l'ARN ____	134
5.2.4.2 Lien potentiel entre LMO2 et les protéines du cytosquelette	134
5.2.4.3 Lien potentiel entre LMO2 et les protéines mitochondriales	135
5.3 LMO2 comme régulateur de la réplication de l'ADN _____	136
5.3.1 LMO2 pourrait faciliter l'assemblage du complexe de pré-initiation de la réplication de l'ADN _____	136
5.3.2 LMO2 pourrait influencer le choix des origines de réplication au moment de leur activation _____	137
5.3.3 LMO2 pourrait modifier l'état de la chromatine aux origines de réplication _____	138

5.3.4 Autres possibilités	139
5.4 Lien entre transcription, réplication et cancer	140
5.4.1 Les facteur de transcription qui influencent la réplication de l'ADN	140
5.4.2 Réplication de l'ADN et instabilité génomique	142
CONCLUSION	145
Bibliographie	146
Annexe 1 : Boucle d'auto-régulation entre LMO2 et deux microARNs	xix
Annexe 2 : Protein Stability and Transcription Factor Complex Assembly Determined by the SCL-LMO2 Interaction	xxi
Annexe 3: LMO2 et SCL partagent des partenaires protéiques avec la protéine de fusion oncogénique NUP98-HOXA9	xxxii

## Liste des figures

Figure 1.1	Schéma de l'hématopoïèse _____	1
Figure 1.2	Fonctions de SCL et LMO2 dans le système hématopoïétique ____	7
Figure 1.3	Organisation du locus <i>LMO2</i> chez l'humain _____	27
Figure 2.1	Parallel between transcription factor activity and origin activity _	56
Figure 3.1	LMO2 interacts with DNA replication proteins in chromatin ____	72
Figure 3.2	LMO2 protein levels are regulated during the cell cycle ____	75
Figure 3.3	LMO2 colocalizes with MCM5 and MCM6 in the nuclei of TF-1 cells, during G0/G1 and the G1/S transition _____	76
Figure 3.4	LMO2 and MCM5 occupy DNA replication origins ____	78
Figure 3.5	Disruption of LMO2 levels impairs S phase progression in murine erythroleukemia (MEL) cells _____	80
Figure 3.6	LMO2 overexpression specifically affects the cell cycle status of early thymic progenitors <i>in vivo</i> _____	84
Figure 3.7	LMO2 specifically interacts with replication proteins ____	94
Figure 3.8	Maps depicting the location of well characterized DNA replication origins and PCR probes _____	95
Figure 3.9	Chromatin immunoprecipitation of TF-1 cells with antibodies against MCM5 or LMO2 _____	96
Figure 3.10	MEL cells can be synchronized in G0/G1 by by flow cytometry	98
Figure 3.11	Cells with reduced levels of LMO2 are not apoptotic ____	99
Figure 3.12	Representative flow cytometry analysis of MSCV-LMO2 induced leukemia cells, compared to cells from a healthy mouse ____	99
Figure 4.1	Identification of SCL and LMO2 protein partners _____	106
Figure 4.2	Protein interactions involving LMO2 could be detected by different techniques _____	108

Figure 4.3	LMO2 partners are involved in a complex protein-protein interaction network	109
Figure 4.4	Identification of LMO2 and SCL partners associated on their target genes	111
Figure 4.5	Construction of a c-Kit <sup>+</sup> Lin <sup>-</sup> bone marrow stem/progenitor cell cDNA library	118
Figure 4.6	Protein-protein interaction networks involving 133 Osprey-available random proteins	119
Figure A.1	Boucle d'auto-régulation entre LMO2 et deux microARNs	xix
Figure A.2	Partenaires protéiques communs entre LMO2, SCL et NUP98-HOXA9, identifiés par double hybride	xxxii

## Liste des tableaux

Tableau I	Gènes les plus fréquemment réarrangés ou mutés dans les leucémies T-ALL chez l'humain _____	18
Table II	Transcription factors influencing DNA replication _____	58
Table III	List of binding partners identified from the LMO2 yeast two hybrid screen _____	120
Tableau IV	Partenaires potentiels de LMO2 interagissant avec d'autres protéines à domaines LIM _____	132

## Liste des abréviations

ADN	Acide déoxyribonucléique
AGM	Aorte-gonade-mésonéphros
AML	Leucémie myéloïde aigue
ARNm	ARN (acide ribonucléique) messenger
ATM	<i>Ataxia telangiectasia mutated</i>
B-ALL	Leucémie lymphoblastique aiguë des cellules B
bHLH	<i>Basic helix-loop-helix</i>
BrdU	5-bromodeoxyuridine
cDNA	ADN complémentaire
ChIP	Immunoprécipitation de la chromatine
ChIP-chip	ChIP couplée à une analyse par micropuces
ChIP-seq	ChIP couplée à une analyse par séquençage
CLP	<i>Common lymphoid progenitor</i>
CRP	<i>Cystein-rich protein</i>
DN	Cellules doublement négatives CD4 <sup>-</sup> CD8 <sup>-</sup>
DP	Cellules doublement positives CD4 <sup>+</sup> CD8 <sup>+</sup>
EMSA	<i>Electrophoretic mobility shift assay</i>
ES	<i>Embryonic stem cells</i>
ETP	<i>Early thymic progenitor</i>
ETS	<i>E26 transformation specific sequence</i>
FDCP mix	<i>Factor-dependent cell Paterson mixed potential</i>
FHL	<i>Four-and-a-half LIM</i>
G-CSF	<i>Granulocyte-colony stimulating factor</i>
GO	<i>Gene Ontology</i>
GST	Glutathione-S-transferase
Ig	Immunoglobuline
IP	Immunoprécipitation
Kb	Kilobase
kDa	Kilodalton

LDB1	Protéine <i>LIM domain-binding protein 1</i>
LIM	Domaine protéique (LIN11, ISL1, MEC3)
LMO	<i>LIM-Only</i>
miR	micro-ARN
MEL	<i>Mouse erythroleukemia cells</i>
MCM	<i>Minichromosome maintenance protein</i>
MLL	<i>Mixed-lineage leukemia</i>
ORC	<i>Origin recognition complex</i>
PAR	<i>Proline and acidic amino acid-rich domain</i>
PCNA	<i>Proliferating cell nuclear antigen</i>
PCR	<i>Polymerase chain reaction</i>
Pre-IC	Complexe de pré-initiation de la réplication
Pre-RC	Complexe de pré-réplication
RB	Protéine du rétinoblastome
RT-PCR	<i>Reverse transcriptase – Polymerase chain reaction</i>
SCL	<i>Stem Cell Leukemia</i>
shRNA	<i>Short hairpin RNA</i>
T-ALL	Leucémie lymphoblastique aiguë des cellules T
TCR	<i>T cell receptor</i>
Tg	Transgénique
TNF	<i>Tumor necrosis factor</i>
UTR	<i>Untranslated region</i>
X-SCID	<i>X-linked severe combined immunodeficiency</i>

## Remerciements

Je tiens d'abord à remercier ma directrice de recherche, Dr Trang Hoang, pour m'avoir soutenue et encouragée tout au long de mon cheminement en vue de l'obtention de mon Ph.D. Vous avez toujours été à l'écoute et surtout ouverte à mes idées et mes opinions, tant d'un point de vue scientifique que personnel. Je vous remercie de l'intérêt que vous avez toujours porté à l'égard de mon projet de recherche. Vous avez su révéler mon potentiel et m'offrir des défis qui m'ont permis de me surpasser.

Je remercie chaleureusement tous les membres passés et présents du laboratoire qui ont tous contribué à mon apprentissage, chacun à leur façon. Merci à tous pour nos discussions, tant scientifiques qu'humoristiques, qui m'ont permis de voir les choses sous des perspectives différentes. Je remercie particulièrement Richard Martin, qui m'a accueillie dans le laboratoire, m'a appris les rudiments de l'hématopoïèse et de la manipulation de levures, et qui m'a donné confiance dès le départ par ses nombreux encouragements. Merci à Benoit et à André, qui ont été là du début à la fin de mes études doctorales dans le laboratoire. Merci à Mathieu : malgré nos différences, mon doctorat n'aurait pas été le même sans toi. Merci à tous ceux et celles avec qui j'ai partagé fous rires et confidences au labo, particulièrement Annie, les deux Julie, Stéphane et Cédric. Merci aux stagiaires que j'ai supervisé, pour votre contribution à l'avancement de mes recherches : Alexandre, Déborah, Francis. Merci à Véronique Lisi, qui est passée de collaboratrice à amie. Merci à Bastien, Cédric, Shanti et Véronique, pour avoir lu une grande partie de cette thèse. Vos commentaires, tous pertinents, ont été très appréciés. Merci à toute la grande famille de l'IRIC pour l'entraide scientifique et le caractère festif de nos nombreuses activités.

Merci à ma famille et mes amis pour votre soutien pendant ses longues études. Merci à Jean-François pour tes encouragements et tes conseils. Je suis chanceuse de partager ma vie avec quelqu'un qui comprend ce que je fais! Merci de ton écoute, de ta disponibilité, de ton humour et de ta grande générosité. Merci à ma fille Laurence, et à mon fils qui arrive, pour m'avoir permis de relativiser mes succès, mes échecs et



mon stress relatif à mon parcours académique. Merci à vous trois pour votre amour inconditionnel qui est si réconfortant. Merci à mes parents, Lise et Roch, ainsi qu'à ma sœur Amélie, pour votre disponibilité, pour avoir partagé avec moi mes moments de bonheur et de découragement, et pour m'avoir offert tout ce support moral si apprécié. Merci à la grande famille Carle/Carmel/Sincennes, de même qu'à mes amies Katherine et Marie-Christine, pour avoir été mes fans profanes (!). Merci aux filles du bac, qui attendent patiemment leur poste de préposée au café ou à la photocopieuse dans mon futur laboratoire!

Enfin, je tiens à souligner l'apport des différents organismes subventionnaires qui ont soutenu mes recherches : Fonds de la Recherche en Santé du Québec (FRSQ), Instituts de Recherche en Santé du Canada (IRSC), Fondation Cole, Programmes de Biologie Moléculaire de l'Université de Montréal. Je désire finalement remercier les membres du jury pour avoir accepté d'évaluer cette thèse.

## Avant-propos

La cellule a mis en place plusieurs mécanismes de contrôle qui lui permettent de réagir adéquatement en réponse à différents événements : signaux de prolifération ou anti-prolifératifs, dommages à l'ADN, stress cellulaire, etc. Le cancer se développe quand plusieurs de ces mécanismes de contrôle sont altérés, ce qui permet à la cellule de survivre et de proliférer de façon aberrante. Certains gènes codent pour des protéines qui sont des régulateurs à large spectre dans la cellule, contrôlant plusieurs différents processus cellulaires. Une altération dans ces gènes peut avoir de profondes conséquences sur la cellule, et ultimement favoriser l'émergence du cancer. Par exemple, la leucémie est souvent associée à des altérations activatrices dans des gènes codant pour des facteurs de transcription. Les facteurs de transcription, de par leur association à l'ADN, déterminent quels gènes seront exprimés ou réprimés dans un contexte spatio-temporel précis. C'est pourquoi un changement dans l'activité d'un facteur de transcription particulier peut avoir un impact déterminant sur l'expression génique et donc sur le devenir de la cellule. Pour ces raisons, il est crucial d'étudier les mécanismes d'action des facteurs de transcription oncogéniques, afin de développer des thérapies plus efficaces et ciblées.

Dans cette optique, le facteur de transcription LMO2 est un bon modèle d'études : 1) LMO2 est un des facteurs de transcription les plus fréquemment activés dans les leucémies lymphoblastiques aiguës des cellules T chez l'enfant; 2) On connaît en partie le rôle de LMO2 dans les cellules hématopoïétiques normales, ce qui permet de faire des parallèles avec le contexte leucémique; 3) La leucémie causée par l'expression ectopique de LMO2 dans les cellules T est bien caractérisée au niveau cellulaire, autant chez l'humain que dans des modèles de souris. Ainsi, il est possible d'étudier efficacement les mécanismes moléculaires par lesquels LMO2 transforme la cellule.

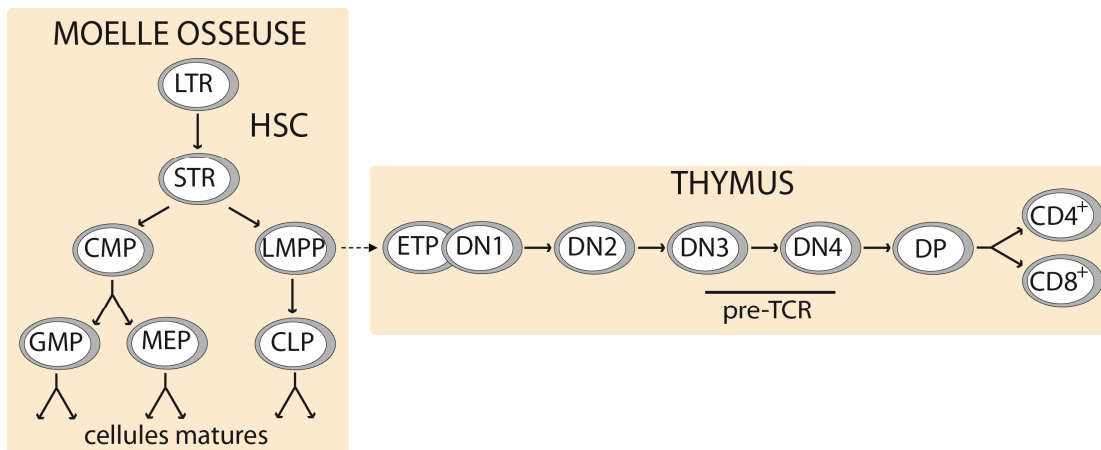
Précédemment, l'étude des gènes cibles de LMO2 a permis d'expliquer partiellement certains effets cellulaires associés à l'expression ectopique de LMO2 dans les cellules T. LMO2 régule l'expression de gènes associés à la différenciation thymocytaire, et

ainsi, l'expression aberrante de LMO2 perturbe la différenciation normale des cellules T. Toutefois, l'identification de gènes cibles n'a pas permis de cerner tous les mécanismes moléculaires par lesquels LMO2 transforme la cellule : par exemple, comment LMO2 engendre-t-il une prolifération accrue des thymocytes? Ainsi, en plus de sa fonction en tant que facteur de transcription, LMO2 pourrait transformer la cellule par d'autres mécanismes moléculaires. De la même façon, d'autres facteurs de transcription oncogéniques, particulièrement les protéines de la famille LMO, pourraient transformer la cellule par des mécanismes d'action semblables. Ainsi, la compréhension du potentiel oncogénique de LMO2 pourrait servir de modèle dans l'étude d'autres oncogènes associés à des leucémies ou même à d'autres types de cancers. Par ailleurs, mes découvertes pourraient aussi préciser le rôle de LMO2 dans l'hématopoïèse normale, particulièrement dans la biologie des cellules souches, qui demeure un domaine mal compris malgré son grand potentiel thérapeutique. Mes recherches doctorales, visant à attribuer de nouvelles fonctions à LMO2 par l'identification de ses partenaires protéiques, s'inscrivent dans cette grande ligne de pensée.

# 1 Introduction

## 1.1 Hématopoïèse

L'hématopoïèse est le processus par lequel les cellules sanguines sont produites et constamment régénérées. Ce processus est initié par la cellule souche hématopoïétique, qui se différencie en divers types de cellules appelées progéniteurs, générant à leur tour les cellules sanguines matures (**Figure 1.1**).



**Figure 1.1** Schéma de l'hématopoïèse.

L'hématopoïèse repose sur les cellules souches au pouvoir de reconstitution à long terme (LTR), qui perdent leur caractère multipotent et leur potentiel d'auto-renouvellement alors qu'elles se différencient. Une sous-population de progéniteurs, possiblement les LMPP (*lymphoid-primed multipotent progenitor*), colonisent le thymus et génèrent les lymphocytes T. HSC : cellule souche hématopoïétique; STR : cellules souches responsables de la reconstitution à court terme; CMP : progéniteur commun myéloïde; GMP : progéniteur granulocyttaire et macrophagique; MEP : progéniteur mégakaryocytaire et érythrocytaire; CLP : progéniteur commun lymphoïde; ETP : progéniteur thymique précoce; DN : cellule double négative CD4<sup>-</sup>CD8<sup>-</sup>; DP : cellule double positive CD4<sup>+</sup>CD8<sup>+</sup>; TCR : *T-cell receptor*.

Une cellule souche possède deux caractéristiques essentielles : elle doit être multipotente, c'est-à-dire qu'elle possède le potentiel de se différencier en plusieurs types de cellules matures, et elle doit avoir une capacité d'auto-renouvellement, c'est-à-dire qu'elle peut générer une autre cellule-souche afin de maintenir constant le réservoir de cellules souches d'un individu au cours de sa vie [1]. Au niveau expérimental, les propriétés d'une cellule souche hématopoïétique sont évaluées par

sa capacité de reconstituer le système hématopoïétique dans un hôte irradié ou immuno-déficient, dépourvu de cellules hématopoïétiques endogènes.

Chez l'humain adulte, l'hématopoïèse est effectuée dans la moelle osseuse. Le sang contient une variété de types cellulaires ayant une durée de vie et des fonctions très variées. Les cellules érythroïdes sont responsables du transport de l'oxygène et ont une durée de vie d'environ 120 jours. Les cellules myéloïdes (granulocytes, macrophages, cellules dendritiques) ont une durée de vie très courte, de l'ordre de quelques jours. Ces cellules sont capables de phagocytose et/ou sont des cellules présentatrices d'antigènes. Les cellules lymphoïdes (lymphocytes B et T) peuvent vivre durant plusieurs années et jouent un rôle crucial dans la réponse immunitaire. L'hématopoïèse est régulée de façon à produire suffisamment, mais sans excès, ces différents types de cellules sanguines.

### **1.1.1 Établissement du système hématopoïétique**

Chez l'humain, le système hématopoïétique est mis en place à la 8<sup>e</sup> journée post-conception, bien avant la formation des os et donc de la moelle osseuse. D'autres organes ont un potentiel hémogénique au cours de l'embryogénèse. Durant le développement, les cellules sanguines dérivent du mésoderme. D'abord, une hématopoïèse dite primitive est amorcée dans le sac vitellin [2] et ne fournit que des cellules sans potentiel d'auto-renouvellement, comme des érythrocytes primitifs, des macrophages et des progéniteurs [2]. Différents sites anatomiques présentent un potentiel hémogénique intermédiaire durant le développement, comme la région AGM (aorte-gonade-mésonephros) et le placenta. Ceux-ci sont responsables de la production des cellules souches hématopoïétiques [2, 3], amorçant une hématopoïèse dite définitive. Ces cellules souches hématopoïétiques vont coloniser le foie fœtal et prendre de l'expansion. Le foie fœtal assume jusqu'à la naissance une hématopoïèse définitive avec une organisation et une hiérarchie telle que celle de l'adulte. Au moment de la naissance, les cellules souches hématopoïétiques colonisent la moelle osseuse, qui devient le site d'hématopoïèse de l'individu pour le reste de sa vie [2].

### **1.1.2 Homéostasie du système hématopoïétique (la cellule souche hématopoïétique)**

Afin de maintenir constant le nombre des différentes cellules sanguines d'un individu, un équilibre doit être établi entre l'auto-renouvellement et la différenciation des cellules souches hématopoïétiques. Bien que d'autres mécanismes de contrôle soient mis en place lors de la différenciation des progéniteurs multipotents et lors du choix de lignage hématopoïétique (myéloïde vs lymphoïde, lymphocytes T vs lymphocytes B, etc.), les cellules souches sont au sommet de la hiérarchie hématopoïétique. Un changement dans leur comportement a une conséquence majeure dans l'homéostasie du système sanguin. La biologie des cellules souches est contrôlée, entre autres, par l'action de différents facteurs de transcription.

Les facteurs de transcription SCL et LMO2, dont il sera question dans cette thèse, sont fortement exprimés dans les cellules souches hématopoïétiques. Ils sont nécessaires à l'émergence de ces cellules lors du développement, et sont aussi importants pour leur maintenance à l'âge adulte. D'autre part, dans la leucémie et plusieurs autres types de cancers, il existe des cellules souches cancéreuses qui sont responsables d'initier et de maintenir la maladie, et ces cellules transformées semblent avoir réacquis certaines propriétés des cellules souches normales. Pour ces raisons, les propriétés des cellules souches hématopoïétiques seront brièvement résumées. Les cellules souches hématopoïétiques ont un cycle cellulaire particulier et ont la propriété de se diviser de façon symétrique et asymétrique. De plus, ces cellules reçoivent différents signaux extracellulaires provenant de leur environnement (niche) qui activent des programmes transcriptionnels favorisant leur quiescence, leur auto-renouvellement ou leur différenciation en réponse à divers stimuli.

#### **1.1.2.1 Cycle cellulaire et division des cellules souches hématopoïétiques**

En conditions d'homéostasie, la plupart des cellules souches hématopoïétiques (environ 75%) sont quiescentes, c'est-à-dire au stade G0 du cycle cellulaire. Elles ne se divisent que très rarement [4-8] afin de maintenir constant le nombre de cellules

matures dans le sang et le nombre de cellules souches dans la moelle osseuse. À la fin de son cycle cellulaire, la cellule souche hématopoïétique a le potentiel de se diviser de trois façons différentes : de façon asymétrique, pour générer une autre cellule souche de même qu'une cellule plus différenciée; ou de façon symétrique, pour générer soit deux cellules différenciées (symétrie d'engagement), soit deux cellules souches identiques à la cellule-mère (symétrie de renouvellement) [1, 9]. Dans certaines conditions (développement embryonnaire, perte sanguine, maladie), un type de division peut être favorisé plutôt qu'un autre. Certains oncogènes (ex. *Nup98-HoxA9* [10]) influencent le type de division effectuée par la cellule souche.

Lorsque l'on parle de division asymétrique de la cellule souche, on se réfère au destin cellulaire différent pour les deux cellules-filles produites. La division asymétrique peut être intrinsèque, par exemple en redistribuant inégalement certains déterminants cellulaires entre les deux cellules-filles, ou extrinsèque, qui serait le résultat d'une « asymétrie environnementale » [9]. En effet, une des deux cellules-filles peut garder ses propriétés souches en demeurant dans la niche hématopoïétique (le site de résidence des cellules souches), alors que l'autre cellule quitte la niche et s'expose à un environnement favorisant la différenciation. Cela a été démontré chez la drosophile, dans les cellules souches germinales [11]. Pour les cellules souches hématopoïétiques, bien que certaines protéines soient distribuées inégalement entre les deux cellules-filles [10, 12], le stroma avec lequel la cellule souche entre en contact dicte le type de division cellulaire préférentiellement adopté par la cellule souche [10], suggérant que les deux types d'asymétrie sont présents.

### **1.1.2.2 La niche hématopoïétique (microenvironnement)**

Une niche de cellules souches est définie par un ensemble de tissus et de matrice extracellulaire *in vivo* près desquels se logent les cellules souches, et via lesquels ces dernières reçoivent des signaux régulant leur auto-renouvellement et leur différenciation. Chez les mammifères, de telles niches ont été identifiées pour les cellules souches hématopoïétiques, neurales, de l'épiderme et de l'intestin [13].

Pour les cellules souches de la moelle osseuse, la niche se trouve dans l'endosteum (région trabéculaire) de l'os [14-17], de même que dans l'endothélium sinusoidal (réseau vasculaire) de la moelle [16, 18]. Les cellules de la niche, de même que les cellules souches hématopoïétiques, expriment à leur surface cellulaire une combinaison de ligands et de récepteurs qui régulent l'adhésion des cellules souches à leur niche, influençant les propriétés intrinsèques des cellules souches hématopoïétiques [19]. La niche envoie aussi des signaux qui bloquent la division cellulaire et favorisent la quiescence des cellules souches [20], pour conserver leur potentiel de reconstitution à long terme. D'autre part, la moelle osseuse est hypoxique [21], et il semble que les niveaux d'oxygène dans la niche contrôlent, du moins en partie, les propriétés des cellules souches [22-25].

La cellule souche peut quitter sa niche; soit pour se différencier, soit dans des cas peu fréquents pour entrer dans la circulation sanguine (processus appelé « mobilisation »). La mobilisation des cellules souches hématopoïétiques est d'ailleurs provoquée en clinique pour des transplantations de moelle osseuse, principalement par stimulation avec la cytokine G-CSF (« *granulocyte-colony stimulating factor* ») [19]. L'interaction entre la cellule souche et la niche hématopoïétique est donc un processus dynamique, influencé entre autres par le remodelage osseux [26], ou même par certains signaux provenant du système nerveux sympathique [27, 28]. La mobilisation des cellules souches hématopoïétiques dans le sang pourrait être utile, par exemple, à la réparation de tissus ou à la défense de l'organisme.

### **1.1.3 SCL et LMO2 comme régulateurs essentiels de l'hématopoïèse**

Certains facteurs de transcription jouent un rôle primordial dans la mise en place du système hématopoïétique. C'est le cas de SCL et de LMO2 dont il sera question dans cette thèse. SCL et LMO2 sont exprimés très tôt lors du développement embryonnaire, et des manipulations génétiques chez la souris ont permis de mettre en lumière un rôle essentiel pour SCL et LMO2 dans l'hématopoïèse, autant primitive que définitive, et dans la spécification des cellules souches hématopoïétiques.



### 1.1.3.1 Expression de SCL et LMO2 durant l'hématopoïèse

Les patrons d'expression de LMO2 et de SCL ont été déterminés par différentes méthodes (*Northern blot*, RT-PCR, hybridation *in situ*, immuno-histochimie). De nombreuses équipes ont déterminé l'expression de SCL et de LMO2 dans différents tissus, mais aussi dans différentes populations hématopoïétiques purifiées à l'aide de combinaisons de marqueurs extracellulaires. Toutes ces études convergent vers des patrons d'expression précis pour SCL et LMO2, qui sont résumés ici.

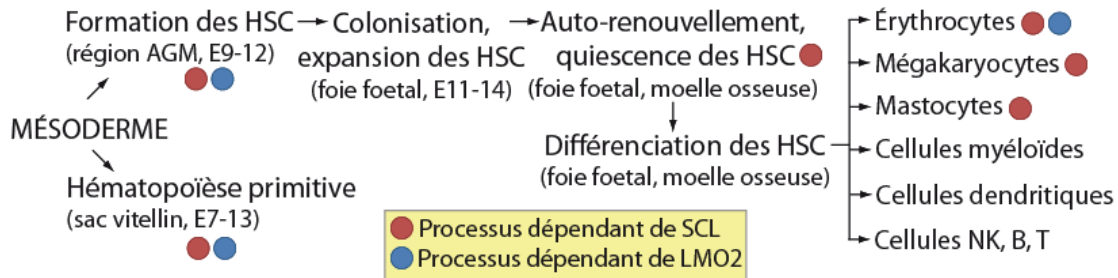
Au cours du développement, l'expression de SCL et de LMO2 est détectée dans le mésoderme embryonnaire et extra-embryonnaire au jour E7.5, dans le sac vitellin et les autres tissus hématopoïétiques en formation [29-34]. Chez la souris et l'humain adulte, l'expression de SCL et de LMO2 est très forte dans les cellules souches hématopoïétiques et les progéniteurs multipotents [35, 36], et leur expression diminue à mesure que la différenciation cellulaire progresse, jusqu'à être indétectable dans les cellules matures [31, 33]. Par exemple, dans la lignée T, l'expression de SCL et de LMO2 est détectée dans le thymus aux stades de différenciation les plus précoces (CD4<sup>+</sup>CD8<sup>-</sup>, DN), et non dans les cellules plus différenciées [37]. Toutefois, l'expression de SCL et de LMO2 est maintenue dans certains lignages, c'est-à-dire au cours de la différenciation érythrocytaire, mégakaryocytaire et mastocytaire [31, 33, 38, 39]. L'expression de LMO2 est aussi maintenue dans un type particulier de lymphocytes B (cellules B du centre germinatif) [40].

Mise à part leur expression dans les tissus hématopoïétiques, SCL et LMO2 sont aussi exprimés dans les cellules endothéliales [33, 34, 41] et dans le cerveau [33, 34, 38, 42-46]. L'expression de SCL n'est pas détectée dans d'autres organes, alors que LMO2 semble être exprimé de façon ubiquitaire, dans une variété de tissus, incluant les poumons, le foie, le rein, l'intestin et les ovaires [44].

### 1.1.3.2 Le rôle essentiel de SCL et LMO2 dans l'hématopoïèse

SCL et LMO2 jouent un rôle crucial dans l'établissement du système hématopoïétique, et ce rôle est conservé chez plusieurs autres organismes dont le système hématopoïétique se compare au nôtre, comme *Xenopus laevis* et le poisson-

zèbre [47-50]. Les études chez la souris seront résumées dans cette section, caractérisant les rôles de SCL et de LMO2 dans l'hématopoïèse de l'embryon et de l'adulte. Les fonctions de SCL et de LMO2 dans le système hématopoïétique sont aussi résumées sous forme de schéma dans la **Figure 1.2**.



**Figure 1.2 Fonctions de SCL et LMO2 dans le système hématopoïétique**

La formation du système hématopoïétique se fait à partir du mésoderme. Une hématopoïèse primitive est amorcée dans le sac vitellin, suivie de l'hématopoïèse définitive dans la région AGM (aorte-gonade-mésonephros) et le foie foetal, caractérisée par l'apparition de HSCs (cellules souches hématopoïétiques). Ces HSCs vont coloniser le foie foetal et la moelle osseuse et se différencier en cellules hématopoïétiques matures. Les étapes dépendantes de SCL et de LMO2 sont indiquées. L'effet de LMO2 n'a pas été investigué dans les HSC adultes, ni dans tous les types de cellules matures.

Les souris hétérozygotes pour *Lmo2* ou *Scl* ne présentent pas de phénotype apparent, alors que les souris nulles pour *Lmo2* ou *Scl* meurent au jour embryonnaire E9.5 [31, 51, 52], à cause de l'absence d'hématopoïèse primitive. Le sac vitellin et les embryons des souris *Lmo2*<sup>-/-</sup> ou *Scl*<sup>-/-</sup> ne contiennent pas d'érythrocytes primitifs. En culture, les cellules de sac vitellin provenant des souris ou des cellules ES *Lmo2*<sup>-/-</sup> ne réussissent pas à former des érythrocytes, comparativement aux cellules de type sauvage [31]. Toutefois, des macrophages primitifs semblent être détectés de par leur phénotype microscopique, bien que des marqueurs moléculaires n'aient pas été employés. De la même façon, les cellules de sac vitellin de souris *Scl*<sup>-/-</sup>, de même que les cellules ES *Scl*<sup>-/-</sup>, ne forment aucune colonie hématopoïétique en culture [51-54]. LMO2 et SCL sont donc nécessaires à l'hématopoïèse primitive.

Des cellules ES inactivées pour *Lmo2* ou *Scl* ont été générées afin d'obtenir des souris chimériques, et ces cellules ne contribuent à aucune lignée hématopoïétique définitive [53-55], alors qu'elles contribuent au développement de tous les autres tissus analysés. Ces résultats indiquent que le défaut hématopoïétique observé est cellulaire-autonome. SCL et LMO2 sont donc essentiels pour l'établissement de l'hématopoïèse primitive autant que définitive. Dans l'hématopoïèse définitive, ces gènes pourraient être nécessaires soit 1) à la génération des cellules souches hématopoïétiques; 2) à la colonisation de celles-ci dans le foie fœtal; ou 3) à leur différenciation en cellules matures. L'étude décrite ci-après pour SCL permet d'écarter les deux dernières hypothèses.

Afin de déterminer si l'expression de SCL est requise à un stade plus tardif du développement hématopoïétique, Schlaeger et collègues [56] ont procédé à la délétion conditionnelle du gène *Scl*. Dans ce modèle, le gène *Scl* a été flanqué de séquences *loxP* (*Scl<sup>fl/fl</sup>*) reconnues par la Cre recombinase, exprimée sous le contrôle du promoteur *Tie2* (*Tie2Cre*), dont l'expression est détectée immédiatement après celle de SCL au cours du développement. La délétion de *Scl* est efficace dans le sac vitellin et la région AGM au jour E10.5, après la spécification des cellules souches hématopoïétiques, mais avant que celles-ci ne colonisent le foie fœtal. Toutefois, les cellules souches hématopoïétiques sont présentes dans le foie fœtal des souris *Scl<sup>fl/fl</sup> Tie2Cre*. Ces résultats montrent que bien que SCL soit nécessaire à la spécification des cellules souches hématopoïétiques, il n'est pas requis pour leur colonisation dans le foie fœtal ni dans la moelle osseuse.

La formation des cellules sanguines chez l'embryon coïncide temporellement et spatialement avec le développement vasculaire, et plusieurs gènes, dont *Scl* et *Lmo2*, sont exprimés à la fois dans ces deux tissus. Différentes évidences pointent en faveur d'un progéniteur commun aux tissus hématopoïétiques et endothéliaux, l'hémangioblaste [57]. Ainsi, la fonction de SCL et de LMO2 dans la vasculogenèse et l'angiogenèse a été investiguée. La vasculogenèse consiste en l'émergence de vaisseaux sanguins primitifs *de novo* à partir du mésoderme ou des précurseurs endothéliaux, alors que l'angiogenèse consiste en le remodelage des cellules

endothéliales existantes pour former des vaisseaux sanguins matures. L'inactivation de *Lmo2* ou de *Scl* ne perturbe pas la vasculogénèse, mais compromet sévèrement l'angiogénèse durant le développement [34, 41, 58]. LMO2 semble aussi être requis pour l'angiogénèse associée aux tumeurs. Effectivement, LMO2 est exprimé dans l'endothélium des vaisseaux sanguins entourant différents types de tumeurs murines et humaines [59]. Des cellules ES *Lmo2*<sup>-/-</sup> ne contribuent pas à la vascularisation des tumeurs chez la souris, signifiant que LMO2 puisse être nécessaire à ce processus. Dans le même ordre d'idées, la surexpression de *Scl* promeut la vascularisation autant *in vivo* qu'*in vitro* [60]. Bien que *Scl* et *Lmo2* soient requis pour la spécification des cellules hématopoïétiques, ces gènes sont facultatifs pour la génération des vaisseaux sanguins primitifs, alors qu'ils sont nécessaires à leur maturation.

Puisque SCL et LMO2 sont nécessaires à la génération des cellules souches hématopoïétiques, ils sont possiblement importants pour l'homéostasie des cellules souches adultes. Le rôle de SCL dans les cellules souches hématopoïétiques adultes sera donc résumé ici, alors que celui de LMO2 n'a pas vraiment été investigué jusqu'à maintenant.

Au point de vue phénotypique, des souris transgéniques pour *Scl* sous le promoteur ubiquitaire *Sil* présentent un plus grand nombre de cellules souches hématopoïétiques [61], alors que la perte d'un allèle de *Scl* fait légèrement diminuer ce nombre [62]. Au niveau fonctionnel, différentes études de transplantation ont été réalisées avec des cellules souches dont l'expression de *Scl* est modifiée. Le premier groupe à avoir fait de telles études a conclu, d'après leur stratégie expérimentale et leurs méthodes d'analyse, que SCL n'affectait pas l'homéostasie des cellules souches hématopoïétiques [63]. Ces résultats ont depuis été contestés par plusieurs équipes, dont la nôtre. La surexpression de *Scl* dans les cellules souches augmente leur capacité de repopulation d'un hôte immuno-déficient [64]. Inversement, les cellules souches dont les niveaux de SCL sont réduits ou nuls présentent des défauts de reconstitution dans un hôte irradié [62, 65, 66]. Le nichage de ces cellules n'est pas affecté par la perte de *Scl* [65], et notre équipe a montré que la diminution des niveaux de *Scl* entraînait une perte de quiescence des cellules souches

hématopoïétiques [62], ce qui pourrait expliquer leur potentiel de reconstitution diminué. Toutefois, un certain niveau de reconstitution est tout de même atteint, suggérant que SCL n'est pas nécessaire à la fonction des cellules souches adultes. Cet effet subtil de SCL peut être expliqué par une redondance avec LYL1, un bHLH de la même sous-famille que SCL. En effet, alors que la perte de *Scl* ou de *Lyl1* affecte plutôt subtilement l'homéostasie des cellules souches hématopoïétiques, la double délétion de *Scl* et de *Lyl1* abolit presque totalement leur capacité de repopulation, puisque ces cellules meurent par apoptose [67]. Ainsi, l'expression de *Scl* ou de *Lyl1* est requise pour la survie des cellules souches hématopoïétiques.

Le rôle de LMO2 et de SCL a aussi été investigué dans les cellules hématopoïétiques plus différenciées. Particulièrement, les niveaux de SCL et de LMO2 influencent la différenciation érythrocytaire. Les souris ou les cellules ES nulles pour *Scl* ou pour *Lmo2* sont incapables de générer des érythrocytes primitifs [31, 51, 52], même si cela peut être la conséquence d'un défaut plus précoce. Ensuite, les souris *Scl<sup>fl/fl</sup> Tie2Cre* présentent des défauts de différenciation des érythrocytes primitifs [56]. De la même façon, les souris *Scl<sup>fl/fl</sup> MxCre* (délétion de *Scl* dans les cellules hématopoïétiques adultes) présentent des défauts de différenciation érythrocytaire et mégakaryocytaire [68-70]. Toutefois, des érythrocytes matures sont produits, et les souris survivent à long terme, bien qu'elles soient légèrement anémiques. Lorsque les cellules souches hématopoïétiques inactivées pour *Scl* sont transplantées dans une souris hôte, la reconstitution du compartiment érythrocytaire se fait efficacement, bien que les souris hôtes soient aussi anémiques [68]. Toutefois, si ces mêmes cellules souches sont transplantées en compétition avec des cellules de type sauvage, ces dernières prennent le dessus dans la reconstitution érythrocytaire, et les cellules inactivées pour *Scl* ne sont pas détectées dans la lignée érythroïde [63, 68]. Ces résultats montrent que SCL est très important, mais pas nécessaire, pour l'érythropoïèse chez l'adulte. La délétion de *Scl* affecte aussi la différenciation en mastocytes [71], alors que son effet sur la différenciation myéloïde est nul ou subtil [63, 66, 69, 72]. Au niveau cellulaire, dans des lignées de progéniteurs myélo-érythroïdes, la surexpression de *Scl* cause la différenciation spontanée (sans induction) en érythrocytes [39, 73], et bloque la différenciation vers la voie myéloïde.

L'effet de LMO2 sur la différenciation érythrocytaire est moins clair. À ce jour, la délétion conditionnelle de *Lmo2* dans la lignée érythroïde n'a pas été générée. D'autre part, selon différentes études, la surexpression de *Lmo2* peut favoriser [74] ou inhiber [75, 76] la différenciation érythrocytaire. Cette divergence peut provenir des différents modèles cellulaires employés pour la différenciation érythroïde, ou encore peuvent s'expliquer par un effet de dosage de LMO2 qui pourrait varier d'un modèle à l'autre. De façon plus générale, LMO2 régule positivement l'expression de nombreux gènes associés à la différenciation érythrocytaire [77-80], suggérant un rôle positif pour LMO2 durant ce processus.

Contrairement à sa fonction dans la lignée érythroïde, LMO2 n'est pas impliqué dans la différenciation lymphoïde normale. En effet, la délétion conditionnelle de *Lmo2* a été effectuée dans les cellules lymphoïdes, en exprimant la Cre sous le contrôle de promoteurs lymphoïdes (*Rag1*, *Cd19* et *Lck*) dans des souris *Lmo2<sup>fl/fl</sup>* [81]. *Lmo2* est ainsi absent dans les progéniteurs lymphoïdes de la moelle osseuse et dans les lymphocytes plus différenciés, et son inactivation n'affecte pas la différenciation des lymphocytes B et T. Puisque *Lmo2* est exprimé dans les stades précoces de différenciation thymocytaire, on ne peut pas exclure que l'absence de phénotype soit due à une redondance avec d'autres protéines de la même famille. Par exemple, *Lmo4* est exprimé dans les différentes populations thymocytaires [82], et son inactivation réduit le volume du thymus [83].

La délétion conditionnelle de *Scl* dans toutes les cellules hématopoïétiques adultes permet de déterminer si ce gène est important dans la hiérarchie hématopoïétique [63, 65]. Ainsi, les souris *SCL<sup>fl/fl</sup> MxCre* présentent une légère prédominance de lymphocytes T dans le sang périphérique, comparativement aux lymphocytes B, suggérant un rôle non essentiel pour SCL dans la différenciation lymphoïde normale. Lorsque surexprimé, *Scl* perturbe la différenciation des cellules B [61, 64, 84], sans affecter significativement la différenciation thymocytaire. L'expression ectopique de *Lmo2*, avec ou sans *Scl*, perturbe la différenciation lymphocytaire T, menant ultimement à des leucémies des cellules T, tel que décrit dans les sections suivantes.

## 1.2 Cancer et leucémie

De façon simpliste, le cancer se définit par la prolifération anarchique de cellules au sein d'un tissu donné. Une cellule cancéreuse est une cellule à l'origine normale, qui a subi des changements au cours des divisions, lui conférant un avantage prolifératif tout en lui permettant d'échapper aux différents contrôles mis en place par l'organisme pour maintenir constant le nombre de cellules d'un tissu donné.

### 1.2.1 Propriétés d'une cellule cancéreuse : le cancer en tant que processus multi-étapes

La cancérogenèse est un processus multi-étapes, signifiant que la cellule doit acquérir diverses altérations génétiques pour être totalement transformée. Par exemple, pour transformer des cellules humaines en culture, plusieurs oncogènes doivent être co-surexprimés [85]. De la même façon, chez l'humain, plusieurs anomalies génétiques différentes sont détectées dans les cellules cancéreuses. L'accumulation d'altérations génétiques séquentielles pourrait expliquer pourquoi la majorité des cancers survient tardivement dans la vie de l'individu [86]. Le caractère multi-étapes du cancer est aussi révélé par les études d'incidence de la leucémie chez les jumeaux identiques. En effet, les jumeaux identiques partagent le même placenta. La circulation placentaire permet à un même bassin de cellules souches, après leur expansion fœtale, de coloniser les deux foies fœtaux, faisant en sorte que les cellules souches hématopoïétiques puissent être génétiquement identiques dans les foies fœtaux des jumeaux. Ainsi, deux jumeaux peuvent avoir la même altération génétique dans un oncogène ou un suppresseur de tumeurs. Toutefois, la leucémie peut se développer après des périodes de latence très différentes, ou encore chez un seul jumeau [87], révélant l'importance d'événements génétiques additionnels qui surviennent différemment pour chacun des jumeaux.

Un exemple éloquent du caractère multi-étapes du cancer est l'apparition de leucémies chez les patients atteints d'immunodéficiences combinées sévères liées au chromosome X (SCID-X1), traités par thérapie génique. La **Section 1.2.3.2** décrit avec plus de détails cette maladie et les essais de thérapie génique effectués.

Brièvement, certains patients traités pour cette maladie ont développé une leucémie due à la thérapie génique, qui a modifié le locus du gène *LMO2*. Bien que les patients atteints de leucémies présentaient tous cette altération, la leucémie s'est développée chez ses individus après des périodes de latence différentes, soulignant à nouveau la nécessité d'autres altérations génétiques pour la leucémogénèse.

De façon intéressante, certains oncogènes présentent un tropisme pour des types de cancers très particuliers. Par exemple, chez la souris, la surexpression de *Lmo2* dans tous les tissus ne génère que des leucémies lymphoblastiques aiguës (majoritairement des leucémies des cellules T, mais quelques leucémies B ont aussi été observées) [88, 89]. De la même façon, dans les leucémies myéloïdes chroniques (CML), la translocation *BCR-ABL1* est retrouvée autant dans les cellules lymphoïdes que myéloïdes, suggérant une cible multipotente, bien que l'on observe l'expansion sélective des cellules granulocytaires [90]. L'environnement cellulaire doit permettre l'expression de la protéine, de même que son association à des partenaires précis, pour que l'expression de cette protéine soit avantageuse pour la cellule plutôt que délétère. De plus, les translocations chromosomiques impliquant *LMO2* et *BCR-ABL1* sont détectées chez une proportion significative de patients sains [91, 92], suggérant la nécessité d'altérations génétiques additionnelles pour l'apparition de cancer.

Au niveau cellulaire, on peut penser que la cellule ne peut être permissive à une telle accumulation de mutations que lorsque son intégrité génomique est compromise. Pour être transformée, la cellule cancéreuse doit à la fois être hautement proliférative (doit échapper aux contrôles normaux du cycle cellulaire), avoir une capacité de survie accrue (inhibition des mécanismes d'apoptose) et posséder un potentiel répliatif illimité, tel que proposé par Hanahan et Weinberg [93, 94].

### **1.2.1.1 Intégrité génomique**

La cellule a développé des mécanismes de contrôle à différents moments durant le cycle cellulaire pour s'assurer que les informations contenues dans son ADN demeurent intactes. Plusieurs protéines sont impliquées dans la détection et la réparation d'erreurs ou de bris dans l'ADN. Des défauts dans ces systèmes de



contrôle pourraient favoriser la mutabilité, et donc le développement du cancer. Le suppresseur de tumeurs *P53* est un des membres les mieux connus de ce système de contrôle, et est fréquemment muté dans différents types de cancers [95]. En réponse aux dommages à l'ADN, *P53* peut induire un arrêt du cycle cellulaire pour permettre la réparation de l'ADN, ou, si le dommage est excessif, peut enclencher des mécanismes d'apoptose. Plusieurs autres gènes impliqués dans la détection et la réparation de dommages à l'ADN, ou dans la surveillance de la ségrégation appropriée des chromosomes durant la mitose, sont inactivés dans certains cancers.

### **1.2.1.2 Prolifération cellulaire**

Les cellules normales ne peuvent proliférer que lorsqu'elles reçoivent des signaux extracellulaires de croissance, leur permettant de passer d'un mode quiescent à un mode prolifératif. Ces signaux peuvent provenir d'un contact direct entre la cellule et a) la matrice extracellulaire ou b) une autre cellule adjacente, ou encore de facteurs de croissance pouvant être diffusés. Ces signaux sont reçus par des récepteurs transmembranaires activant des cascades de signalisation qui promeuvent la prolifération. Plusieurs oncogènes (*HER2/neu* [96], *RAS* [97]) agissent en activant ces voies de signalisation en l'absence de signal extracellulaire (indépendance envers les signaux de croissance), menant à une prolifération cellulaire incontrôlée.

À l'inverse, plusieurs signaux antiprolifératifs sont dirigés vers les cellules à un moment ou un autre afin de les empêcher de proliférer et les maintenir dans un état quiescent, pour maintenir l'homéostasie dans un tissu donné. Les voies de signalisation activées de cette façon (ex.  $TGF\beta$  [98]) peuvent forcer les cellules à sortir du cycle cellulaire de façon permanente ou réversible, dépendamment des signaux reçus. La vaste majorité des signaux antiprolifératifs sont intégrés à l'intérieur de la cellule par le biais des protéines de la famille RB (« *retinoblastoma* »). RB bloque l'activité des facteurs de transcription de la famille E2F, responsables de l'expression de nombreux gènes essentiels au cycle cellulaire [99]. Dans les cellules cancéreuses, l'activité de RB peut être bloquée à différents

niveaux, soit dans la signalisation en amont de RB (*TGF $\beta$*  [98]), dans des régulateurs de RB (*CDKN2A/B* [100]) ou via RB lui-même [101].

### **1.2.1.3 Survie cellulaire et apoptose**

En plus de pouvoir proliférer, les cellules cancéreuses doivent survivre, en échappant à la mort cellulaire programmée (apoptose). Lorsque déclenchée, la mort cellulaire par apoptose se fait de façon très ordonnée et relativement rapide : les membranes cellulaires sont détruites, suivies de la structure du cytoplasme et du noyau; le cytosol est ensuite expulsé, les chromosomes sont dégradés et le noyau est fragmenté.

Le mécanisme d'apoptose est gouverné par deux voies principales d'activation, les voies extrinsèque et intrinsèque (résumées dans [102]). La voie extrinsèque implique des récepteurs de surface (TNF, FAS) : suite à la liaison de leur ligand, ces récepteurs recrutent et activent certaines caspases, des protéases dégradant des substrats cellulaires du noyau et du cytoplasme. La voie intrinsèque met particulièrement en jeu la mitochondrie, dont l'ouverture des pores est régulée par la famille de protéines BCL-2 pro- et anti-apoptotiques. L'ouverture des pores provoque la relâche du cytochrome C par la mitochondrie et l'activation de caspases. La voie intrinsèque est activée en réponse à différentes anomalies cellulaires (dommages à l'ADN, signalisation dérégulée par l'action d'oncogènes, stress cytotoxique). Par exemple, suite à des dommages à l'ADN, la protéine P53 augmente l'expression du gène pro-apoptotique *Bax* [103]. Pour échapper à la mort cellulaire programmée, les cellules cancéreuses peuvent surexprimer des gènes anti-apoptotiques (*BCL2* [104]) ou acquérir des mutations inactivant des gènes pro-apoptotiques (*P53* [95]).

### **1.2.1.3 Potentiel répliatif illimité**

En plus de la dépendance aux facteurs de croissance et de la mort cellulaire programmée, la cellule normale possède un autre mécanisme de contrôle l'empêchant de proliférer indéfiniment : la sénescence. En effet, la cellule normale ne peut se diviser qu'un nombre restreint de fois. Les télomères (séquences répétées d'ADN aux extrémités des chromosomes) raccourcissent légèrement lors de chaque phase S du

cycle cellulaire, dû à l'incapacité des polymérase à compléter la réplication des extrémités 3' de l'ADN. Lorsque l'érosion des télomères devient trop sévère, les extrémités des chromosomes ne sont plus protégées, et la cellule entre en apoptose ou encore en sénescence, un état d'arrêt irréversible du cycle cellulaire. Certains oncogènes (*Ras* [105], *Raf* [106], *E2f* [107], *Scf/Lmo* (Tremblay M. et Hoang T., *manuscrit en préparation*)) induisent la sénescence. C'est pourquoi la cellule cancéreuse doit développer un potentiel réplicatif illimité. Par exemple, dans plus de 85% des cancers humains, le gène *TERT* est surexprimé [108], encodant la sous-unité catalytique de l'enzyme télomérase qui permet de conserver la longueur des télomères en ajoutant des répétitions d'ADN aux extrémités des chromosomes.

### **1.2.2 La cellule souche cancéreuse**

À l'origine, on croyait qu'une tumeur cancéreuse était un tissu homogène, formé de cellules toutes identiques. Toutefois, à l'analyse, plusieurs tumeurs solides sont hétérogènes, tant au niveau du degré de différenciation des cellules, de leur taux de prolifération ou de leur capacité d'invasion. Récemment, l'identification de cellules souches cancéreuses a amélioré notre compréhension de la progression du cancer. Les cellules souches cancéreuses ont initialement été identifiées dans les leucémies [109], mais sont aussi retrouvées dans les tumeurs solides, notamment dans le cancer du sein [110], le cancer colorectal [111] et le cancer du cerveau [112].

Les cellules souches cancéreuses constituent la sous-population de cellules cancéreuses capable de régénérer une tumeur après transplantation dans une souris hôte immuno-déficiente. C'est donc la cellule initiatrice de la tumeur, de la même façon qu'une cellule souche normale est au sommet de la hiérarchie du tissu qu'elle génère. La cellule souche cancéreuse partage plusieurs caractéristiques avec une cellule souche normale : certains marqueurs phénotypiques, profil transcriptionnel, capacité de repopulation d'un tissu donné (auto-renouvellement et différenciation), faible prolifération (quiescence) (résumé dans [113]). Cette dernière caractéristique peut expliquer la réapparition du cancer chez certains patients, après les traitements classiques de chimiothérapie qui visent les cellules prolifératives.

Selon le type de cancer, l'origine de la cellule cancéreuse varie. Dans certaines tumeurs, la cellule souche du tissu d'origine est elle-même la cible de transformation, alors que dans d'autres tissus, une cellule plus différenciée peut acquérir des altérations génétiques qui lui confèrent des propriétés de cellules souches. Par exemple, dans les leucémies T-ALL, les translocations chromosomiques ont lieu dans les cellules T, et non dans les cellules souches hématopoïétiques (voir **Section 1.2.3.2**). Ainsi, la surexpression de *Lmo2* dans les T-ALL confère aux thymocytes primitifs des propriétés d'auto-renouvellement qui en font des cellules souches leucémiques [114]. L'identification de la population de cellules souches cancéreuses au sein d'un cancer donné peut permettre de mieux comprendre les mécanismes de transformation et de développer des thérapies ciblées efficaces à long terme.

### **1.2.3 La leucémie de type T-ALL**

La leucémie T-ALL est un cancer très agressif chez l'humain, et représente un groupe hétérogène de leucémies différant tant dans le phénotype immunologique que dans les anomalies génétiques et les caractéristiques cliniques comme la réponse à la thérapie. La leucémie T-ALL est, comme son nom l'indique, une leucémie aigue des cellules T, se caractérisant chez l'humain par un grand nombre de blastes (cellules immatures) dans la moelle osseuse et la circulation, ainsi qu'un grossissement des organes lymphatiques, particulièrement le thymus. Cette leucémie représente environ 15% des leucémies lymphoblastiques aigues chez l'enfant, et 25% chez l'adulte [115].

Dans certains types de leucémies, en opposition aux T-ALL, une seule translocation chromosomique est retrouvée majoritairement chez les patients, soulignant l'importance d'un oncogène particulier dans le processus leucémogénique. Par exemple, la translocation menant au chromosome de Philadelphie (protéine de fusion BCR-ABL1), est détectée dans environ 95% des patients atteints de leucémie myéloïde chronique (CML) [116]. De la même façon, la leucémie aigue promyélocytaire (APL) est aussi une maladie mono-oncogénique, puisque dans plus de 95% des cas, on retrouve une translocation chromosomique générant la protéine de

fusion PML-RAR $\alpha$  [117]. Certains oncogènes semblent donc être nécessaires (mais pas suffisants) à l'apparition d'un type particulier de leucémie. En comparaison, une variété d'altérations chromosomiques est détectée chez les patients atteints de T-ALL. Les réarrangements chromosomiques détectés dans les leucémies T-ALL impliquent différents gènes régulant des processus cellulaires souvent dérégulés dans les cancers, comme la différenciation, le cycle cellulaire ou la survie. En fait, des leucémies T-ALL peuvent présenter plus d'une altération chromosomique, en accord avec le postulat de maladie multi-génique (cancer en tant que processus multi-étapes).

### 1.2.3.1 Les principaux oncogènes associés à la leucémie T-ALL

Le **Tableau I** dresse la liste des gènes les plus fréquemment réarrangés ou mutés dans les leucémies T-ALL, avec des pourcentages représentatifs de plusieurs études indépendantes [115, 118].

<b>Gènes</b>	<b>Fréquence des réarrangements/mutations</b>
<i>CDKN2A /CDKN2B</i>	80-96%
<i>SCL /LYL1 /TAL2</i>	60%
<i>LMO1 /2</i>	59%
<i>NOTCH1</i>	56%
<i>TLX1 (HOX11) /TLX3 (HOX11L2)</i>	27%
<i>ABL1</i>	8%
<i>HOXA10 /11</i>	5-7%
<i>CALM-AF10 (PICALM-MLLT10)</i>	5-10%
<i>N-RAS</i>	5%
<i>FLT3</i>	5%
Fusions <i>MLL</i>	4-8%

**Tableau I. Gènes les plus fréquemment réarrangés ou mutés dans les leucémies T-ALL chez l'humain**

Les gènes *CDKN2A/B* codent pour des protéines de la famille INK4, des régulateurs négatifs du cycle cellulaire. Ces suppresseurs de tumeurs sont inactivés dans la quasi-

totalité des T-ALL. Le rôle de leur mutation dans la leucémogénèse est évidemment de permettre aux cellules d'échapper aux contrôles normaux du cycle cellulaire.

Il est intéressant de noter que la plupart des gènes mutés ou réarrangés dans les leucémies T-ALL codent pour des facteurs de transcription, soulignant l'importance de cette classe de protéines dans la leucémogénèse. La dérégulation des programmes transcriptionnels est un thème récurrent dans les leucémies aiguës. Toutefois, contrairement à d'autres types de leucémies (AML, B-ALL), pour lesquels on détecte fréquemment des protéines de fusion générant des facteurs de transcription chimériques [119, 120], la leucémie T-ALL est principalement associée à la dérégulation de l'expression des facteurs de transcription dans leur état sauvage.

Tel que décrit dans le **Tableau I**, certains gènes, ou familles de gènes, sont mutés ou réarrangés dans plus de 50% des leucémies T-ALL. L'altération de ces gènes semble donc nécessaire au développement des leucémies T-ALL. De plus, les oncogènes impliqués dans les leucémies T-ALL sont souvent surexprimés en l'absence d'anomalies dans leur locus [121]. Ces cas n'étant pas comptabilisés dans les études précédemment citées [115, 118], on peut penser que l'expression de *SCL/LYL1*, de *LMO1/2* et de *NOTCH1* soit altérée dans une très grande proportion des T-ALL.

On peut supposer que les translocations impliquant des protéines semblables sont fonctionnellement équivalentes dans la leucémogénèse. Par exemple, certains patients présentent des translocations chromosomiques impliquant le gène *LMO1*, alors que d'autres patients ont plutôt des anomalies dans le locus *LMO2*. Bien que *LMO1* ne soit pas normalement exprimé dans le système hématopoïétique, sa grande homologie avec *LMO2* peut expliquer pourquoi les translocations chromosomiques impliquant l'un ou l'autre de ces gènes engendrent des leucémies identiques. De la même façon, SCL fait partie des facteurs de transcription de la classe B de la famille des bHLH, tout comme *LYL1*, *TAL2* et *BHLHB1*. *TAL2* et *BHLHB1* ne sont pas non plus exprimés normalement dans les cellules T, mais ils sont tous impliqués, à des fréquences différentes, dans des réarrangements chromosomiques à l'origine de leucémies T-ALL. D'ailleurs, le domaine bHLH de SCL partage plus de 80%

d'homologie avec ceux de *LYL1* et de *TAL2*, et tous s'hétérodimérisent avec les mêmes partenaires, reconnaissant des séquences d'ADN semblables. En accord avec ces observations, l'expression ectopique de *SCL* et de *LYL1* semble être mutuellement exclusive dans les leucémies, suggérant que ces oncogènes induisent des perturbations semblables dans la leucémogénèse, puisque l'altération d'un seul gène est requise pour la leucémogénèse. Ainsi, les rôles de *LMO1*, ainsi que de *LYL1*, *TAL2* et *BHLHB1* seront très peu discutés dans cette thèse, puisque leurs mécanismes d'action semblent similaires à ceux de *LMO2* et *SCL*, respectivement.

Mis à part *LMO2* et *SCL* (et leurs homologues respectifs), *NOTCH1* est le seul autre oncogène connu à être muté dans plus de 50% des leucémies T-ALL. D'ailleurs, notre équipe a récemment montré que dans un modèle de souris, l'expression simultanée de *Scf*, *Lmo1* et *Notch1* est suffisante pour transformer complètement la cellule T [122]. *NOTCH1* a aussi été découvert de par son implication dans des translocations chromosomiques impliquant le locus du *TCR* dans des T-ALL [123], même si ces translocations sont en fait très rares chez les patients. Dans les cellules leucémiques humaines, on retrouve plutôt des mutations activatrices de *NOTCH1*, dans des régions encodant différents domaines de la protéine (domaine d'hétérodimérisation, domaine PEST), ce qui engendre son activation constitutive et/ou augmente sa stabilité protéique [124-126]. Les mutations de *NOTCH1* sont retrouvées chez les patients de tous les sous-groupes génétiques de T-ALL, incluant les groupes *HOX11*<sup>+</sup>, *HOX11L2*<sup>+</sup>, *SCL*<sup>+</sup>, *LYL1*<sup>+</sup>, *MLL-ENL*<sup>+</sup> et *CALM-AF10*<sup>+</sup> [124]. Ces mutations sont aussi retrouvées dans des leucémies murines de divers génotypes, par exemple dans des tumeurs de souris transgéniques pour *Scf* [122, 127].

Contrairement à *SCL* et *LMO2*, *NOTCH1* est normalement exprimé dans les cellules T et est un régulateur majeur de la différenciation thymocytaire (revu dans [128]). *NOTCH1* est impliqué autant dans la spécification des cellules T que dans diverses étapes de leur différenciation (par exemple le choix B vs T, ou le choix *CD4*<sup>+</sup> vs *CD8*<sup>+</sup>), tel que révélé par différentes études d'inactivation conditionnelle [128]. De la même façon, lorsque *Notch1* est constitutivement activé, la différenciation thymocytaire peut être perturbée à différentes étapes, tout dépendant du modèle de

souris utilisé (expression rétrovirale, transgènes sous le contrôle de différents promoteurs). Dans ces modèles murins, des leucémies des cellules T se développent [129-132]. Il est donc concevable que chez l'humain, une mutation activatrice de *NOTCH1* puisse engendrer des défauts de différenciation thymocytaire favorisant l'apparition de leucémie. D'autre part, *Notch1* promeut l'auto-renouvellement des cellules souches au détriment de leur différenciation [133, 134]. Il est donc probable que des mutations activatrices de *NOTCH1* confèrent aux thymocytes un potentiel d'auto-renouvellement qui contribue à son rôle oncogénique.

### 1.2.3.2 LMO2 et SCL en tant qu'oncogènes

Les gènes *SCL* et *LMO2* ont été découverts chez l'humain par leur implication dans des translocations chromosomiques à l'origine de leucémies T-ALL [30, 135-138]. Ces translocations juxtaposent la région codante de *SCL* ou de *LMO2* au promoteur de la chaîne  $\alpha$ ,  $\beta$  ou  $\delta$  du TCR, engendrant l'expression ectopique de *SCL* ou de *LMO2* dans les cellules T. Ces translocations proviennent de l'action des recombinases RAG, activées durant les stades précoces de la différenciation lymphocytaire pour effectuer la recombinaison V(D)J des chaînes du TCR, pour générer le répertoire immunologique. Les gènes *LMO2* et *SCL* possèdent une séquence cryptique pouvant être reconnue par les RAG, ce qui cause une recombinaison aberrante entre le gène du TCR et celui de *SCL* ou *LMO2* [91, 139]. Environ 3% des patients atteints de T-ALL présentent la translocation *TCR-SCL*, contre 7% pour *TCR-LMO2*. Le gène *SCL* peut aussi être activé dans les T-ALL suite à une délétion d'environ 90-100kb en amont du gène *SCL*, ce qui juxtapose la région codante de *SCL* aux régions régulatrices du gène *SIL*, qui est actif dans les cellules en prolifération [140]. Cette délétion engendre donc l'expression ectopique de *SCL* dans la lignée T (9-30% des patients atteints de T-ALL présentent la délétion *SIL-SCL*). Dans environ 4% des patients atteints de T-ALL, *LMO2* est activé via une délétion dans un régulateur négatif en amont du gène *LMO2* [141].

Le potentiel oncogénique de *LMO2* a aussi été révélé de façon totalement inattendue dans des essais de thérapie génique. En effet, l'immunodéficiences combinée sévère



liée au chromosome X (SCID-X1) est causée par une mutation de la chaîne  $\gamma$  du récepteur des interleukines-2, -4, -7, -9, -15 et -21 (*IL2RG*), et entraîne la mort des bébés atteints [142]. Une thérapie génique a été menée chez 20 de ces patients, lors de laquelle les cellules souches et progéniteurs hématopoïétiques des patients ont été prélevés, puis infectés avec un rétrovirus exprimant la chaîne  $\gamma$ , et enfin réintroduits chez les patients. Cette thérapie s'est avérée efficace dans le sens où 90% des patients ont été efficacement guéris de l'immunodéficiences. Toutefois, quelques années après le traitement (moyenne de 2-3 ans), 4 patients sur 9 d'une première cohorte, et 1 patient sur 10 d'une deuxième cohorte, ont développé une leucémie des cellules T, dont 4/5 ont été causées par mutagenèse insertionnelle dans le locus du gène *LMO2* [143-145]. La mutagenèse insertionnelle est observée lorsque le rétrovirus s'insère dans le génome à proximité d'un oncogène ou d'un suppresseur de tumeurs, modifiant ainsi son expression, ce qui peut ultimement transformer la cellule. Les cellules leucémiques des patients présentaient d'autres mutations et anomalies chromosomiques, suggérant que bien que *LMO2* soit oncogénique, d'autres événements de mutations sont requis pour l'apparition de la leucémie, en accord avec la vision du cancer en tant que processus multi-étapes.

Tel que décrit précédemment, différentes altérations chromosomiques, engendrant l'expression ectopique de *SCL* ou de *LMO2* dans le thymus, sont détectées chez les patients T-ALL avec une fréquence totale d'environ 40%, suggérant que *SCL* et *LMO2* sont importants pour la leucémogénèse. De plus, *SCL* et *LMO2* sont aussi exprimés de façon aberrante chez plusieurs patients T-ALL en l'absence de réarrangements chromosomiques impliquant leurs loci respectifs. Effectivement, une première étude (37 patients) a révélé que *SCL* est exprimé dans les T-ALL de plus de 60% des patients, présentant ou non des anomalies du locus *SCL* [146]. Une étude indépendante (59 patients) [121] est arrivée à un pourcentage légèrement plus faible (49% des échantillons totaux, dont 75% ne présentaient pas d'anomalie chromosomique du locus *SCL*). Cette même étude a révélé que 60% des échantillons exprimaient *LMO1* et/ou *LMO2*, et que l'expression de *LMO1/2* était presque toujours associée à l'expression de *SCL* ou de *LYL1* [121]. Une autre étude indépendante a montré que souvent, l'expression de *SCL* et de *LMO2* se fait à partir

des 2 allèles du patient atteint (42% des cas pour *SCL* et 64% pour *LMO2*) [147], suggérant des mécanismes transcriptionnels pour l'activation de ces gènes, en plus des réarrangements chromosomiques. Toutes ces études montrent que *SCL* (ou *LYLI*) et *LMO1/2* sont surexprimés à des fréquences élevées chez les patients atteints de leucémies T-ALL, suggérant que leur expression est requise dans la leucémogénèse.

Alors que les leucémies *HOX11*<sup>+</sup> sont associées à un bon pronostic (80-90% de survie après cinq ans, en 2003), les leucémies *SCL*<sup>+</sup> ou *LYLI*<sup>+</sup> (qui pour la plupart expriment aussi *LMO1/2*) sont plutôt associées à un mauvais pronostic (30-40% de survie après cinq ans, en 2003) [87, 121]. Il est donc nécessaire de mieux comprendre le fonctionnement de ces oncogènes afin de développer des thérapies ciblées.

Différents modèles murins ont permis d'étudier l'effet oncogénique de *SCL* et de *LMO2*. De façon surprenante, les souris transgéniques pour *Scl* (*Scl*<sup>tg</sup>) sous le promoteur *Sca-1*, le promoteur lymphoïde *Cd2* ou le promoteur ubiquitaire *Sil*, ne développent pas de leucémie, même après une période de 16 mois [84, 148-150]. Toutefois, sous le promoteur lymphoïde *Lck*, les souris *Scl*<sup>tg</sup> développent des leucémies T-ALL après une période de latence relativement longue (moyenne de 10-12 mois) [151, 152]. En comparaison, des souris transgéniques exprimant *Lmo2* (*Lmo2*<sup>tg</sup>) sous le contrôle de différents promoteurs lymphoïdes (*Cd2*, *Thy1.1*) ou ubiquitaires (*metallothionein-1*) développent toutes des leucémies T-ALL, avec un temps de latence moyen de 9-10 mois [88, 153-155]. De façon intéressante, même lorsque *Lmo2* est surexprimé dans tous les tissus, les souris développent uniquement des leucémies T [88], récapitulant le phénotype des leucémies humaines. Cela suggère que *LMO2* peut être oncogénique uniquement dans les cellules T. La longue période de latence observée suggère que *LMO2* et *SCL*, seuls, ne sont pas suffisants pour induire la leucémie, et que d'autres événements de transformation sont requis.

Durant la période asymptomatique (stade pré-leucémique), les souris *Lmo2*<sup>tg</sup> présentent des défauts de différenciation thymocytaire [88, 155]. Les cellules CD4<sup>+</sup> et CD8<sup>-</sup> (double négatives, DN) ayant réarrangé la chaîne β du TCR s'accumulent dans le thymus de ces souris avec le temps, suggérant un possible défaut de différenciation

vers le stade double positif, ou DP. Des cellules DP sont toutefois observées dans le thymus, quoiqu'en nombre réduit, indiquant que le blocage de différenciation est partiel. D'ailleurs, les leucémies développées dans ces souris expriment souvent les marqueurs CD4 et CD8 [88, 154]. Les cellules DN des souris *Lmo2*<sup>tg</sup> sont hautement prolifératives[88], ce qui peut évidemment contribuer à leur accumulation.

Une étude récente a montré que l'expression ectopique de *Lmo2* dans les thymocytes conférait des propriétés de cellules souches aux thymocytes du stade DN3 [114]. Effectivement, des thymocytes transgéniques pour *Lmo2* sont capables de repeupler le thymus de souris irradiées, même après plusieurs transplantations en série, tout en retenant leur capacité de se différencier en thymocytes matures. L'expression de *Lmo2* dans les thymocytes entraîne une augmentation de l'expression d'environ 80 gènes, dont une majorité est normalement exprimée dans les cellules souches hématopoïétiques, plutôt que dans les thymocytes. En contraste, des gènes associés au développement des cellules T sont plutôt réprimés par l'expression de *Lmo2*. Ces différentes études montrent que LMO2, seul, peut modifier les programmes de différenciation et d'auto-renouvellement des thymocytes. Au contraire, les souris transgéniques pour *Scl* ne présentent pas de défaut de différenciation thymocytaire [148, 149, 151, 152], à l'exception des souris *Scl*<sup>Scal-Tg</sup> pour lesquelles la cellularité thymique est réduite [84], suggérant que le blocage de la différenciation thymocytaire est dépendant de l'expression de LMO2.

Alors que l'expression ectopique de *Scl* ou de *Lmo2* seul génère des leucémies plutôt tardivement, ces deux gènes collaborent ensemble dans la leucémogénèse. Effectivement, les cellules leucémiques de patients présentent fréquemment des altérations génétiques à la fois pour *SCL* et *LMO2*, suggérant une coopération entre ces deux gènes. D'autre part, des souris transgéniques à la fois pour *Scl* et *Lmo2*, sous le promoteur lymphoïde *Cd2*, développent des leucémies T-ALL avec une période de latence moyenne de 7 mois (accélération de 2-3 mois par rapport à *Lmo2*<sup>tg</sup>) [148], indiquant une coopération entre les deux oncogènes dans la leucémogénèse. De la même façon, lorsque des souris *Sil-Scl* (ne générant pas de leucémie) sont croisées avec des souris *Lck-Lmo1* (générant des leucémies après 5-16 mois de latence), des

leucémies T-ALL commencent à apparaître à 10 semaines, avec un temps de latence moyen de 4 mois [150]. Ainsi, sous les promoteurs *Cd2* et *Sil*, l'expression de *Scl* n'est pas suffisante pour l'apparition de la leucémie, mais *Scl* favorise tout de même la leucémogénèse en collaboration avec d'autres oncogènes comme *Lmo1* et *Lmo2*. Toutefois, cette période de latence est encore relativement longue, suggérant que d'autres événements de transformation sont requis.

Au stade pré-leucémique, l'accumulation de cellules DN dans le thymus est beaucoup plus importante dans les souris *Scl<sup>tg</sup>/Lmo2<sup>tg</sup>* que dans les souris *Lmo2<sup>tg</sup>* [148]. Il y a donc synergie entre LMO2 et SCL au stade pré-leucémique pour aggraver le défaut de différenciation des thymocytes. Ce défaut est observé dès la naissance, indiquant que la coopération entre les deux oncogènes survient très tôt dans le développement du thymus. Contrairement aux souris transgéniques pour *Lmo2*, dont les cellules leucémiques peuvent exprimer les marqueurs CD4 et CD8, les cellules leucémiques *Scl<sup>tg</sup>/Lmo2<sup>tg</sup>* n'expriment pas ces marqueurs (leucémies DN), alors que le réarrangement de la chaîne  $\beta$  du TCR est observé, confirmant qu'il s'agit bien de cellules T. Alors que les cellules DN s'accumulent dans les thymus des souris *Lmo2<sup>tg</sup>*, tout comme dans les souris bi-transgéniques, il semble que les cellules *Lmo2<sup>tg</sup>* puissent passer au stade DP avant l'apparition de la leucémie, possiblement de par la longue période de latence précédant l'apparition de tumeurs. En contraste, les cellules leucémiques bi-transgéniques sont uniquement DN.

*Scl* et *Lmo* coopèrent individuellement avec la présence d'autres gènes (*v-ABL* [156], *CKII $\alpha$*  [151], signalisation par le pre-TCR [122]) ou l'inactivation de suppresseurs de tumeurs (*p53* [157], *Cdkn2a* [89]) pour accélérer le processus leucémogénique. Notre équipe a d'ailleurs montré la collaboration entre *Scl*, *Lmo1* et *Notch1* [122]. Les souris bi-transgéniques pour *Scl* et *Lmo1* développent des leucémies après une longue période de latence d'environ 4 mois, et les cellules leucémiques présentent presque toutes des mutations dans le gène *Notch1*, soulignant l'importance de ce gène dans la leucémogénèse induite par *Scl* et *Lmo1*. D'ailleurs, les souris transgéniques à la fois pour *Scl*, *Lmo1* et *Notch1* développent des leucémies après seulement 30 jours, soit le temps suffisant pour l'expansion clonale d'une cellule souche leucémique en

transplantation. Cela signifie que l'expression de ces trois oncogènes est suffisante pour transformer les thymocytes, sans nécessiter d'événements de mutations supplémentaires. En contraste, la surexpression de certains oncogènes, comme *Pim-1* [150] ou *N-ras* [157], n'accélère pas l'apparition des leucémies induites par SCL.

## 1.3 Le gène *LMO2*

### 1.3.1 Régulation de l'expression de *LMO2*

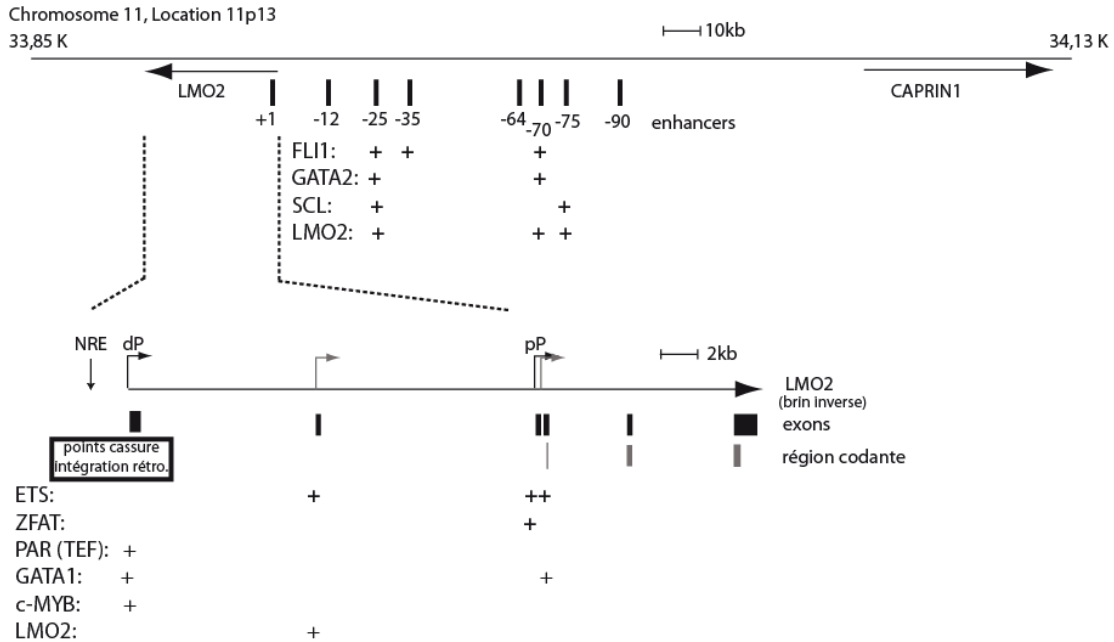
L'expression de *LMO2* est régulée au niveau transcriptionnel par l'association de différents facteurs de transcription à son promoteur. L'expression protéique de *LMO2* est aussi régulée au niveau post-transcriptionnel, par l'action de micro-ARNs de même que par l'ubiquitination.

#### 1.3.1.1 Régulation transcriptionnelle de *LMO2*

Chez l'humain, le gène *LMO2* se situe sur le chromosome 11. Tel qu'illustré dans la **Figure 1.3**, le gène *LMO2* contient six exons, dont les trois derniers codent pour la protéine de 158 acides aminés. La transcription se fait à partir de différents promoteurs, dont les deux les plus étudiés ont été nommés promoteur distal (en amont de l'exon 1) et proximal (en amont de l'exon 3) [158]. Plusieurs sites d'initiation de la transcription à partir du promoteur proximal sont répertoriés, mais tous les transcrits encodent la même protéine.

Alors que le promoteur proximal est plutôt utilisé de façon ubiquitaire dans différents tissus, l'utilisation du promoteur distal est restreinte aux tissus hématopoïétiques (moelle osseuse, foie fœtal) [158, 159]. Des souris transgéniques contenant un gène rapporteur sous le contrôle du promoteur distal de *Lmo2* ont été générées, et cette construction suffit pour récapituler l'expression dans les tissus hématopoïétiques [159]. Toutefois, bien qu'en essai de souris transgénique, le promoteur proximal ne puisse pas à lui seul diriger l'expression dans les cellules hématopoïétiques, ce promoteur semble être aussi largement utilisé dans ces cellules [160].

Des analyses de conservation de séquence, couplées à des études de ChIP-chip contre certaines histones acétylées, ont démontré la présence d'*enhancers* dans le locus *Lmo2*, dispersés sur une distance d'environ 100kb en amont du gène [161]. Ces



**Figure 1.3 Organisation du locus *LMO2* chez l'humain**

*Haut* : La position du gène *LMO2* sur le chromosome 11 est représentée par une flèche, de même que le gène le plus près, *CAPRINI*. La position des *enhancers*, relativement au site d'initiation de la transcription à partir du promoteur distal, est indiquée. La présence de différents facteurs de transcription sur les *enhancers* est indiquée par le signe +. *Bas* : La région agrandie, tel que précisé par les pointillés, représente le gène *LMO2* sur le brin inverse. Les différents sites d'initiation de la transcription sont indiqués par des flèches recourbées. La localisation de l'élément régulateur négatif (NRE) est indiquée. La région où ont lieu les points de cassure des réarrangements chromosomiques et les sites d'insertion rétrovirale est représentée par un encadré. Les exons et la région codante sont représentés par des boîtes noires et grises, respectivement. La présence de facteurs de transcription sur les différents promoteurs de *LMO2* est indiquée par le signe +. dP : promoteur distal; pP : promoteur proximal.

*enhancers* ont été testés en transgénèse chez la souris et contribuent positivement à l'expression de *Lmo2* dans divers tissus. Ceux-ci sont liés par différents facteurs de transcription, dont les membres du complexe SCL/LMO2, et certaines protéines de la famille ETS, particulièrement FLI1. Bien que la protéine SCL ait été détectée au

locus *Lmo2* sur certains *enhancers* [161], l'expression de *Lmo2* ne semble pas être fortement régulée par SCL, puisque dans le sac vitellin de souris *Scf*<sup>-/-</sup> [52], de même que dans le foie fœtal des souris conditionnellement inactivées pour *Scf* [56], le transcrit de *Lmo2* est détecté au même niveau que pour les souris de type sauvage. Une redondance avec d'autres bHLH de la même famille que SCL, comme par exemple LYL1, pourrait expliquer ces observations.

Alors que le locus entier de *LMO2* est très conservé chez les mammifères, seuls le promoteur proximal et la séquence codante sont conservés chez le poisson et la grenouille [160]. Le promoteur proximal humain contient trois sites ETS qui sont requis pour l'activité transcriptionnelle. Trois protéines de la famille ETS (FLI1, ETS1 et ELF1) ont été détectées par ChIP sur le promoteur proximal [160]. D'ailleurs, chez le poisson-zèbre, ETS1 a aussi été identifié comme un régulateur positif du promoteur de *lmo2* [162]. Chez l'humain, le promoteur proximal de *LMO2* est aussi régulé par la protéine à doigts de zinc ZFAT [163].

L'exon 1 de *LMO2* contient un site consensus lié par des protéines à domaine PAR, et il a été montré qu'une protéine de cette famille, TEF, était active en essai de transactivation du promoteur *LMO2* [159]. Le promoteur distal est régulé par le facteur de transcription GATA-1 de même que par c-MYB [164].

90% des différents points de cassure des translocations chromosomiques impliquant le locus *LMO2* se situent dans une région de 6,5 kb comprenant le promoteur distal, qui est donc souvent absent dans les réarrangements chromosomiques [158]. C'est aussi dans cette région qu'ont eu lieu les insertions rétrovirales à l'origine de leucémies dans le traitement de thérapie génique. Le gène *LMO2* contient aussi un élément répresseur (NRE, *negative regulatory element*) dans la région -2187 à -1915 impliquant des histones déacétylases [165], mais la protéine responsable de cette répression n'est toujours pas identifiée. En essai de transactivation, cette région est nécessaire pour empêcher l'expression de *LMO2* dans des lignées de cellules T. Il est intéressant de noter que cette région est absente dans plusieurs réarrangements chromosomiques impliquant le locus *LMO2*.

Un transcrit alternatif à partir d'un autre promoteur présumé a aussi été identifié, contenant uniquement deux exons, et codant pour une protéine de 151 acides aminés, dont les sept premiers diffèrent de la protéine de 158 acides aminés communément étudiée [166]. Cette protéine alternative comprend les deux domaines LIM intacts, et retient la capacité d'interagir avec ses partenaires SCL et GATA-1 [167]. Toutefois, certaines propriétés de la protéine sont modifiées, telles que son activité transcriptionnelle et son interaction avec la protéine LDB1. Ce troisième promoteur est régulé positivement par le facteur de transcription GATA-1, et négativement par le facteur de transcription PU.1, de la famille ETS. Un autre promoteur, nommé promoteur intermédiaire, a été identifié en amont de l'exon 2 [168]. Dans les lignées cellulaires et les cellules primaires de patients T-ALL, ce promoteur est lié par les facteurs de transcription de la famille ETS (FLI1, ERG), de même que par LMO2 lui-même.

### **1.3.1.2 Régulation des niveaux protéiques de LMO2 par l'action de micro-ARNs**

LMO2 est exprimé dans les cellules souches hématopoïétiques, et son expression diminue à mesure que les cellules se différencient. Cela suggère que l'expression de LMO2 doit être étroitement contrôlée dans les cellules hématopoïétiques afin de leur permettre de se différencier adéquatement. En plus du contrôle transcriptionnel, les niveaux de LMO2 sont aussi régulés au niveau protéique, particulièrement par l'action de micro-ARNs, qui ciblent l'ARNm de LMO2 pour la dégradation.

Différentes équipes, incluant la nôtre, ont montré que miR-223 régulait directement LMO2 dans les cellules hématopoïétiques. Nous avons observé que l'expression de LMO2 était inversement corrélée à celle de miR-223 dans différentes lignées hématopoïétiques (**Annexe 1**) (Lisi V, Sincennes MC et al, *manuscrit en préparation*), suggérant que la présence de miR-223 empêche l'expression de LMO2. miR-223 et LMO2 sont tous deux détectés dans les lignées de progéniteurs érythroïdes TF-1 et K562. Ces cellules peuvent se différencier en érythrocytes



matures suite à une stimulation par des cytokines, ce qui permet d'étudier l'expression de LMO2 et de miR-223 dans un contexte dynamique.

Lorsque les K562 se différencient en érythrocytes, l'expression de miR-223 s'éteint, alors que celle de LMO2 augmente [169]. Le même phénotype a été observé dans des cellules primaires de sang de cordon se différenciant en érythrocytes [170]. Inversement, lorsque les K562 se différencient plutôt en mégakaryocytes, l'expression de miR-223 est plus forte, alors que LMO2 devient indétectable [169]. Au point de vue fonctionnel, la surexpression de miR-223, autant dans les cellules TF-1 (**Annexe 1**) (Lisi V, Sincennes MC et al, *manuscrit en préparation*), K652 [169] ou de sang de cordon [170], compromet de façon significative la différenciation en érythrocytes matures, et fait diminuer l'expression de LMO2. Dans les cellules de sang de cordon, l'utilisation de shRNAs contre LMO2 provoque le même phénotype que la surexpression de miR-223 [170]. La surexpression de miR-223 dans différentes lignées de cellules B fait aussi diminuer l'expression de LMO2 [171, 172]. Des méthodes bioinformatiques nous ont permis d'identifier LMO2 comme étant fort probablement une cible directe de miR-223. En accord avec ces prédictions, miR-223 semble cibler directement l'ARNm de LMO2, tel qu'observé dans des essais de transfection transitoire avec des gènes rapporteurs [169-172].

En plus de miR-223, nous avons identifié miR-363 comme régulateur de LMO2 (**Annexe 1**) (Lisi V, Sincennes MC et al, *manuscrit en préparation*). Comme pour miR-223, l'expression de miR-363 est inversement corrélée avec celle de LMO2 dans différentes lignées hématopoïétiques. La surexpression de miR-363 dans les cellules TF-1 compromet la différenciation érythrocytaire en diminuant les niveaux de LMO2. LMO2 semble aussi être une cible directe de miR-363 tel que prédit de façon bioinformatique.

Chez la drosophile, miR-9a régule directement dLMO, l'orthologue de LMO2 [173]. Tel que discuté dans la **Section 1.3.3**, dLMO est bien connu pour son rôle dans le développement de l'aile de la drosophile. Le mutant gain-de-fonction de dLMO *ms1096* est en fait un mutant dont l'ARNm est plus stable, dû particulièrement à une

perte du site de liaison pour miR-9a. Chez la mouche, dLMO interagit génétiquement avec miR-9a [173]. De plus, miR-9a régule les niveaux protéiques de dLMO, autant *in vivo* qu'*in vitro*, et miR-9a cible l'ARNm de dLMO en essais de gène rapporteur luciférase. Le site de liaison de miR-9a est conservé chez LMO2 humain, suggérant que ce micro-ARN peut aussi réguler les niveaux protéiques de LMO2 dans un contexte hématopoïétique ou dans d'autres tissus. miR-9 est exprimé dans les cellules hématopoïétiques [174, 175] et est dérégulé dans différents cancers, mais son rôle le mieux caractérisé est dans la neurogenèse [176], où LMO2 est aussi exprimé.

### **1.3.1.3 Régulation des niveaux protéiques de LMO2 par l'ubiquitination**

En plus d'être régulés par l'action de micro-ARNs, les niveaux protéiques de LMO2 sont aussi contrôlés par le système ubiquitine/protéasome. Nous avons montré que LMO2 est une protéine très instable, ayant une demi-vie d'environ 30 minutes [177] (article présenté en **Annexe 2**). La protéine LMO2 est stabilisée lorsque le protéasome est inhibé par la molécule MG132, suggérant que LMO2 est dégradé par le protéasome. Un autre groupe a montré que LMO2 est ubiquitiné *in vivo* et *in vitro* par l'action de l'ubiquitine ligase RLIM [178].

La protéine LMO2 est stabilisée par l'interaction avec ses différents partenaires, comme SCL et LDB1 [177, 179]. Il est possible que les différentes interactions protéiques ne permettent plus à l'ubiquitine ligase de reconnaître son motif de liaison à LMO2, ou encore que le site d'ubiquitination de LMO2 ne soit plus exposé lorsque LMO2 est lié à ses partenaires. D'autres expériences sont requises pour évaluer ces hypothèses. Tout comme LMO2, LMO1 est aussi stabilisé par l'utilisation de MG132, de même que par son association avec SCL [177] (article présenté en **Annexe 2**). LMO4 est aussi stabilisé par son interaction avec LDB1 [179].

### **1.3.2 Les protéines à domaines LIM**

Le domaine LIM a été nommé d'après les premières protéines identifiées contenant ce domaine : LIN11 (nématode), ISL1 (rat) et MEC3 (nématode). Le domaine LIM

est retrouvé chez tous les eucaryotes dont le génome est caractérisé, incluant la levure et certaines plantes. Le domaine LIM est constitué d'environ 55 acides aminés, formant deux doigts de zinc riches en histidine et cystéine. Même si, à l'origine, le domaine LIM a été comparé aux doigts de zinc des facteurs de transcription de la famille GATA [180, 181], il semble que le domaine LIM ne possède pas la capacité de lier l'ADN directement *in vivo*. Le domaine LIM représente en fait une plateforme d'interactions protéine-protéine [182-184], et ne semble pas lier préférentiellement un motif protéique particulier. Les domaines LIM peuvent interagir avec une grande variété de partenaires, incluant des protéines de structure, des kinases, des facteurs de transcription et d'autres protéines à domaines LIM.

Chez l'humain, on compte une soixantaine de gènes encodant des protéines à domaines LIM [185]. La fréquence des domaines LIM chez l'humain est comparable à celle d'autres domaines d'interaction protéine-protéine, comme les domaines *SRC-homology-2* (SH2) ou SH3 [185]. Chez l'humain, les protéines à domaines LIM contiennent généralement entre deux et cinq domaines LIM, qui peuvent se retrouver au centre de la protéine autant qu'aux extrémités C- et N-terminales. Les domaines LIM peuvent être juxtaposés à d'autres domaines : homéodomaine, SH3, PDZ, kinase, etc. Les protéines à domaines LIM forment donc une famille très diversifiée en termes de structure, mais aussi de fonction : en effet, elles ne forment pas une famille fonctionnellement homogène, comme par exemple la famille bHLH qui regroupe exclusivement des facteurs de transcription. Les différentes protéines à domaine LIM jouent des rôles cellulaires diversifiés, principalement dans la régulation de la transcription ainsi que dans la constitution du cytosquelette. On peut classer ces protéines en différentes sous-familles, selon la similarité de leurs domaines LIM : les protéines LIM à homéodomaine, les protéines LMO, les LIM kinases, les protéines CRP, FHL, etc. Plus grossièrement, on peut aussi regrouper les différentes sous-familles en deux groupes, soit les protéines LIM nucléaires (LMO et LIM à homéodomaine) et cytoplasmiques (toutes les autres sous-familles).

Les protéines LIM cytoplasmiques (ex. ZYXIN, PAXILLIN, CRP1-3, FHL1-5, LIMK1-2) forment un groupe plutôt hétérogène. Parmi cette famille, la majorité des

protéines est impliquée dans la structure et/ou la régulation du cytosquelette [186-193]. Particulièrement, plusieurs protéines (ZYGIN [194], PAXILLIN [195], TESTIN [196], HIC-5 [197], FBLIM1 [198], FHL1-3 [199-201], PINCH1-2 [202-204], LMO7 [205], PDLIM2 [206], ZASP [207]) font partie des plaques d'adhésion focale, un complexe macromoléculaire où les intégrines s'accumulent et où les filaments d'actine sont ancrés. Les plaques d'adhésion focale effectuent le contact entre la matrice extracellulaire et le cytosquelette, et régulent entre autres l'adhésion et la motilité cellulaire [208, 209]. Les LIM kinases (LMK1, LMK2) sont des protéines contenant deux domaines LIM ainsi qu'un domaine kinase. LMK1 joue un rôle central dans l'organisation du cytosquelette d'actine, en phosphorylant la protéine cofiline, ce qui modifie la dynamique de polymérisation des filaments d'actine [210, 211]. LMK1 est régulée par la voie Rho impliquée dans la réorganisation du cytosquelette d'actine [211, 212]. Les trois membres de la sous-famille *cystein-rich protein* (CRP) sont associées aux filaments d'actine, et interagissent avec la protéine à domaines LIM ZYGIN, localisée aux plaques d'adhésion, de même qu'à l' $\alpha$ -actinin [213]. D'ailleurs, les souris *Crp3<sup>-/-</sup>* présentent des défauts dans l'architecture des muscles cardiaques et squelettiques [214]. Certaines protéines LIM cytoplasmiques ont aussi été associées à d'autres fonctions cellulaires, telles que l'apoptose [215, 216], l'inactivation génique par les micro-ARNs [217] ou encore la régulation de la mitose [218, 219].

En plus de leurs fonctions dans le cytoplasme, la majorité des protéines LIM cytoplasmiques possèdent aussi la capacité de faire la navette entre le noyau et le cytoplasme [184, 220-234]. Dans le noyau, ces protéines s'associent à divers facteurs de transcription [231, 235-239], particulièrement à des récepteurs nucléaires [223, 229, 240-245], et modulent leur activité transcriptionnelle. D'ailleurs, les membres de la sous-famille FHL (*four-and-a-half LIM*) sont davantage caractérisés en tant que cofacteurs transcriptionnels [241, 242, 246-257] que pour leur rôle cytoplasmique [199-201, 258-260]. Certaines protéines LIM (TRIP6 et LPP) ont été détectées au niveau des télomères, et pourraient jouer un rôle dans la protection des télomères en s'associant avec la protéine POT1 [261, 262].

Les protéines LIM à homéodomaine (ISL1-2, LHX1-9, LMX1a-b) contiennent à la fois deux domaines LIM et un homéodomaine. Cette famille de protéines représente des facteurs de transcription, qui contactent l'ADN via leur homéodomaine, et qui sont principalement impliqués dans le développement et la spécification du destin cellulaire. Par exemple, chez la souris, LHX1 est impliqué dans le développement de certaines structures de la tête [263], LHX2 dans le développement de l'œil et du cortex cérébral [264] et LHX3 et LHX4 sont importants pour le développement de l'hypophyse [265, 266]. Les protéines LIM à homéodomaine ont des patrons d'expression bien précis, restreints spatialement et temporellement, en accord avec leurs fonctions dans la spécification de types cellulaires.

Les protéines de la famille LMO sont nommées ainsi pour « *LIM-Only* ». Ce sont de petites protéines contenant presque exclusivement deux domaines LIM. Tel que discuté dans la section suivante, leur principale fonction est de participer à l'assemblage de larges complexes transcriptionnels.

Les protéines de la famille LMO et LIM à homéodomaine s'associent à un cofacteur appelé *LIM-domain binding1* (LDB1), et cette interaction est déterminante pour leurs fonctions transcriptionnelles. LDB1 ne peut s'associer aux protéines LIM cytoplasmiques [267]. D'ailleurs, les domaines LIM des protéines dites nucléaires ont une grande similarité de séquence, comparativement aux autres domaines LIM. La liaison de LDB1 avec les protéines LIM nucléaires semble moduler leur stabilité et leur pouvoir transcriptionnel, mais permet surtout d'établir une plateforme de recrutement pour former de larges complexes multiprotéiques, permettant entre autres l'association avec d'autres protéines LIM [80, 268], impliquant parfois l'homodimérisation de LDB1 [269]. La délétion de *Ldb1* chez la souris engendre de multiples défauts durant la gastrulation, incluant entre autres un développement anormal de la tête et des îlots sanguins du sac vitellin [270].

### **1.3.3 Les protéines de la famille LMO**

Chez l'humain et la souris, la famille LMO comprend quatre membres, LMO1-4. Tous sont associés à des complexes transcriptionnels régulant le développement et la

différenciation dans divers tissus. De plus, ce sont aussi tous des oncogènes puissants lorsqu'exprimés de façon aberrante. La séquence protéique des membres de la famille LMO est très conservée entre les espèces, particulièrement entre l'humain et la souris (deux acides aminés changés pour LMO1, un seul pour LMO2, et aucun pour LMO3-4), ce qui suppose un rôle fondamental pour ces protéines. Les protéines LMO se ressemblent aussi entre elles. Celles qui partagent le plus d'homologie sont LMO1 et LMO3 (98% d'homologie dans les domaines LIM), suivies de LMO2 (60% homologie avec LMO1), et de LMO4 (55% homologie avec LMO1 et 46% avec LMO2), qui est le plus divergent de la famille. Chez la drosophile, il n'y a qu'un seul membre de la famille LMO, dLMO, qui est conservé à 65% avec LMO2 et à 85% avec LMO1. dLMO est aussi retrouvé dans des complexes transcriptionnels gouvernant le développement et la différenciation, tel que décrit ci-après.

Le gène *dLMO* a été identifié de par son homologie avec *LMO2* [271]. Son expression est détectée dans plusieurs tissus autant durant le développement que chez la mouche adulte. Le phénotype des allèles de mouche *Beadex*, présentant des anomalies dans la marge de l'aile (semblable au phénotype *notch*), est attribuable à une mutation dans un régulateur négatif de *dLMO* (région 3'UTR), ce qui engendre la surexpression de *dLMO* [272-274]. Une mutation dans la région codante de *dLMO* cause aussi la perte de fonction de ce gène chez l'allèle *heldup-a* [274]. D'ailleurs, la mutation *heldup-a* supprime le phénotype *Beadex*. Chez la mouche, *dLMO* interagit génétiquement avec *Apterous*, codant pour une protéine LIM à homéodomaine (orthologue de *Lhx2* et *Lhx9*), de même qu'avec *Chip*, l'orthologue de *Ldb1* [273, 274]. Chez la mouche, dLMO compétitionne avec APTEROUS pour la liaison à CHIP [273], récapitulant les interactions protéine-protéine observées chez les mammifères. Les niveaux stœchiométriques de dLMO et de APTEROUS sont donc déterminants pour le développement de l'aile chez la mouche, dLMO agissant comme régulateur négatif de l'activité transcriptionnelle de APTEROUS [275, 276]. Comme chez les mammifères (discuté dans la section suivante), dLMO est aussi retrouvé dans des complexes transcriptionnels qui régulent positivement la transcription, comprenant PANNIER (orthologue de GATA), CHIP (orthologue de LDB1), ACHAETE et DAUGHTERLESS (deux facteurs bHLH comme SCL et E2A) [277].

*LMO1* a été découvert de par son implication dans des translocations chromosomiques menant à des leucémies T-ALL [135, 278], tout comme *LMO2* [30, 135]. Exactement comme pour *Lmo2*, des souris transgéniques pour *Lmo1*, sous le contrôle d'un promoteur spécifique des cellules T (*Lck*), développent des leucémies de type T-ALL après une longue période de latence [153, 279]. De plus, toujours à l'instar de *LMO2*, *LMO1* s'associe à SCL et collabore avec celui-ci pour accélérer la leucémogénèse dans des modèles de souris transgéniques [150, 280]. Toutefois, en contraste avec *LMO2*, *LMO1* n'est pas normalement exprimé dans les cellules hématopoïétiques. *LMO1* est majoritairement exprimé dans le cerveau et le système nerveux [46, 281]. D'ailleurs, *LMO1* a récemment été identifié comme oncogène dans les neuroblastomes [282], un cancer du système nerveux sympathique, fréquent chez les enfants. L'expression de *LMO1* dans les neuroblastomes a été corrélée avec le degré d'avancement de la maladie, de même qu'avec le taux de survie. L'utilisation de shRNAs contre *LMO1* inhibe la croissance des cellules de neuroblastomes, alors que la surexpression de *LMO1* a l'effet inverse. Au niveau moléculaire, la fonction de *LMO1* dans le développement du cerveau et du système nerveux, de même que dans les neuroblastomes, n'a pas été caractérisée.

Le gène *LMO2* a été identifié par son homologie avec *LMO1*, et est aussi associé à des translocations chromosomiques à l'origine de leucémies T chez l'humain [30, 135]. Le rôle de *LMO2*, autant dans l'hématopoïèse normale que dans la leucémie, a été décrit dans les sections précédentes, alors que ses fonctions moléculaires sont décrites dans les sections suivantes. Bien que *LMO2* soit exprimé de façon plutôt ubiquitaire, son rôle dans des tissus non-hématopoïétiques n'a pas été décrit jusqu'à maintenant. *LMO2* est aussi exprimé dans différents types de cancers avancés [283-287], et son expression est corrélée à un bon pronostic chez les patients atteints de lymphome/leucémie des cellules B [284, 288] ou de cancer du pancréas [285].

*LMO3* a été identifié par homologie avec *LMO1/2* [46], et est aussi exprimé dans le système nerveux [46]. Comme *LMO1*, *LMO3* est un oncogène associé aux neuroblastomes [289]. Les cellules de neuroblastome surexprimant *LMO3* prolifèrent

plus vite, et forment plus rapidement des tumeurs lorsqu'injectées chez la souris. De plus, l'expression de *LMO3* chez les patients atteints de neuroblastome est corrélée avec un mauvais pronostic. Au niveau moléculaire, *LMO3* s'associe au facteur de transcription neuronal HEN2, de la famille bHLH [289]. Ce complexe promeut l'expression du gène *Mash1* [290] qui est aussi dérégulé dans les neuroblastomes. La fonction de *Lmo1* et de *Lmo3* dans le développement normal du système nerveux a aussi été investigué. Les souris *Lmo1*<sup>-/-</sup> ou *Lmo3*<sup>-/-</sup> ne présentent pas de phénotype particulier, alors que l'inactivation à la fois de *Lmo1* et de *Lmo3* cause la mort des souris moins de 24h après la naissance [291]. La cause de ces décès n'a pu être établie, mais il est probable que le défaut provienne du système nerveux, considérant les patrons d'expression de *Lmo1* et de *Lmo3*.

*LMO4* a été identifié par son interaction avec LDB1 [82, 292, 293], et aussi en tant qu'auto-antigène dans le cancer du sein [294]. *LMO4* est exprimé de façon ubiquitaire, ayant été détecté dans tous les tissus testés, entre autres dans le système nerveux, les cellules épithéliales, le cerveau et le thymus [82, 292]. *LMO4* est surexprimé dans des cas de cancer du pancréas [295], mais son rôle en tant qu'oncogène dans le cancer du sein est beaucoup mieux caractérisé. *LMO4* est surexprimé dans plus de 50% des cas de cancer du sein [296], et l'expression élevée de *LMO4* est corrélée avec un mauvais pronostic [297]. Des souris transgéniques pour *Lmo4* sous le contrôle d'un promoteur spécifique de la glande mammaire développent des cancers avec une période de latence de 13-18 mois [297]. En culture, *LMO4* contrôle la prolifération et la motilité des cellules de glande mammaire [297]. *LMO4* influence aussi le cycle cellulaire de ces cellules, en modulant l'expression de régulateurs des phases G2/M [298], de même qu'en contrôlant la duplication des centrosomes et la formation du faisceau mitotique [299]. La surexpression de *LMO4* bloque la différenciation des cellules de glande mammaire en culture [296], alors que l'utilisation de *shLMO4* a l'effet inverse [297]. Puisque *LMO2* empêche la différenciation thymocytaire dans des modèles de souris transgéniques [88, 155], *LMO4* est le deuxième membre de la famille *LMO* à bloquer la différenciation du tissu dans lequel il est surexprimé, et pour lequel il est oncogénique.



Les souris *Lmo4*<sup>-/-</sup> meurent au stade embryonnaire, ou peu après la naissance, de problèmes variés du système nerveux, associés à un défaut de fermeture du tube neural [291, 300, 301]. L'inactivation conditionnelle de *Lmo4* dans différents tissus a souligné l'importance de ce gène dans la prolifération des cellules épithéliales de glande mammaire durant la gestation [302], de même que dans la régulation du contrôle sensoriel et moteur [303] et le développement de certains neurones de la rétine [304]. De plus, LMO4, de même que son orthologue unique dLMO chez la drosophile, régulent la sensibilité des souris et des mouches à la cocaïne [305, 306].

LMO4 s'associe à plusieurs différents complexes transcriptionnels, qui incluent souvent LDB1. Toutefois, l'impact de ces interactions sur le pouvoir oncogénique de LMO4 n'a toujours pas été mesuré. LMO4 se lie au bHLH proneural neurogenin2 en association avec LDB1 [307]. De plus, alors que LMO3 s'associe à HEN2 [289], et LMO2 à HEN1 et HEN2 [308, 309], LMO4 se lie au bHLH HEN1 [309]. LMO4 interagit aussi avec les facteurs de transcription DEAF1 [293], GET-1 [310], GRHL3 [311], de même qu'avec BRCA1 [312] et ER $\alpha$  [313]. De façon intéressante, dans des cellules épithéliales en culture, LMO4 est retrouvé dans un complexe transcriptionnel regroupant LDB1, GATA6 et la sérine-thréonine kinase LKB1 [314], alors que dans la moelle épinière, LMO4 s'associe à SCL, GATA2 et à LDB1 pour induire la différenciation en interneurones V2b, aux dépens des interneurones V2a [315]. Ces derniers complexes sont extrêmement similaires au complexe transcriptionnel principal décrit pour LMO2 dans le contrôle de l'hématopoïèse et dans les leucémies, tel que discuté dans la section suivante.

### **1.3.4 Le rôle transcriptionnel de LMO2 : interactions protéiques et gènes cibles**

*LMO2* et *SCL* ont été découverts presque simultanément pour leur implication dans des translocations chromosomiques à l'origine de leucémies T [30, 135-138]. De la même façon, leur inactivation chez la souris provoque des phénotypes hématopoïétiques identiques [31, 51, 52]. Pour ces raisons, il a rapidement été proposé que *SCL* et *LMO2* puissent posséder des fonctions similaires. Étant donné

que SCL est un facteur de transcription de la famille bHLH, et que certaines protéines à domaines LIM soient aussi des facteurs de transcription, l'idée que LMO2 puisse aussi influencer la transcription a rapidement été investiguée et confirmée.

L'interaction protéique entre LMO2 et SCL a été détectée dans des cellules érythroïdes [316], de même que dans des leucémies T humaines et murines [148, 317]. Parallèlement, l'interaction entre LMO2 et les facteurs de transcription GATA1/2 dans des cellules érythroïdes a aussi été rapportée [318]. Enfin, un complexe pentamérique composé de LMO2, SCL, GATA1, E2A et LDB1 a été décrit autant dans les cellules érythroïdes [318-320] que dans les thymocytes [321]. Ce complexe pentamérique est maintenant connu pour réguler la transcription de nombreux gènes cibles. Ce complexe est dynamique, en ce sens que selon le gène cible et le type cellulaire, différents partenaires peuvent s'associer au complexe pour réguler finement la transcription. De plus, plusieurs membres de ce complexe pourraient être remplacés par des membres de la même famille : LMO2 interagit autant avec GATA1 que GATA2/3 [318, 322]; LMO2 interagit autant avec le bHLH SCL que ses plus proches parents LYL1 ou TAL2 [317]; SCL peut s'hétérodimériser aux bHLH de classe I (protéines E) E2A, E2-2 ou HEB [323-326]. Ainsi, certaines composantes du complexe pentamérique pourraient être interchangeables dépendamment du contexte cellulaire ou transcriptionnel.

Dans les cellules érythroïdes, le complexe pentamérique reconnaît principalement une séquence bipartite formée d'un site *GATA* (reconnu par les protéines GATA) et d'une boîte E (reconnue par l'hétérodimère SCL/E2A) [320, 327]. Ce complexe active la transcription dans les cellules érythroïdes, et les cinq partenaires sont nécessaires à l'activité transcriptionnelle, selon un principe de « tout ou rien » [320]. Plusieurs gènes cibles répondant à ces critères ont été identifiés et leur expression est régulée par le complexe. Ces gènes sont principalement des facteurs de transcription importants pour l'hématopoïèse (*Gata1* [328], *Gata2* [72], *Nf-e2* [70], *Gfi-1b* [329], *Lmo2* [161]) ou encore des protéines de structure des érythrocytes (*α-globin* [77], *β-globin* [78] *glycophorin A* [79], *Epb42* [80]). D'ailleurs, des études de ChIP-seq contre SCL confirment que les gènes cibles de SCL dans les cellules érythroïdes sont

majoritairement impliqués dans la fonction et la différenciation des cellules érythrocytaires [327, 330]. Le complexe pentamérique régule l'expression de gènes ayant d'autres fonctions, comme le récepteur tyrosine kinase *c-kit* [331] important pour la biologie des progéniteurs hématopoïétiques. Dans ce cas particulier, le complexe pentamérique ne semble pas lier l'ADN via la séquence bipartite *GATA*/boîte E, mais est plutôt recruté par un motif de liaison à SP1, un facteur de transcription qui peut s'associer au complexe [331]. Le gène *CyclinD1* est aussi régulé par SCL via une interaction avec SP1 [332]. D'autres gènes cibles du complexe ont été identifiés, notamment *Cdkn1a* [62, 329] et *Id1* [62]. Nous avons aussi démontré la présence de LMO2 sur les promoteurs de deux micro-ARNs dans les cellules érythroïdes (**Annexe 1**).

Dans les thymocytes, le complexe pentamérique reconnaît de façon préférentielle une autre séquence bipartite formée de deux boîtes E [321]. Toutefois, des études de ChIP-seq contre SCL dans des cellules T ont montré que les sites *GATA* sont présents dans 79% des gènes cibles [330]. Dans ce type cellulaire, SCL et LMO2 sont beaucoup associés à une répression de la transcription, en plus de leur rôle d'activateur. Il a été remarqué que le phénotype des souris double transgéniques pour *Scl* et *Lmo* (blocage de la différenciation, apparition de leucémie) était semblable à celui observé chez les souris inactivées pour le gène *E2a* [280]. Les protéines E2A et HEB jouent un rôle important dans la différenciation thymocytaire [333, 334], alors que SCL et LMO2 ne sont pas normalement exprimés dans ces cellules. Puisque SCL s'associe physiquement à E2A et HEB, il a été proposé que l'expression ectopique de SCL et LMO2 dans les thymocytes puisse interférer avec l'activité normale des protéines E. Dans les leucémies induites par SCL et LMO, plusieurs gènes cibles connus de E2A et HEB sont réprimés, notamment *Rag1/2* [335], *pre-Ta* [37, 335, 336], *Tcra* [337], *Cd4* [37, 335], *Cd5* [335] et *Cip1* [338]. SCL et LMO2 s'associent aux promoteurs de ces gènes cibles [37, 335, 336] et recrutent des co-répresseurs tels que SIN3A/HDAC [335], ce qui transforme le complexe activateur en répresseur. D'ailleurs, les leucémies *Scl/Lmo2* sont sensibles au traitement avec des inhibiteurs d'histones déacétylases [335].

L'interférence fonctionnelle avec E2A et HEB n'est probablement pas le seul mécanisme par lequel SCL et LMO2 affectent la transcription dans les leucémies T. D'ailleurs, une étude de ChIP-seq contre SCL dans les cellules T a récemment montré qu'une proportion importante de gènes cibles de SCL possédaient des motifs de liaison à RUNX ou à ETS1, et que ces protéines étaient importantes pour la liaison de SCL à ces gènes cibles [330]. Cette caractéristique a aussi été observée dans les progéniteurs hématopoïétiques multipotents, où une grande proportion de gènes sont liés à la fois par le complexe pentamérique, par RUNX1 et par ERG et FLI-1, deux protéines de la famille ETS [339]. De façon intéressante, les motifs de liaison à RUNX1 et à ETS1 sont beaucoup moins présents dans les gènes cibles de SCL identifiés dans la lignée érythroïde [330]. Ces études suggèrent que le contexte cellulaire détermine substantiellement les fonctions transcriptionnelles du complexe SCL/LMO2, et que RUNX et/ou ETS1 puissent jouer un rôle dans l'activité oncogénique induite par SCL et LMO2.

Contrairement à ce qui avait été proposé initialement, SCL et LMO2 ne sont pas exclusivement activateurs dans l'hématopoïèse normale et répresseurs dans les leucémies. Dans les leucémies T, SCL et LMO2 régulent positivement l'expression des gènes *Raldh2* [322] et *NKX3.1* [340], en étant recrutés sur un site *GATA* unique via la protéine GATA3. De la même façon, dans l'hématopoïèse normale, le complexe inhibe l'expression du gène *Cdkn2a* via l'association à des boîtes E [72], et réprime aussi l'expression de *miR-142* en s'associant à une séquence bipartite *GATA*/boîte E [341]. D'ailleurs, des études de ChIP-chip et de ChIP-seq, combinées à des analyses d'expression, ont montré que SCL pouvait activer aussi bien que réprimer l'expression de différents gènes cibles, autant dans les cellules érythroïdes [330] que dans les cellules T [330, 342]. Toutefois, le complexe SCL a tendance à activer la transcription de ses gènes cibles dans les cellules érythroïdes [327, 330, 343, 344], alors que dans les cellules T, la proportion de gènes activés et réprimés est semblable [330, 342]. De plus, l'identité des gènes cibles est très différente dans les deux types cellulaires. Selon le type cellulaire, les gènes cibles régulés par SCL sont principalement associés à la fonction et à la différenciation des cellules érythroïdes et des thymocytes, respectivement [327, 330].

### 1.3.5 Possibilité d'un rôle non-transcriptionnel pour LMO2

Au niveau moléculaire, le seul rôle décrit pour LMO2 est son implication dans le complexe transcriptionnel pentamérique SCL/LMO2/GATA/LDB1/E2A. Quelques autres partenaires protéiques de LMO2 ont été identifiés, presque tous étant des protéines impliquées dans la transcription et/ou le remodelage de la chromatine. Précisément, des interactions ont été rapportées entre LMO2 et la protéine BEX2 [308], les bHLH HEN1/2 [308, 309], le facteur de transcription de la famille ETS ELF2 [345] et l'histone lysine déméthylase KDM5A (JARID1A ou RBP2) [346]. Bien que ces partenaires protéiques puissent influencer la fonction de LMO2 dans l'hématopoïèse normale autant que dans la leucémogénèse, le rôle fonctionnel de ces interactions n'a pas été décrit. Ces partenaires pourraient être associés au complexe pentamérique, mais pourraient aussi lier LMO2 indépendamment de SCL.

L'expression de LMO2 est détectée dans plusieurs tissus en l'absence de SCL, par exemple dans le poumon, le rein et le foie [44], suggérant que LMO2 pourrait avoir des fonctions indépendantes de son association au sein du complexe pentamérique. De plus, puisque LMO2 ne lie pas l'ADN directement, on peut penser que LMO2 puisse s'associer à des complexes protéiques non-transcriptionnels, régulant d'autres processus cellulaires. D'ailleurs, par criblage de double hybride, un partenaire non-transcriptionnel de LMO2 a été identifié : MLLT4 (AF6) [347], une protéine impliquée dans les jonctions cellule-cellule [348]. De façon intéressante, cette protéine est un partenaire de translocation récurrent de l'oncogène *MLL* (*Mixed-lineage leukemia*) [349]. Les réarrangements chromosomiques impliquant *MLL* sont associés principalement aux leucémies B et myéloïdes, et plus rarement aux leucémies de type T-ALL. Toutefois, le gène de fusion *MLL-AF6* a été récemment détecté chez un patient atteint de T-ALL [350]. Bien qu'AF6 puisse être une protéine importante dans la leucémogénèse, de par son association à la fois à *MLL* et à LMO2, son rôle dans ce processus n'a toujours pas été élucidé.

Dans des modèles de souris transgéniques, *Lmo2* est un oncogène plus puissant que *Scl*, puisque sous le contrôle d'un même promoteur *Cd2*, la surexpression de *Lmo2*

seul [154], mais pas *Scl* [149], génère des leucémies. De plus, la surexpression de LMO2 [88, 155], mais pas de SCL [148, 149, 151, 152], perturbe la différenciation thymocytaire. Ces observations pourraient être en accord avec l'hypothèse que LMO2, en plus de son rôle transcriptionnel en complexe avec SCL, puisse posséder des fonctions additionnelles.

Dans cette optique, j'ai identifié une nouvelle fonction pour LMO2 dans la réplication de l'ADN, tel que présenté dans le **Chapitre 3**. Des publications récentes pointent en faveur d'un rôle direct pour certains facteurs de transcription de différentes familles (*c-MYC*, *MYB*, *RB/E2F*, *HOX*) dans la réplication de l'ADN. Ces protéines interagissent directement avec des protéines de la machinerie de réplication et sont détectées aux origines de réplication sur l'ADN. De façon surprenante, ces facteurs de transcription sont aussi des oncogènes majeurs dans les leucémies aigües, soulevant la possibilité que ce nouveau rôle puisse faire partie de leur potentiel de transformation dans les leucémies. Le rôle de ces facteurs de transcription dans la réplication de l'ADN est discuté dans le **Chapitre 2**.

## 1.4 Problématique et choix de la méthodologie

Tel que décrit précédemment, LMO2 est un facteur de transcription dont l'expression aberrante dans les thymocytes génère des leucémies T-ALL. Bien entendu, le rôle transcriptionnel de LMO2 est crucial dans la transformation des thymocytes. Notamment, l'expression aberrante de LMO2 perturbe le programme transcriptionnel de nombreux gènes impliqués dans la différenciation thymocytaire [37, 335-338], avec comme conséquence un blocage de la différenciation des thymocytes au stade double négatif (DN) [88, 155]. De plus, LMO2 provoque l'auto-renouvellement des cellules DN3 en modifiant leur transcriptome : des gènes spécifiques des cellules souches hématopoïétiques se retrouvent activés dans ces cellules aux dépens des gènes de différenciation thymocytaire [114]. Certains de ces gènes (*c-Kit* [331], *pre-T $\alpha$*  [37, 335, 336], *Nf-e2* [70]) sont des cibles directes de LMO2.

Bien que la démonstration du rôle transcriptionnel de LMO2 dans la leucémogénèse ne soit plus à faire, on peut se demander si ses gènes cibles suffisent à expliquer

totallement son potentiel oncogénique. Le cancer est un événement multi-étapes : la cellule doit survivre, proliférer, modifier sa réponse à l'environnement, éteindre ses différents mécanismes de contrôle. Dans la leucémogénèse induite par LMO2, d'autres événements collaborateurs sont requis, indiquant que LMO2 ne contrôle pas tous les aspects du cancer. En plus de l'inhibition de la différenciation et de l'induction d'auto-renouvellement, LMO2 module-t-il d'autres processus cellulaires? Si oui, est-ce que ces effets de LMO2 dépendent de son rôle transcriptionnel? Par exemple, LMO2 cause une prolifération accrue des thymocytes au stade pré-leucémique, autant dans des modèles de souris [88] que dans les essais de thérapie génique chez l'humain [143]. Pourtant, LMO2 ne semble pas contrôler la transcription des gènes du cycle cellulaire. Afin de préciser le rôle transcriptionnel de LMO2, et de déterminer si LMO2 pourrait avoir des fonctions indépendantes de la transcription, il est utile de mieux décrire son interactome.

Certains partenaires protéiques de LMO2 sont bien connus, comme SCL et LDB1. Il est possible que de telles interactions soient bien caractérisées en raison de leur forte affinité, ou encore de leur fréquence dans la cellule. En accord avec cette hypothèse, nous avons déterminé que 70% du pool nucléaire de LMO2 est associé à SCL dans les cellules érythroïdes en culture [177]. Toutefois, les interactions plus rares, ou de plus faible affinité, pourraient être informatives quant à la régulation fine de la fonction de LMO2. Certaines interactions peuvent avoir lieu dans un type cellulaire particulier, dans une organelle précise, à un moment précis (cycle cellulaire, réponse à l'environnement). Ces interactions ne peuvent pas être détectées par toutes les techniques servant à identifier les interactions protéine-protéine. Les différentes techniques employées pour identifier les partenaires protéiques, de par leurs principes et leurs méthodologies différentes, répondent à des questions biologiques différentes. À titre d'exemple, une brève comparaison sera effectuée entre la technique de double hybride et l'analyse par spectrométrie de masse.

La spectrométrie de masse est un outil très puissant pour étudier les complexes protéiques dans des contextes biologiques particuliers. Par exemple, notre équipe l'a employée afin d'identifier les protéines qui s'associent à SCL [329] et celles qui sont

recrutées aux promoteurs des gènes cibles de SCL et LMO2 (**Chapitre 4**). Différentes méthodes de purification sont requises avant de soumettre l'échantillon à la spectrométrie de masse. Ainsi, les conditions de purification et/ou d'extraction protéique peuvent détruire certaines interactions moins fréquentes, plus transitoires ou de faible affinité. À l'opposé, avec la technique de double hybride, les interactions sont testées en paires, en l'absence de toute manipulation biochimique, ce qui permet la détection d'interactions plus transitoires et plus rares. La spectrométrie de masse présente toutefois l'avantage d'identifier des complexes protéiques en préservant le contexte biologique dans lequel ces interactions se retrouvent, en plus d'identifier des interactions protéiques qui ne soient pas nécessairement directes. Bien que la technique de double hybride puisse générer un plus grand répertoire d'interactions protéiques directes, cette méthode n'offre aucune information quant au contexte spatio-temporel au cours duquel ces interactions ont lieu.

En dépit des critiques au sujet de la technique de double hybride (beaucoup de faux positifs, interactions détectées hors contexte), je soutiens que cette méthodologie a sa place dans l'étude d'interactions protéine-protéine, précisément parce qu'elle n'est pas biaisée spatio-temporellement. Évidemment, seule la combinaison de plusieurs techniques, dans leurs limitations respectives, pourrait éventuellement permettre de décrire extensivement l'interactome d'une protéine donnée. Pour ces raisons, et pour d'autres contraintes biologiques énumérées dans la **Section 5.1** (Discussion), j'ai choisi d'identifier les partenaires d'interaction de LMO2 par la technique de double hybride de levure. Je propose que l'exploitation de la technique de double hybride puisse être un moyen efficace de mieux caractériser la fonction de LMO2 dans l'hématopoïèse normale et la leucémie.

## 1.5 Objectifs de la thèse

LMO2 est un oncogène important dans les leucémies T-ALL, en plus d'être un régulateur majeur de l'hématopoïèse normale. Son rôle au sein de complexes transcriptionnels est plutôt bien décrit, mais ne permet pas d'expliquer la totalité des phénotypes observés dans les cellules transformées par LMO2. Puisque LMO2 ne lie



pas l'ADN directement, il n'est pas exclu que cette protéine d'échafaudage puisse être retrouvée dans différents complexes multiprotéiques régulant des fonctions cellulaires diversifiées. Dans cette optique, l'objectif principal de ma thèse doctorale était d'identifier les partenaires protéiques de LMO2 par un criblage de double hybride chez la levure, et de caractériser de nouvelles fonctions potentielles de LMO2 dans les cellules hématopoïétiques.

Ainsi, mon criblage de double hybride a révélé des interactions entre LMO2 et des protéines de réplication de l'ADN. J'ai démontré que LMO2 contrôlait la réplication de l'ADN dans les cellules hématopoïétiques et ce, indépendamment de son association au sein de complexes transcriptionnels (**Chapitre 3**). De plus, la surexpression de LMO2 dans les cellules hématopoïétiques perturbe spécifiquement le cycle cellulaire des progéniteurs thymocytaires, qui sont des cibles de transformation par LMO2. Cette nouvelle fonction pour LMO2 permet donc de mieux caractériser la fonction de LMO2 dans les cellules hématopoïétiques et possiblement durant la leucémogénèse.

L'analyse plus globale des interactions entre LMO2 et ses partenaires potentiels (**Chapitre 4**) suggère que LMO2 recrute plusieurs protéines gouvernant la transcription de ses différents gènes cibles. De plus, LMO2 semble être impliqué dans d'autres processus cellulaires, en plus de sa fonction transcriptionnelle et de sa fonction nouvellement décrite dans la réplication de l'ADN. La caractérisation de ces nouvelles fonctions pourra nous éclairer davantage sur les mécanismes moléculaires de transformation par LMO2.

## **2 Cross-regulation of transcription and replication, two DNA-dependent processes**

Marie Claude Sincennes<sup>1</sup>, Trang Hoang<sup>1,2</sup>

<sup>1</sup> Institute of Research in Immunology and Cancer (IRIC) – University of Montreal,  
Montréal, Québec, Canada H3C 3J7

<sup>2</sup> Pharmacology, Biochemistry and Molecular Biology Departments, University of  
Montréal, Québec, Canada H3C 3J7

This chapter corresponds to a manuscript to be submitted (review article)

## **2.1 Contribution des co-auteurs**

MCS : Rédaction du manuscrit

TH : Rédaction du manuscrit

## 2.2 Summary

Oncogenic transcription factors that are misregulated in acute leukemias typically control complex gene expression networks that regulate major aspects of cell survival, proliferation and differentiation. Less well known is the possibility that these oncogenes, by their DNA-binding abilities, may also regulate other DNA-dependent processes, more specifically DNA replication. Genome-wide studies in *Drosophila* and mammalian cells underscored the proximity between replication origins and transcriptional promoters. This could be explained in part by the local open chromatin structure in transcriptional units. In addition to this hypothesis, emerging evidence favours a more direct involvement of transcription factors in DNA replication. The examples of c-MYC, MYB, RB/E2F and HOX proteins are discussed in this review. These transcription factors are major oncogenes in acute leukemias, suggesting that their overexpression could perturb cell behaviour by at least two molecular mechanisms, i.e. transcription and DNA replication.

## 2.3 Introduction

Upstream regulators of transcription that control cell fate in the hematopoietic system are often dysregulated in acute leukemias, as exemplified by the frequent implication of oncogenic transcription factors and chromatin modifiers in recurrent chromosomal translocations [351, 352]. This naturally led to the postulate that oncogenic transcription factors establish gene expression programs that control cell survival, cell proliferation and differentiation as well as cellular metabolism, considered as hallmarks of cancer [94]. Indeed, one of the best studied oncogenic transcription factors, c-MYC, has a major impact on all these processes, via its capacity to control gene expression, and therefore, c-MYC is viewed as a master regulator of hematopoietic cell fate. Interestingly, many transcription factors directly bind to DNA, raising the possibility that these proteins may influence other DNA-dependent processes. For example, emerging evidence indicates that c-MYC as well as HOX proteins may also control DNA replication, as discussed in this review. In parallel, studies on DNA replication also point to an overlap between replication initiation sites and transcription units.

A convergence between different DNA-dependent processes has already been demonstrated. It is well known, for example, that chromatin-remodelling complexes are needed to regulate DNA accessibility during transcription, replication and repair, and that the same chromatin modifiers are necessary for these different processes (for review, see [353, 354]). Like chromatin modifiers, some DNA-binding proteins are also important for more than one DNA-based process. The coupling between DNA replication and DNA repair, for example, is well documented (for review, see [355]). On the other hand, the link between transcription and other DNA-dependent processes remains to be further investigated. As discussed below, recent evidence raises the possibility that the coupling between transcription and replication may be more direct than previously anticipated, particularly in metazoans.

#### **2.4 Proximity between replication origins and gene promoters**

Replication origins are regions in the genome at which DNA replication is initiated. The origin recognition complex (ORC) directly binds to replication origins [356], followed by the association of CDC6, CDT1 and MCM2-7, which collectively form pre-replication complexes (pre-RC) at each origin. During the G1/S transition, at the time of initiation, CDC6 and CDT1 leave the origins, while other replication proteins are recruited, leading to the firing of replication origins and the initiation of DNA replication (reviewed in [357]). In *S. cerevisiae*, origins contain an ARS (autonomously replicating sequence) consensus sequence (ACS) that is essential, but not sufficient, for origin activity [358]. Unlike *S. cerevisiae*, origins in *S. pombe* are not sequence-specific but are characterized by AT-rich regions [359]. In higher organisms, however, there is no consensus DNA sequence for replication origins [358]. Many groups estimated that in mammals, more than 30 000 replication origins exist in order to replicate all chromosomes during the S phase of normal duration [360-363]. To date, less than 4% of these potential replication initiation sites are known, and most of them have been identified by genome-wide studies, which have to be validated individually to confirm that they really possess origin features. To date, the small number of well-characterized origins is an important issue for studying DNA replication in mammals.

Replication origins are mainly identified by purification of nascent DNA strands or by chromatin immunoprecipitation (ChIP) with antibodies against ORC and MCM proteins. The ChIP technique results in the identification of pre-RC binding sites, but do not necessarily reflect active origins. In addition, ORC proteins have replication-independent functions (reviewed in [364]), suggesting that ORC binding sites may not exclusively reflect replication origins. Strictly speaking, active replication initiation sites in mammals are identified by the nascent strand purification. Hence, the choice of methods could result in differences in the type of origins purified (i.e. “potential” origins *versus* “active” origins).

One could argue that if replication and transcription are intimately linked, replication origins should reside near gene promoters. In agreement with this hypothesis, studies in *Drosophila*, mouse and human cells indicated frequent colocalization between replication origins and promoter regions, although replication origins are also found elsewhere.

In *Drosophila*, it appears that most if not all characterized replication origins have been identified by ORC binding. MacAlpine et al [365] used a high resolution genomic microarray covering the left arm of chromosome 2, combined with ChIP against RNA PolII and ORC, and showed a significant overlap between ORC and RNA Pol II binding sites. Consistent with this, in the *Drosophila* cell line Kc167, genome-wide ChIP-chip was performed against ORC2, and the authors determined that 89% of ORC binding sites overlapped with transcriptional start sites [366]. Moreover, by comparing the ChIP-chip data with expression microarrays, the authors showed that ORC binding sites were greatly enriched at the promoters of actively transcribed genes. The correlation between replication and gene expression has also been observed in *Drosophila* polytene tissues, in which underreplicated regions correlate with the absence of gene expression, both in heterochromatin and in euchromatin [367, 368]. Recently, the modENCODE Consortium integrated more than 700 data sets mapping many chromosomal features such as transcripts, transcription factors, histone modifications and replication proteins [369]. They also correlated ORC binding with transcriptional start sites. Moreover, they analyzed ORC

binding in highly occupied target (HOT) regions, bound by many different transcription factors. They observed that the enrichment for ORC binding sites increases with the complexity of transcription factor-bound regions. This suggests that ORC proteins are more likely to bind regions occupied by many transcription factors.

In mammals, most of the well characterized DNA replication origins are in proximity to a gene promoter, namely *LaminB2/Timm13* [370], *Hprt* [371], *c-Myc* [372], *Hsp70* [373],  *$\beta$ -globin* [374, 375], *G6pd* [376], *Top1* [377], *Mcm4/Prdkc* [378], *Dbf4* [379], *Fmr1* [380], *Ifna2* [381], *rDNA* [382], *Igh* [383, 384] and *Rdh/Ori6* [385]. Nonetheless, other well characterized origins also localize inside genes (*Dnmt1* [386], *Rpe* [387], *Ors8/Elmo1* [388]) or downstream (*Dhfr* [389, 390], *Gmcsf* [391]). During the last three years, these observations have been confirmed by mapping replication origins in mammalian cells through genome-wide studies.

In murine asynchronous ES cells, Sequeira-Mendes et al [363] hybridized nascent DNA strands on tiling arrays covering 0,4% of the mouse genome. By purifying nascent fragments that were 300-800nt in length, they identified 97 replication initiation sites, 85% of which mapped to annotated transcriptional units. Specifically, 44% were found at promoter regions. They repeated their hybridization with shorter fragments (100-600nt) and found that 38 replication initiation sites could be reproducibly identified by the two methods of purification. In this subset, 97% mapped to transcriptional units and 78% overlapped with promoter regions. Moreover, by randomly selecting a few origins, the authors demonstrated that replication initiation sites are juxtaposed to transcriptional start sites in almost all cases studied. The authors also asked the question if origin-containing promoters were active in transcription. Strikingly, they found with a significant degree of correlation that genes having a replication origin in their promoters were expressed early in development. The authors therefore suggest that replication origins are often associated with promoters of actively transcribed genes. Consistent with these data, it was previously reported that replication origins within the *Igh* locus were only active in cells in which the locus is expressed [383]. It will be interesting to examine this

phenomenon at the single cell level. Recently, a similar genome-wide study was performed in three different murine cell lines (ES, P19 and MEF cells) [392], and the authors observed that replication origins are significantly overrepresented at transcriptionally active promoters and at transcriptional start sites in all cell lines.

The same type of study was also performed in human cells. Cadoret et al [361] purified short nascent DNA strands from asynchronous HeLa S3 cells that were hybridized on the Human Encyclopedia of DNA Elements (ENCODE) ChIP-on-chip microarray, containing representative genomic regions covering about 1% of the human genome [393]. They found a strong correlation between origin localization and GC-rich regions, which are associated with high gene density. Moreover, nascent strands were very frequently localized within or close to CpG islands. Finally, 34% of the identified nascent strands overlap with gene promoter regions.

Recently, Karnani et al [394] took a genomic approach to identify replication initiation sites in HeLa cells. They isolated origins by two different methods, i.e. purification of nascent strands coupled with lambda exonuclease digestion, and anti-5-bromodeoxyuridine (BrdU) immunoprecipitation. After hybridization of replication initiation sites on high-density tiling arrays representing the ENCODE region, they identified in total 985 replication initiation sites, of which 150 were identified reproducibly by both techniques. They validated their data by quantitative PCR and by ChIP against ORC2 and ORC3 performed on a few randomly selected origins, showing that more than 65% of potential initiation sites identified by one method or the other were bound by ORC proteins, while about 90% of origins identified by the two methods were positive for ORC binding. They focused their analyses on the 150 reproducibly identified origins, and showed that 68% were located at transcriptional start sites within 10kb. Furthermore, a significant proportion of initiation sites overlapped with RNA Pol II binding sites.

Together, these studies point to a variable but substantial degree of overlap between transcription units and replication origins. These variations could be due in part to the fact that the various chips cover only a small proportion of the genome.



## 2.5 Epigenetic factors

DNA accessibility must be controlled for transcription or DNA replication to occur. However, exposed DNA is more fragile and the cell must have developed mechanisms to restrict DNA accessibility. One efficient way is to carry out many DNA-dependent processes at the same time, on the same opened DNA portion. An example in favor of this hypothesis is the nucleosome positioning nearby transcriptional promoters and replication origins. It is well established that nucleosomes preclude the binding of some transcription factors to DNA ([395, 396], reviewed in [397]). Similarly, replication origins are also depleted of bulk nucleosomes in *Drosophila* [366] and in yeasts [398-401]. Moreover, rapid nucleosome turnover occurs in the vicinity of ORC binding sites as in actively transcribed regions [402].

Transcription factors can recruit histone modifiers to their target genes to change the chromatin environment, allowing the assembly of the RNA pol II complex. One may argue that this open chromatin environment in promoter regions can also be favorable to the binding of replication proteins to origins. This is probably the case for many origins. In *Drosophila*, ORC binding was correlated with a set of specific histone marks, mostly H3 and H4 acetylation patterns, as well as with the histone variant H3.3 [369]. In HeLa cells, origin localization is also correlated with active histone marks, namely H3Ac, H3K4 di- and tri-methylation, a chromatin signature associated with active transcription [394]. In mammalian cells, histone H4 at replication origins is acetylated (residues K5, K12) and mono-methylated (residue K20) by HBO1 and PR-SET7, respectively, and these chromatin marks are important for pre-RC loading on origins [403, 404].

Despite the above, replication origins are not exclusively defined by epigenetic marks. Cadoret and Prioleau evaluate that in HeLa cells, about 45% of origins do not display H3 and H4 acetylation, H3K4 mono, di and trimethylation [405]. Additionally, in mice, the same origins are active in both the transcriptionally active and the heterochromatic, silent X chromosomes [406]. Accordingly, in B cells,

after *Igh* locus rearrangement, only one rearranged allele is expressed, whereas origin activity is detected on both alleles [383].

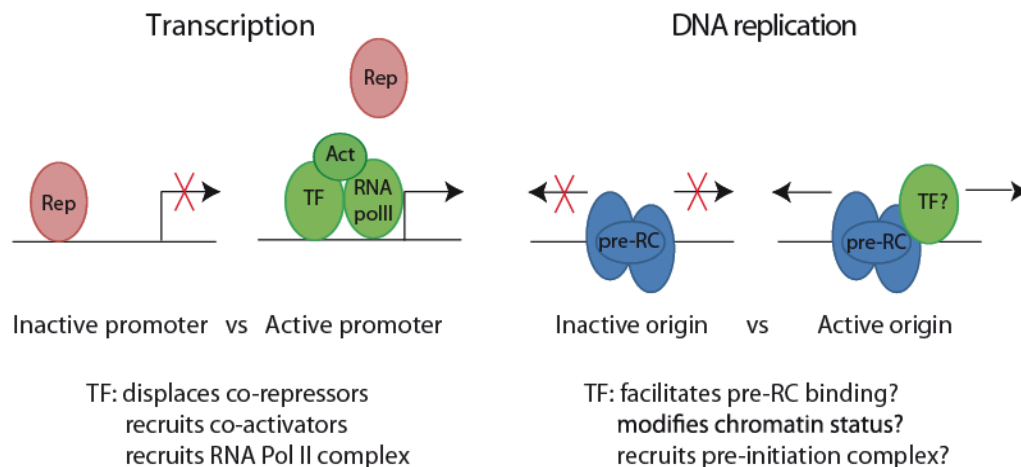
How is DNA duplicated in regions of dense chromatin? In fact, ORC proteins interact with heterochromatin protein 1 (HP1) and both have been implicated in gene silencing as well as in heterochromatin formation and maintenance [364]. However, the link between this property of ORC proteins and their function in DNA replication cannot be generalized. In *S. cerevisiae*, the gene silencing and replication functions of ORC proteins can be uncoupled by specific mutations, arguing that these two functions of ORC are independent [407-409]. However, the *S. pombe* counterpart of HP1, Swi6, is important for the initiation of replication in silent chromatin [410]. In metazoans, the different functions of ORC proteins have been studied independently, but the association of ORC proteins on heterochromatin could be important both for heterochromatin formation and for the firing of replication origins in these regions. It is improbable that transcription factors influence the replication in dense chromatin regions, even if the cell must have developed mechanisms to bring accessibility to the replication machinery at these sites. However, it is interesting to note that in *Drosophila*, ORC binding sites have been correlated with transcriptional start sites in both euchromatin and heterochromatin [369], even if *Drosophila* may be unique in that regard.

The proximity between replication origins and transcriptional units could be explained in part by the local chromatin structure, but other factors must influence origin selection, since active histone marks are not always present on replication origins.

## **2.6 Origin selection: another task for transcription factors?**

In higher organisms, most origins do not fire in each cell cycle, and studies at specific and genome-wide loci in metazoans evaluate the efficiency of firing to be around 5-20% [358, 392]. Some origins are active only in a given cell type or at specific developmental stages, while others seem to be activated in stress conditions. This allows the cell to strictly coordinate the timing and duration of S phase in response to

different environmental cues. What determines which origin will be active in a given cell type or in specific conditions? This remains an open question. In metazoans, ORC proteins do not exhibit sequence specificity in their DNA binding properties, and other replicating proteins depend on ORC for their recruitment to replication origins. In addition, ORC and MCM proteins are very abundant in the cell, probably to allow the licensing of all potential origins during G1 [358]. Since the pre-replication complex formed by ORC and MCM proteins seems to assemble on every replication origin, without any selection, this suggests that some “limiting factors” exist to regulate origin activation [358, 411]. In that regard, a parallel can be made with the basal transcription machinery, which does not show promoter specificity (**Figure 2.1**). In transcription, some promoters are active, while others are not, depending on cell type or extracellular signals from the microenvironment, as cells receive differentiation, stress or proliferation signals. The selection of promoter activity is in large part accomplished by transcription factors, ubiquitous or cell-type specific, that bind to the promoter of their target genes following their activation or recruitment.



**Figure 2.1 Parallel between transcription factor activity and origin activity.**

In transcription, the activity of a promoter is determined by the specific binding of transcription factors (TF) that mediate protein-protein interactions leading to the recruitment of RNA Pol II complex. In DNA replication, transcription factors could regulate origin activity by different proposed mechanisms. Rep: co-repressor; Act: co-activator; pre-RC: pre-replication complex.

In replication, “specificity” factors that could regulate origin activity in different conditions have not yet been identified. The fact that the majority of replication sites identified so far lie next to a transcriptional unit may not be a coincidence. One may envisage the possibility that a subset of transcription factors could also regulate origin selection by their dual interactions with DNA and with one or several proteins from the basal replication machinery. This dual role has been observed in viral genomes as well as in yeast, where transcription factors can directly influence DNA replication (reviewed by [411]). In mammalian cells, potential candidates for such a dual activity are oncogenic transcription factors. Indeed, these factors are tightly regulated in normal conditions and their dysregulated expression leads to the development of cancer. While their transcriptional properties on cell cycle genes are well demonstrated, recent advances suggest that these transcription factors can also be closely involved in DNA replication, particularly in metazoans (**Table II**). Inappropriate control of DNA replication could lead to genomic instability and cancer [412, 413].

## 2.7 c-MYC

More than 20 years ago, the oncoprotein c-MYC was proposed to directly regulate DNA replication [414], although these results were controversial [415]. Recently, c-MYC associated proteins were identified by mass spectrometry after affinity purification of MYC-containing protein complexes from the H1299 human tumor cell line [416]. MCM5 and MCM7 were co-precipitated with c-MYC. By immunoprecipitation, MYC was found associated with proteins from the pre-replication complex, namely MCM proteins, CDC6, CDT1 and ORC2. While it is not clear whether these interactions are direct or not, these observations indicate that c-MYC can be found in a complex with replication proteins. Additionally, c-MYC colocalizes with BrdU foci in primary human fibroblasts during early S phase, and c-MYC binds to the *c-Myc* and *Lamin B2* replication origins. Elegant experimental strategies were employed to separate the role of c-MYC in transcription *versus* DNA replication [416], notably by the use of *Xenopus* cell-free extracts replication assays, in which there is no transcription nor protein synthesis. These results demonstrate that

TF	Binding to pre-RC	Origin binding	Presence BrdU foci	Functional assays	Ref.
c-MYC	MCM5/7	c-Myc	yes	overexpression accelerates S phase progression	[416, 419]
	Other pre-RC proteins	Lamin B2		depletion impairs replication in <i>Xenopus</i> cell-free extracts depletion impairs replication in mammalian cells	
MYB	ORC2	ACE3 (Dm)	yes	deletion impairs chorion gene amplification	[422, 423]
E2F	ORC1/2	ACE3 (Dm) Lamin B2 GM-CSF c-Myc	N/A	mutation impairs chorion gene amplification	[431-433]
RB	ORC1/2 MCM7	Lamin B2 Hsp70 $\beta$ -globin GM-CSF	N/A	inhibits replication in <i>Xenopus</i> cell-free extracts	[425-430, 434]
HOXC10	N/A	Lamin B2	N/A		[437]
HOXA13, HOXD11	N/A	Lamin B2 MCM4 Fmr1 Top1 c-Myc	N/A N/A		[439]
HOXC13	CDC6 ORC1/2	Lamin B2 MCM4 Top1	yes	depletion has no effect on cell cycle and BrdU incorporation	[437, 438]
HOXD13	CDC6 geminin	Lamin B2 MCM4 Fmr1 Top1 c-Myc	N/A	overexpression stimulates binding of pre-RC (CDC6, MCM7) to origins overexpression stimulates replication of plasmid DNA overexpression in 293 cells increases BrdU incorporation (rescued by shRNA)	[439]

**Table II. Transcription factors influencing DNA replication**

List of transcription factors (TF) and their interaction with proteins from the pre-replication complex (pre-RC) as well as with DNA replication origins. Their presence in BrdU foci is also specified, as well as their impact on DNA replication by functional assays.  
*Dm: Drosophila melanogaster*

c-MYC influences DNA replication by a non-transcriptional mechanism.

One proposed possibility is that c-MYC attenuates the ATM/ATR checkpoint response [417]: c-MYC interacts with MIZ1, perturbing the MIZ1-dependent chromatin recruitment of TOPBP1, an activator of ATR kinase function at stalled forks [418]. Consistent with this, Werner DNA helicase protein (WRN) is required to escape ATR-dependent DNA damage response following c-MYC overexpression [419]. Alternatively, c-MYC could be more directly involved in the replication process, possibly by regulating origin activity/selection. Indeed, c-MYC influences the number of origins to be fired without affecting the speed of DNA synthesis [416]. In primary human fibroblasts, overexpression of *c-Myc* accelerates S phase progression [419], as measured by BrdU incorporation after G1/S synchronization by double thymidine block. Inversely, a rat cell line from which c-MYC is depleted presents a delay in S phase completion after release from G1/S block.

## 2.8 MYB

The oncogenic transcription factor MYB promotes the G1/S transition in vertebrates [420], at least by activating the transcription of cell cycle genes [421]. A non-transcriptional role for MYB in DNA replication has been proposed in *Drosophila* at the *chorion* locus [422], a well-characterized model to study DNA replication in the fly. MYB was detected at the *chorion* locus, containing different replication origins. The presence of MYB on the specific *ACE3* origin was confirmed by EMSA and by ChIP from primary ovary cells. In the *Drosophila* Kc cell line, MYB interacts with ORC2 as revealed by immunoprecipitation. Deletion of MYB binding sites from the *ACE3* locus reduced the levels of DNA amplification from this locus. MYB is colocalized with BrdU foci in wildtype flies [423]. In addition, the absence of *Myb* in *Drosophila* ovaries does not affect the formation of ORC2-positive foci, but prevents replication initiation from these foci, as demonstrated by the absence of BrdU staining [422]. Thus, MYB could regulate DNA replication at the *chorion* loci and maybe elsewhere, as its colocalization with BrdU foci is observed in different cell types [423]. The same group also observed that deletion of *Myb* in *Drosophila* causes

lethality and genomic instability [423], but the link between this phenotype and a possible role for MYB in DNA replication remains to be established. At this point, it is not known whether MYB regulates replication at other loci and whether this function of MYB is conserved in mammals.

## 2.9 RB and E2F

The retinoblastoma protein (RB) is an important cell cycle regulator of the G1/S transition. This tumor suppressor interacts with E2F and is part of a complex that represses the expression of its target genes (reviewed in [99]). RB negatively regulates DNA replication by repressing the transcription of many replication genes [424, 425]. In addition, accumulating evidence suggests that RB could also influence DNA replication by a more direct mechanism, but the role of RB in this process remains controversial. For example, the effect of RB on DNA replication of *Xenopus* cell-free extracts *in vitro* has been reported to be negative [426] or null [425], according to two independent studies. Despite these discrepancies, evidence favors a role for RB in the regulation of DNA replication, as described below, although this role needs to be further investigated.

Different protein-protein interactions involving RB and replication proteins have been reported: RB associates with the large subunit of Replication factor C (RF-C) [427] and MCM7 [426]. While hypo-phosphorylated RB binds to ORC1 [428], the phosphorylated form directly stimulates DNA polymerase alpha activity [429]. In addition, RB is detected at replication foci in primary cells during G1 [430]. Recently, E2F1 and RB have been detected to the *Lamin B2* and the *GM-CSF* replication origins in HeLa cells [428], and E2F2/3/4 were associated with the *c-Myc* origin in Raji cells [431]. However, the function of the association of RB and E2F with replication proteins at origins remains to be further characterised.

The function of E2F in DNA replication has been investigated in the fly. In *Drosophila*, mutations in E2F impair *chorion* gene amplification and prevent ORC2 localization to the amplifying loci [432], suggesting a role for E2F in the replication of this locus. Accordingly, immunoprecipitations in *Drosophila* ovary extracts

indicate that a complex is formed between Rbf (a RB homologue), E2F1 and ORC1/2 in these cells, and that E2F1 is recruited to the *chorion* replication origin [433]. These results indicate that E2F could regulate DNA replication specifically to the *chorion* locus. In mammalian cells, a role for RB in specific replication loci, after irradiation-induced DNA damage, has also been investigated. In IMR-90 and primary human T cells synchronized with hydroxyurea and  $\gamma$ -irradiated, RB is dephosphorylated and associates with 3 different replication origins (*Lamin B2*, *HSP70* and  *$\beta$ -globin*) [434]. The authors demonstrated that the association of RB on these origins had no effect on the transcriptional level of the nearby genes *HSP70* and *Lamin B2*, whereas  *$\beta$ -globin* was not expressed at all in these cells, suggesting that the association of RB on these sites is not transcriptionally-linked. Strikingly, the timing of association between RB and these replication sites parallels the order in which these origins fire in the cell, suggesting a possible role for RB in origin firing following DNA damage.

## 2.10 HOX proteins

To date, at least 5 different HOX proteins have been detected at replication origins (HOXA13, HOXC10, HOXC13, HOXD11, HOXD13), and for two of them (HOXC13 and HOXD13), their potential role in DNA replication has been investigated, as discussed below. In addition, many different HOX proteins have been shown to interact, either directly or not, with the replication inhibitor GEMININ [435], suggesting that this could be a common mechanism by which HOX proteins might regulate DNA replication.

de Stanchina et al [436] aimed to identify novel proteins bound to the *lamin B2* origin by a yeast one hybrid screen. They identified and confirmed by EMSA and CHIP the binding of HOXC10 and HOXC13 on this DNA sequence. HOXC13 colocalizes with BrdU foci in early S phase, indicating that this transcription factor is present at many replication sites [437]. Accordingly, HOXC13 was also detected at the *MCM4* and *TOP1* replication origins by CHIP in T98G cells, but not at the  *$\beta$ -globin* origin [438]. Moreover, the association of HOXC13 with the *lamin B2* origin is cell cycle regulated, since the binding is detected only in G1 and during the G1/S transition, the



period of the cell cycle when pre-RCs are formed. By GST pull-down assays, HOXC13 was shown to interact with CDC6 and ORC1. The interaction with CDC6 and ORC2 was confirmed by immunoprecipitation and by FRET, and seems to be stronger during the G1/S transition. However, disruption of HOXC13 expression with shRNAs does not affect cell cycle progression nor BrdU incorporation. This could be explained by a redundancy of the presence of other HOX proteins on replication origins in the cells studied (T98G and U2OS).

The homeotic proteins HOXA13, HOXD11 and HOXD13 were also detected at five different replication origins, namely *LaminB2*, *MCM4*, *FMRI*, *TOP1* and *c-Myc* [439]. The association of HOXD13 with origins appears preferentially in G1 and is antagonized by the protein GEMININ. HOXD13 directly associates with CDC6 and GEMININ, and these interactions are mutually exclusive. Interestingly, transfection of HOXD13 in NIH3T3 cells stimulates the binding of pre-replication proteins (ORC1, CDC6, MCM7) to replication origins [439]. Overexpression of HOXD13 does not affect the transcriptional level of genes located nearby these origins, while the transcriptional level of replication genes has not been measured. Thus, HOXD13 binding to origins is not linked to its function in transcription. Rather, HOXD13 stimulates the replication of plasmid DNA containing the *laminB2* and *MCM4* replication origins. Additionally, transfection of HOXD13 increases BrdU incorporation in synchronized HEK293 cells, whereas reducing HOXD13 overexpression with shRNAs prevents this effect. Therefore, HOXD13 may directly control DNA synthesis.

## **2.11 Concluding remarks and perspectives**

To date, the exact mechanism by which transcription factors influence DNA replication has not been dissected. Transcription factors may a) help in the loading of replication complexes on DNA; b) modify the chromatin environment near origins; c) regulate the number of origins to be fired; d) facilitate (or prevent) the firing efficiency of specific origins; d) influence the speed of fork progression.

Transcription factors may influence DNA replication by a combination of different mechanisms as well.

### **2.11.1 Are transcription factors important for origin selection/activation?**

Despite the large numbers of origins in mammals, less than 50% appear to be commonly used between different cell types [392]. In addition, in a particular cell type, only a minority of origins appears to be constitutively active whereas a large proportion of origins are either dormant or flexible [358, 392]. It is possible that these flexible origins might be activated in particular conditions, for example during development, differentiation or in stress conditions. Hence, the number of origins to be fired could be inversely proportional to the S phase length. In addition, when stress signals elicit rapid cell expansion, stress-induced transcription factors could influence the number of origins to be fired. In support of this hypothesis, Cadoret et al [361] found that 69 out of 283 (24%) origins were bound by c-JUN, c-FOS or both transcription factors. Inversely, 27,5% of all c-JUN binding sites corresponded to replication origins. Another example is ILF2, a subunit of the heterodimeric NFAT transcription factor, that has been found in a complex with its heterodimerization partner ILF3, as well as MCM and ORC proteins, on the *β-globin* locus [440]. In hematopoietic cells, knockdown of ILF2 perturbs both the transcription of *globin* genes and replication from the *β-globin* locus, without affecting replication at the *Lamin B2* locus. This specific intervention of NFAT proteins could be important for erythroid differentiation or, again, when cells undergo rapid expansion following stress response.

During the process of lineage diversification, it is well established that tissue-specific transcription factors impart cell type specificity to promoter activity. Similarly, origin activity also varies between cell types. It is possible that transcription factors bring specificity in the origin firing process. Of course, the local active chromatin state within replication initiation sites may favor the assembly of both replication and transcription complexes at specific loci. Transcription factors could influence DNA

replication by modifying the local chromatin structure, instead of regulating the firing of origins. However, in the examples proposed in this review, many transcription factors associate directly with replication proteins, arguing in favor of a more direct involvement of some transcription factors in the regulation of DNA replication, rather than mediated through the modification of local chromatin structure. Transcription factors are unlikely to regulate all replication origins: it is possible that DNA replication at constitutive replication origins, in heterochromatin or in gene-poor regions occur in the absence of transcription factor binding. Instead, transcription factors could be involved in fine tuning DNA replication at specific loci.

### **2.11.2 Early versus late origins**

During S phase, replication origins are fired according to a temporal program, some origins firing early, while others fire later. What governs the timing of origin firing remains an open question. The fact that the timing of RB binding to replication origins parallels the timing of origin firing following environmental insult by irradiation [434] is interesting, even though the direct implication of RB in origin firing has not been demonstrated. Whether this coupling is also observed during normal cell cycle, or with other transcription factors, remains to be demonstrated. Many genome-wide studies in *Drosophila* [365, 441], mouse [442-444] and human cells [445-448] point to a correlation between replication timing and gene expression, supporting the hypothesis that actively transcribed genes are often replicated early in S phase. This correlation seems to be a unique feature of higher eukaryotes, not observed in the budding yeast *S. cerevisiae* [449]. Since early origins are found near actively transcribed regions, we could hypothesize that transcription factors may preferentially regulate early replication origins.

### **2.11.3 Oncogenic transcription factors as regulators of DNA replication**

Interestingly, transcription factors that have thus far been associated with replication (c-MYC, E2F, MYB and HOX) are potent oncogenes. With the exception of HOXD13 and c-MYC, a direct involvement in DNA replication remains to be

demonstrated. It is interesting to note that both *HOXD13* [450] and *c-MYC* [451] are misregulated in childhood leukemias as a result of chromosomal translocations. Moreover, seventy per cent of chromosomal translocations in acute leukemias involve transcription factors [452], underscoring their importance as primary oncogenes. Undoubtedly, their role as master regulators of gene expression, including genes in the DNA replication category could contribute to their transforming potentials. Additionally, evidence reviewed here supports a transcription-independent role in DNA replication. This in turn could induce a replicative stress, leading to replication errors and genomic instability, a hallmark of transformed cells [453].

Replicative stress can arise from a premature entry into S phase with too few pre-RCs assembled, with the consequence of incomplete DNA replication during S phase which could lead to DNA strand breaks and chromosomal aberrations [454, 455]. Alternatively, re-replication from the same origin during S phase could also generate genomic instability and replication fork stalling [412, 413]. The control of DNA replication is critical for preservation of genomic integrity and prevents DNA synthesis under suboptimal conditions. For example, a hypomorphic allele of *Mcm4* leads to chromosomal instability and adenocarcinomas in mice [456], as reduced *Mcm2* expression results in cancer, particularly lymphomas [457]. In keeping with this, a perturbation of the timing and/or efficiency of origin firing at specific loci by oncogenic transcription factors could lead to genomic instability and eventually tumorigenesis. For example, *c-Myc* overexpression triggers inappropriate entry into S phase after ionizing irradiation [458]. While this might be due to impaired G1/S checkpoint, another possibility which is not mutually exclusive with the first one is that c-MYC facilitates S phase progression by favouring DNA replication under conditions that would normally be non permissive. In addition, this could provide an alternative interpretation for oncogene collaboration in cell transformation, as exemplified by the synergy between *Myc* and *Ras* oncogenic potential [459]. In primary fibroblasts, a constitutively active RAS induces senescence, a state of prolonged cell cycle arrest [105]. Under these conditions, Herold et al proposed that co-expression of c-MYC could trigger aberrant DNA replication even in unfavourable conditions and overcome RAS-induced growth arrest [417]. This in turn

could lead to genomic instability and cell transformation as discussed above. In the same perspective, other oncogenic transcription factors could transform the cell, in part by forcing cells to replicate their DNA under conditions when they should not [412, 413].

## **2.12 Acknowledgements**

We wish to sincerely apologize to the many colleagues whose important contributions we were not able to cite due to space constraints, and to thank Dr Alain Verreault for critical reading of the manuscript.

### **3. LMO2 regulates DNA replication in hematopoietic cells**

Marie Claude Sincennes<sup>1</sup>, Déborah Ley<sup>1</sup>, André Haman<sup>1</sup>, Francis Migneault<sup>1</sup>, Trang Hoang<sup>1,2</sup>

<sup>1</sup> Institute of Research in Immunology and Cancer (IRIC) – University of Montreal, Montréal, Québec, Canada H3C 3J7

<sup>2</sup> Pharmacology, Biochemistry and Molecular Biology Departments, University of Montréal, Québec, Canada H3C 3J7

This chapter corresponds to a manuscript to be submitted

### **3.1 Contribution des co-auteurs**

MCS : Design expérimental, réalisation des expériences, analyse des données, rédaction du manuscrit

DL : Réalisation des expériences (Figures 3.2 A, B et 3.3) et analyse des données

AH : Réalisation des expériences (contribution aux Figures 3.1C et 3.5)

FM : Réalisation des expériences (Figures 3.1A, 3.2C, 3.7) et analyse des données

TH : Design expérimental, analyse des données, rédaction du manuscrit

## 3.2 Summary

The LMO2 oncogenic transcription factor is often misexpressed in T-cell acute lymphoblastic leukemias (T-ALL), due to chromosomal rearrangements or accidental retroviral insertion in patients from a gene therapy trial. How LMO2 transforms hematopoietic cells remains elusive. Here we show that LMO2 is recruited to a significant number of DNA replication origins in hematopoietic cells by interaction with three essential components of the pre-initiation complex: DNA polymerase  $\delta$ , DNA primase 1 and MCM6. This interaction requires DNA and occurs at the G1-S transition. Accordingly, lowering LMO2 levels decreases the speed of DNA synthesis and delays G1-S progression upon serum stimulation. Conversely, LMO2 overexpression induces T-ALL in mice, and specifically drives cell cycle entry in thymocyte progenitors without affecting bone marrow-derived common lymphoid progenitors or B cells. These results are consistent with a novel role for LMO2 in directly promoting DNA synthesis in hematopoietic cells and, possibly, during T-cell transformation.

## 3.3 Introduction

The gene encoding the oncogenic transcription factor LMO2 (LIM-only 2) is frequently activated and misexpressed in the T lineage as a consequence of recurrent chromosomal rearrangements (reviewed in [460]) or deletion of a negative regulatory element [141] in T cell acute lymphoblastic leukemia (T-ALL). Moreover, five patients in a X-SCID gene therapy trial developed T-ALL associated with *LMO2* gene activation by retroviral insertion [143-145]. These observations underscored the importance of *LMO2* in T cell transformation. Direct evidence for the role of LMO2 as an oncogene came from several mouse models in which LMO2 expression was directed to the thymus [88, 154, 155]. Despite the fact that LMO2 is a potent oncogene activated by multiple mechanisms both in humans and in mice, how LMO2 transforms T cells remains to be determined. Loss of function studies revealed a critical role for LMO2 at the onset of hematopoiesis during development [31, 55]. LMO2 overexpression in transgenic mouse models prevents thymocyte differentiation [88, 155] and increases T cell proliferation [88]. Furthermore, a major



consequence of LMO2 activation in the X-SCID gene therapy was uncontrolled proliferation of T cells in two patients long before leukemia onset [143]. The molecular mechanisms responsible for this high proliferation have not been determined.

LMO2 (LIM only 2) is a small protein containing essentially two LIM domains. These zinc finger domains typically establish multiple protein-protein interactions (reviewed in [185]). In erythroid cells and leukemic T cells, LMO2 is in a complex with the SCL transcription factor, which is also a T cell oncogene (reviewed in [461]). In this complex, LMO2 entertains direct interactions with SCL [316, 317], GATA1/2 [318, 320] and LDB1 [320, 462] to activate or repress the transcription of target genes ([37, 335, 336, 463], reviewed in [461]). Despite increased proliferation, most cell cycle genes were not affected by global analysis of LMO2-expressing thymocytes, except for *Cdkn1a* which was two-fold decreased [114]. *Cdkn1a* is a direct target of the SCL-LMO2 complex which expression can either be activated or repressed, depending on the stoichiometry of activator vs ETO2 co-repressor [62, 329, 464]. The regulation of *Cdkn1a* has important functional consequences in hematopoietic stem cells, in erythroid progenitors and megakaryocytes [62, 329, 464]. However, *Cdkn1a*-deficiency did not affect thymocyte development [465], suggesting the implication of additional events.

LMO2 does not bind directly DNA and thus, may entertain additional protein-protein interactions that could influence other cell processes. Emerging evidence indicates that c-MYC, one of the most studied oncogenic transcription factor, can be part of non-transcriptional complexes involved in DNA replication [416], accelerating S phase progression [419]. In addition, the oncogenic transcription factor HOXD13 also associates with DNA replication origins and regulates their activity [439]. DNA replication is tightly regulated, and defects in this process could ultimately lead to cancer.

In this study, we investigated whether LMO2 could be part of new protein complexes, by performing a proteome-wide interaction screen. We show that LMO2

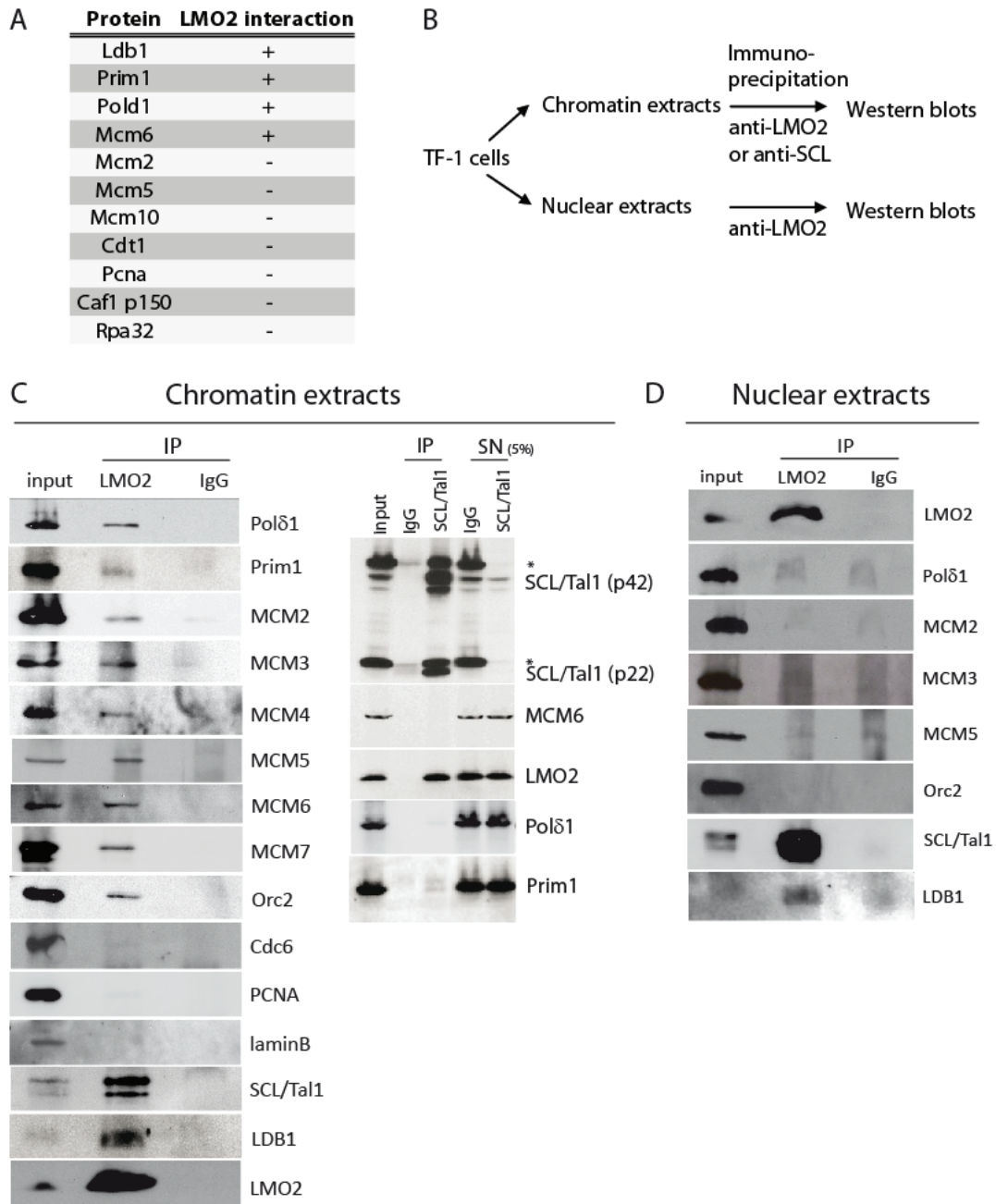
associates with three replication proteins MCM6, DNA primase and DNA polymerase delta, and that LMO2 influences cell cycle progression in hematopoietic cells, supporting a new role for the LMO2 oncogene.

## 3.4 Results

### 3.4.1 LMO2 interacts with DNA replication proteins

Hematopoietic stem cells and progenitors represent a minority within the bone marrow compartment, with a frequency of less than  $10^{-3}$ , and therefore, are not amenable to large scale proteomics approaches. In order to identify LMO2 protein partners in these rare cells, we performed a yeast-two hybrid screen using LMO2 as bait. For this purpose, we purified c-KIT<sup>+</sup> Lin<sup>-</sup> progenitors from mouse bone marrow cells to construct a cDNA library. In addition to known LMO2-interacting proteins such as LDB1, we identified new interactions with three basic components of the replication machinery, namely MCM6, DNA polymerase delta (POLD1) and DNA primase (PRIM1). These interactions were reproducibly confirmed by independent yeast two hybrid assays with full length cDNAs for the three genes and were not observed with negative controls (SV40 large T antigen, p53, Lamin C) (**Suppl. Figure 3.7**). Furthermore, interaction was limited to POLD1, PRIM1 and MCM6 and was not observed with MCM2, MCM5, MCM10, CDT1, PCNA, CAF1 or RPA32 (**Figure 3.1A**). Together, these results indicate that LMO2 specifically interacts with MCM6, POLD1 and PRIM1.

DNA replication is an ordered process and its timing is tightly controlled (reviewed in [357]). Hence, the pre-replication complex is formed on DNA during late G1 and dissociates at later stages. Furthermore, protein-protein interactions in this complex often require the presence of DNA and are observed on chromatin but not in the absence of DNA templates [438, 466, 467]. We therefore assessed whether LMO2 interacts with endogenous replication proteins in undifferentiated hematopoietic progenitors (human CD34<sup>+</sup> TF-1 cells [468]) and whether this requires DNA or not. To this end, we prepared DNA-containing chromatin extracts as well as DNA-free nuclear extracts from TF-1 cells (**Figure 3.1B**). Immunoprecipitation with an anti-



LMO2 antibody efficiently brought down LMO2 in both types of extracts. The anti-LMO2 antibody co-immunoprecipitated SCL in TF-1 nuclear extracts as expected [177] (**Figure 3.1D**), indicating that the SCL-LMO2 interaction does not require DNA. In chromatin extracts, SCL also co-immunoprecipitated with LMO2 (**Figure 3.1C**). There were marked differences, however, between chromatin extracts and nuclear extracts when replication proteins were analysed. In chromatin extracts, POLD1, PRIM1 and MCM6 were reproducibly brought down by the LMO2 antibody but not by control Ig. Furthermore, the non-replicative protein Lamin B was not found in the immune pellet, indicating that the observed associations were specific. In sharp contrast, replication proteins were not found in LMO2 immunoprecipitation of nuclear extracts, even if SCL detection appeared stronger (**Figure 3.1D**). Together, these observations indicate that the association between LMO2 and POLD1, PRIM1 and MCM6 proteins occurs *in vivo* and depends on the presence of DNA.

Origin recognition complex (ORC) proteins bound to DNA recruit CDC6 together with CDT1 followed by MCM2-7. This is known as the pre-replication complex. ATP hydrolysis leads to the release of both CDT1 and CDC6 from replication origins. This permits the completion of MCM2-7 loading and the transition of the pre-replication complex to the pre-initiation complex (pre-IC) (reviewed in [357, 469]). DNA polymerases are then recruited to the origins. To gain insight into the role of LMO2, we assessed the presence of other replication proteins with chromatin-bound LMO2. As shown in **Figure 3.1C**, we detected the presence of MCM2-7 and ORC2 but not CDC6, nor PCNA. The processivity factor PCNA needs the presence of a RNA/DNA primer to assemble, thus the association of polymerases on replication origins precedes the association of PCNA, which is more involved in the elongation step of DNA replication [470, 471]. These observations are consistent with the view that LMO2 associates with the pre-initiation complex following the completion of MCM2-7 loading onto DNA, prior to the elongation step.

In hematopoietic cells, LMO2 is mostly found associated with SCL in transcription complexes that include LDB1, E2A or HEB and GATA proteins ([320], reviewed in [461]). As illustrated in **Figure 3.1C**, immunoprecipitation with an antibody against

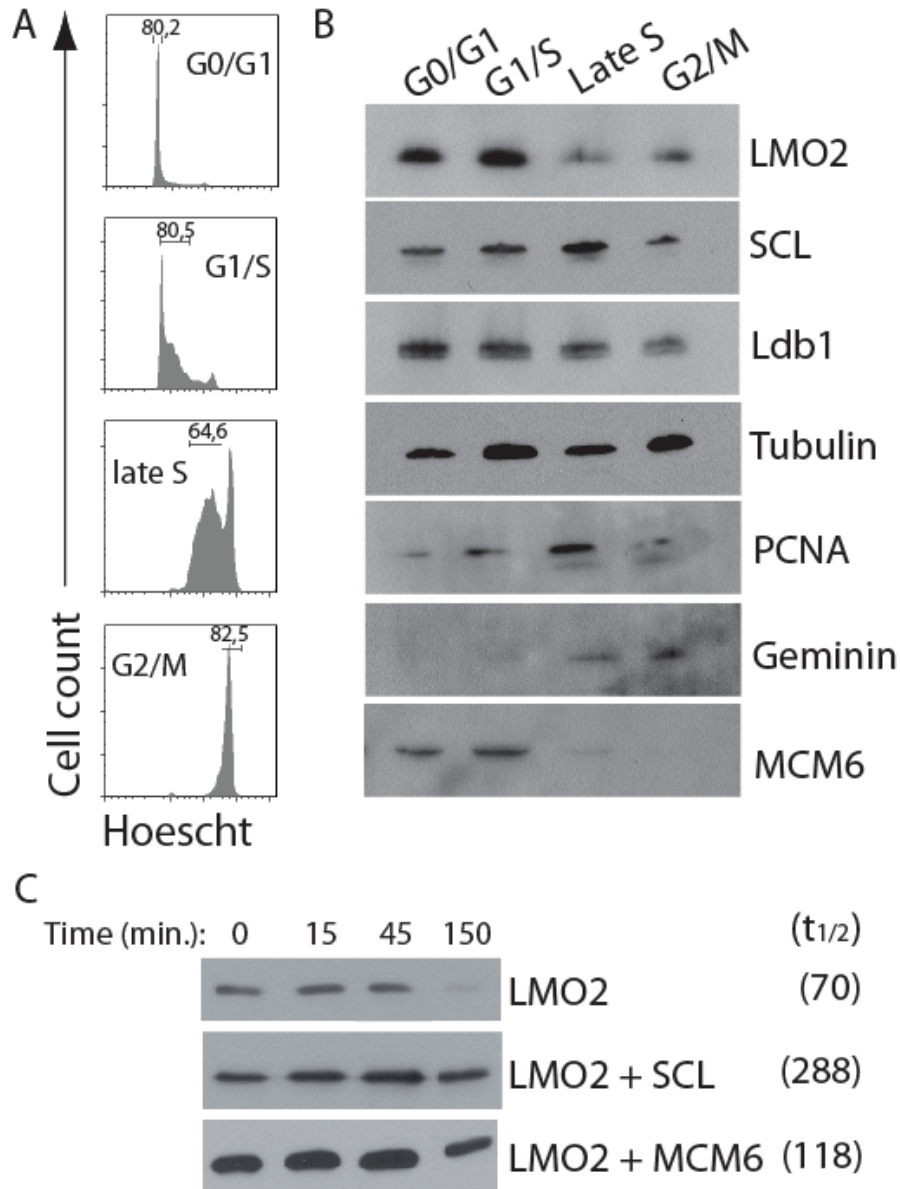
SCL efficiently brought down LMO2. In contrast, MCM6, POLD1 or PRIM1 did not co-immunoprecipitate with SCL, even though these three replication proteins directly interact with LMO2. Almost all of the SCL protein was immunoprecipitated by the anti-SCL antibody, indicating that the absence of replication proteins was not due to ineffective immunoprecipitation. We therefore conclude that replication proteins do not associate with the SCL transcription complex.

### **3.4.2 LMO2 levels vary during the cell cycle**

Expression levels of some replication proteins like geminin [472] fluctuate during the cell cycle. We therefore assessed the expression levels of LMO2 during the cell cycle by immunoblotting of TF-1 cells that were purified into four fractions, G0/G1, G1-S transition, late S and G2/M, according to their fluorescence with the Hoescht dye (**Figure 3.2A**). PCNA protein levels are highest in S whereas geminin expression is higher in G2/M cells, as expected (**Figure 3.2B**). Western blotting indicates that LMO2 is expressed at higher levels during the G0/G1 and G1/S phases of the cell cycle, and is almost absent in late S and G2/M, comparable to the expression of MCM6. In contrast, LDB1 and SCL expression levels remain constant during the cell cycle. Of note, LMO2 is an unstable protein with a short half-life [177] which is known to be stabilized by interaction with SCL [177] and LDB1 [179]. Since LMO2 interacts with MCM6 and exhibits a similar expression pattern during the cell cycle, we performed cycloheximide chase and observed that LMO2 half-life is two-fold increased in the presence of MCM6 (**Figure 3.2C**). Thus, LMO2 is stabilized by MCM6 and both LMO2 as well as MCM6 nuclear levels are regulated during the cell cycle, being reproducibly highest in late G1 or early S phase. Furthermore, this process is likely to be independent of SCL and LDB1.

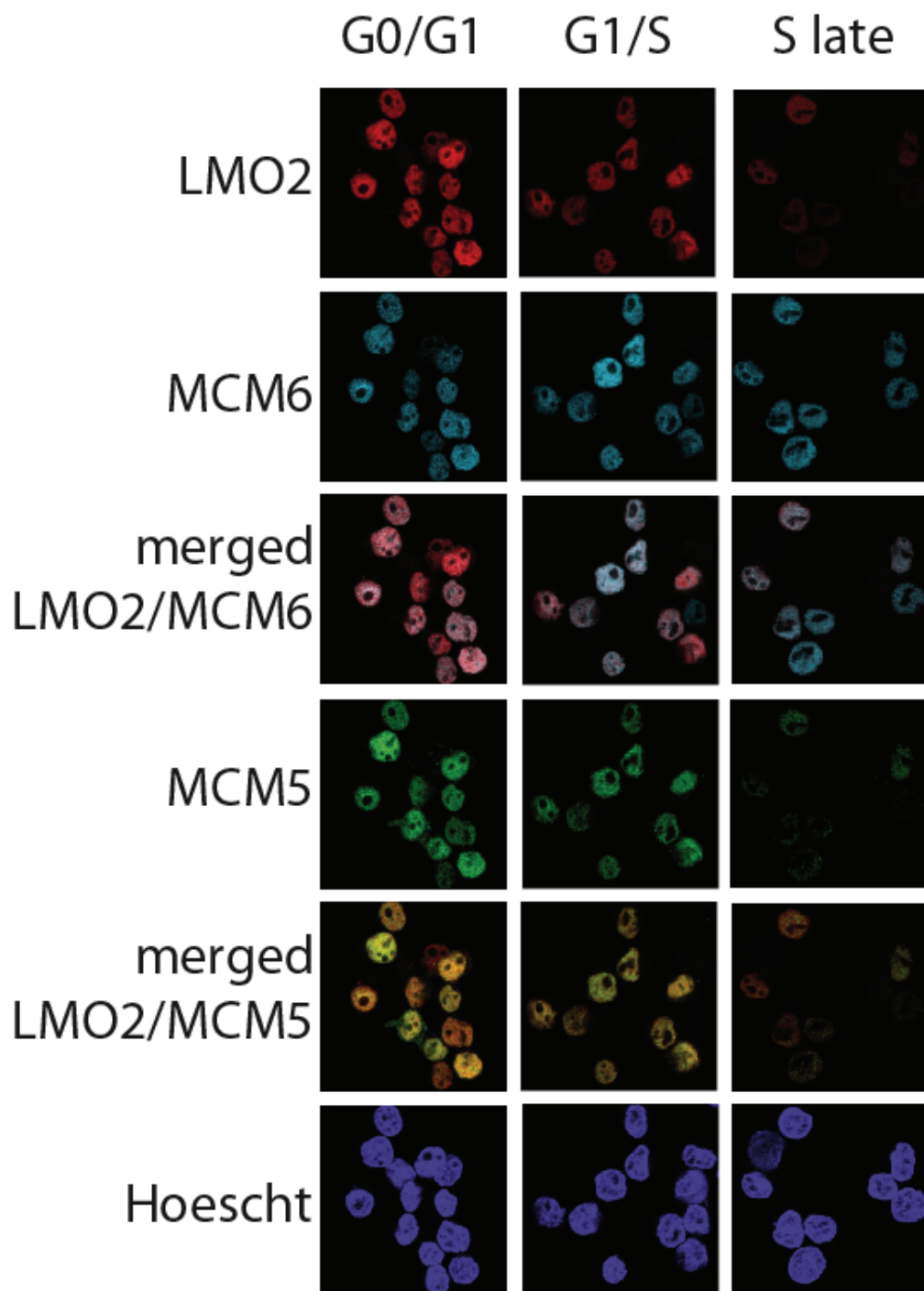
### **3.4.3 LMO2 colocalizes with replication proteins**

We then addressed the question whether LMO2 colocalized with DNA replication proteins in intact cells. Immunofluorescence analysis in TF-1 cells using an anti-LMO2 antibody (**Figure 3.3**) concurs with immunoblotting experiments indicating that LMO2 protein levels are highest in G1 and G1-S. Similarly, the replication



**Figure 3.2 LMO2 protein levels are regulated during the cell cycle.**

A) Purification of TF-1 cells according to their cell cycle status. TF-1 cells were stained with the Hoescht dye and purified by flow cytometry according to their DNA contents and reanalyzed for purity. B) LMO2 expression is higher in G0/G1 and G1/S. Nuclear extracts (40 $\mu$ g per lane) from cell cycle subpopulations purified in A) were analyzed by immunoblotting with the indicated antibodies. Data shown are typical of 4 experiments. C) LMO2 half-life is increased in presence of its interacting partners. Transfections were carried out in 293T cells with expression vectors encoding LMO2 alone (top), or in combination with SCL (middle) or MCM6 (bottom). Cells were treated with cycloheximide (100 $\mu$ g/ml) for the indicated times, and nuclear extracts were prepared. Samples were analyzed by IB with an anti-LMO2 antibody.



**Figure 3.3 LMO2 colocalizes with MCM5 and MCM6 in the nuclei of TF-1 cells, during G0/G1 and the G1/S transition**

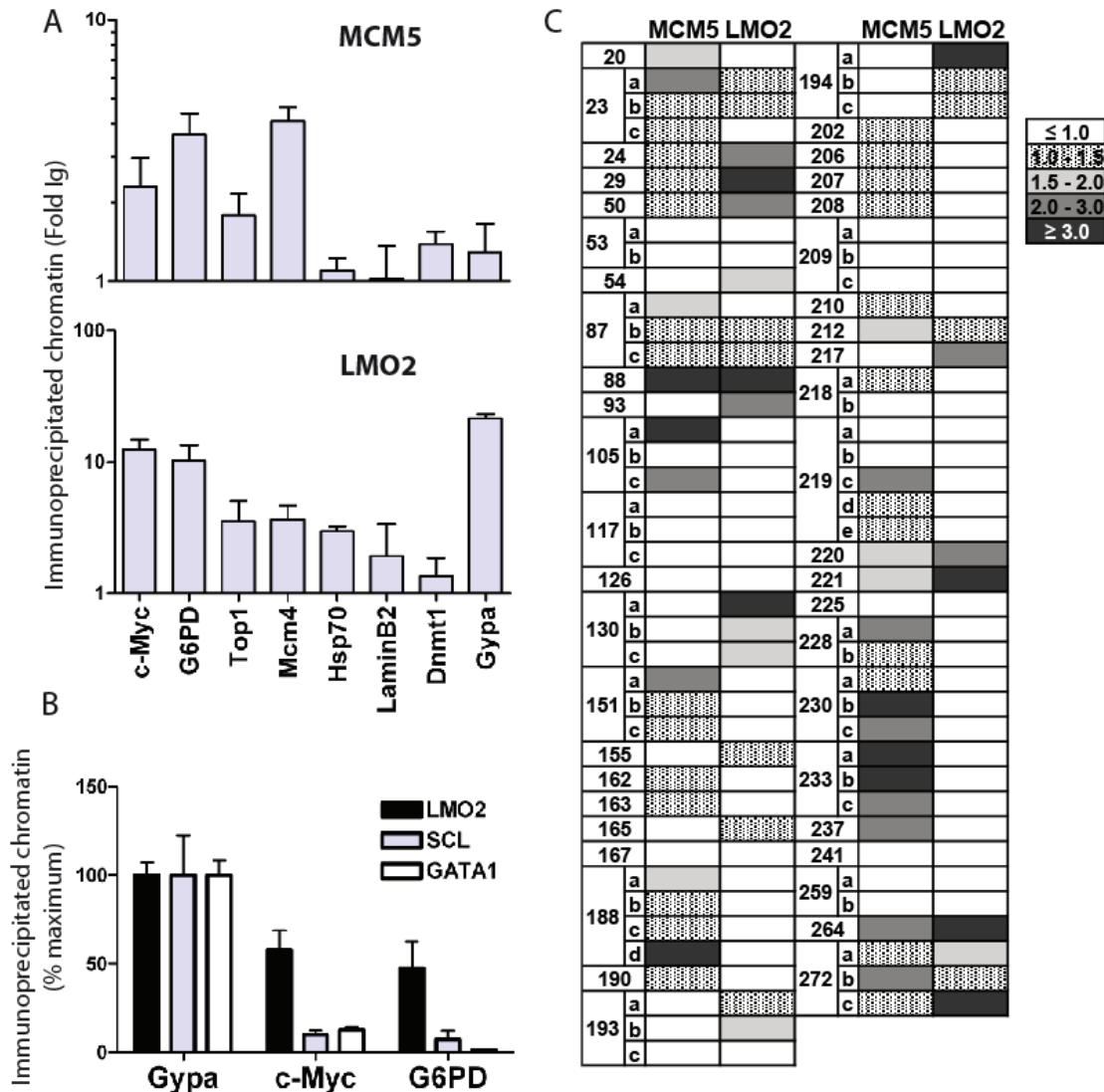
TF-1 cells purified according to their cell cycle status were triple-stained with antibodies directed against LMO2 (red), MCM6 (cyan) and MCM5 (green). Hoescht-stained nuclei are in blue. Data are typical of at least 3 experiments.

proteins MCM5 and MCM6 levels were highest in G1-S, although these proteins persist in late S phase. Finally, this approach demonstrated a significant degree of colocalization between LMO2, MCM5 and MCM6 during G0/G1 and the G1-S transition. These results concur with co-immunoprecipitation data and suggest that LMO2 associates with the pre-IC complex.

#### 3.4.4 LMO2 binds to DNA replication origins

Since LMO2 associates with MCM6, POLD1 and PRIM1 on DNA, we assessed whether LMO2 associates with DNA replication origins. We used chromatin immunoprecipitation (ChIP) of TF-1 cells to determine the presence of LMO2 on well characterized replication origins as well as newly identified origins by genome-wide studies. We first studied seven well characterized human DNA replication origins: *LAMINB2* [370], *c-MYC* [372], *MCM4* [378], *TOP1* [377], *G6PD* [376], *HSP70* [373] and *DNMT1* [386] (**Figure 3.4A**). We performed ChIP using antibodies against LMO2 and MCM5. We found that four out of seven replication origins (*c-MYC*, *G6PD*, *TOP1*, *MCM4*) are bound by MCM5 in TF-1 cells (**Figure 3.4A**). LMO2 was also detected at these four replication origins and was recruited to two of these (*c-MYC*, *G6PD*) to high levels. LMO2 is known to be recruited to Ebox/GATA sites on DNA by SCL and GATA1 [320]. However, LMO2 detection by ChIP did not correlate with the presence of E boxes and GATA sites (**Suppl. Figure 3.8**). As a control for the ChIP assay, we confirmed that LMO2, SCL and GATA1 but not MCM5 occupy the *GLYCOPHORIN A (GYPA)* promoter, a well defined LMO2 transcriptional target [79] (**Figure 3.4A, B**), indicating that the procedure is discriminatory. Furthermore, GATA1 and SCL are practically absent from the two origins that show the highest enrichment with LMO2 immunoprecipitation, i.e. *c-MYC* and *G6PD* (**Figure 3.4B**). The simultaneous presence of MCM5 and LMO2 in absence of SCL and GATA1 at the *c-MYC* and *G6PD* origins, as well as the lack of correlation with typical LMO2 consensus sequence, suggest that LMO2 is recruited to DNA replication origins by MCM proteins but not SCL nor GATA1.





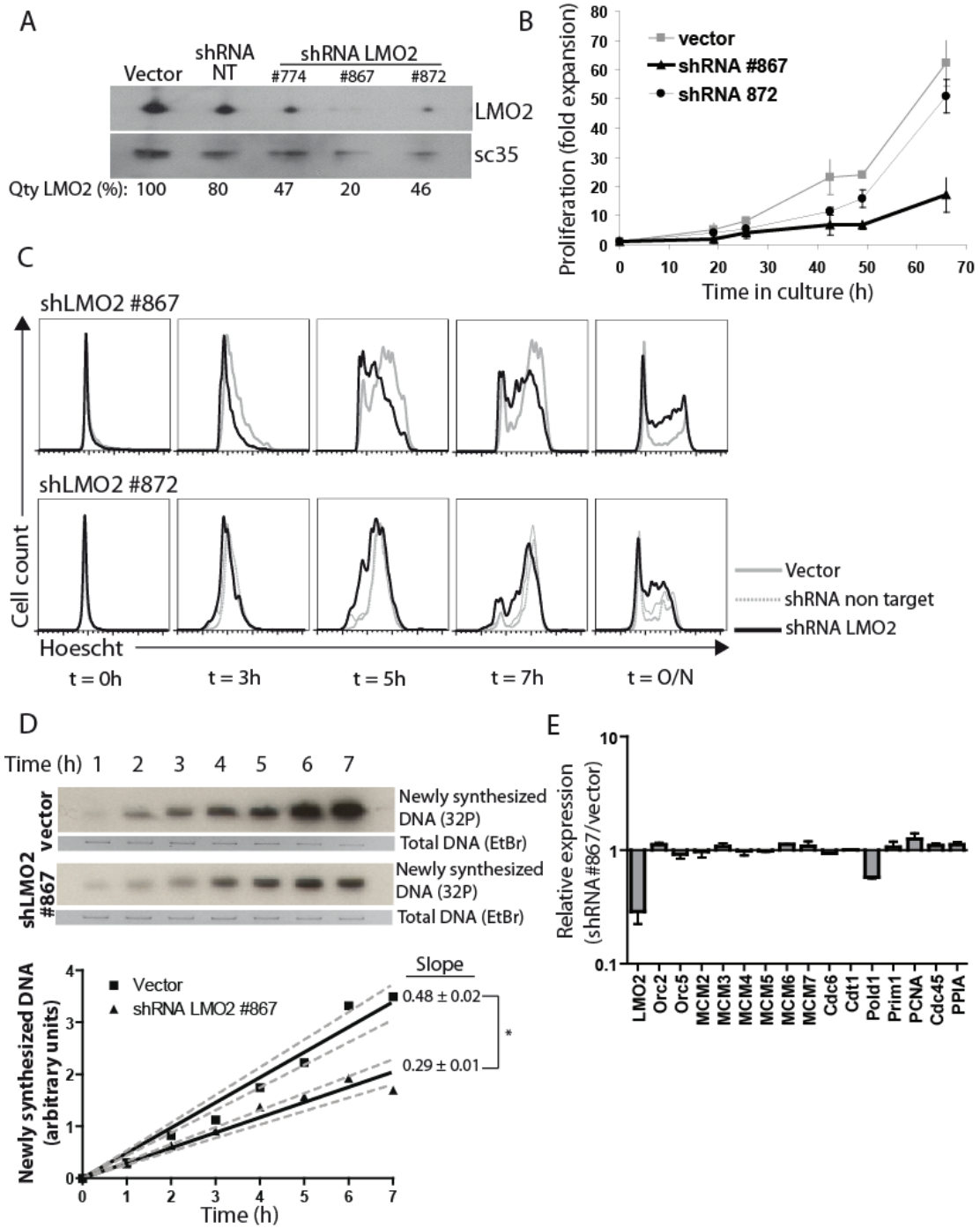
**Figure 3.4 LMO2 and MCM5 occupy DNA replication origins**

A, B) Chromatin immunoprecipitation of TF-1 cells with antibodies against MCM5, LMO2, SCL or GATA1. Immunoprecipitated chromatin levels are indicated as ratio of enrichment over control isotype Ig, after normalization with a negative control DNA sequence (*c-Kit* -20kb upstream sequence) B) Immunoprecipitated chromatin was normalized to maximum occupancy levels observed with the *Glycophorin A* promoter (Gypa). C) Heat map representing the levels of immunoprecipitated chromatin, after normalization as in A). Initiation zones are numbered according to Cadoret et al[361]. When more than one primer pair was used for a defined region, primer pairs are designated by letters. Data represent the mean of two independent experiments performed in duplicate. The exact mean and SD are shown in **Suppl. Figure 3.9**.

The presence of LMO2 at well characterized origins led to the question whether LMO2 might also be associated with new replication initiation zones identified by unbiased methods. Recently, two independent studies in human cells [361, 394] purified origin-centered nascent DNA strands using lambda exonuclease digestion, and Karnani et al [394] also included anti-bromodeoxyuridine (BrdU) immunoprecipitation. Nascent DNA strands were hybridized on the Human ENCODE ChIP-on-chip microarray. Importantly, 45 regions identified by Cadoret et al showed an overlap within a 1kb distance with 47 origins identified by Karnani et al. We therefore focused on this subset since these DNA regions are more likely to contain highly efficient origins in human cells. We designed 80 PCR primer pairs covering a) the overlapping region, when the two origins directly overlapped; or b) each region in addition to the “gap” region(s), when the origins were separated by a short DNA segment. After validation, these primer pairs were used to detect the presence of these origins in chromatin extracts from TF-1 cells immunoprecipitated with anti-MCM5 and anti-LMO2 antibodies. As shown in **Figure 3.4C** and **Suppl. Figure 3.9**, MCM5 is detected to 29 replication initiation zones (64%), a proportion that reflects the results obtained with well-characterized origins (57%). LMO2 is detected on 19 initiation zones (42%), and moreover, 57% of LMO2-bound DNA regions were also positive for MCM5-binding. These results indicate that LMO2 is recruited to a significant proportion of initiation zones in TF-1 cells.

### **3.4.5 S phase progression depends on LMO2 levels in hematopoietic cells**

We next assessed the possibility that LMO2 might influence cell cycle progression, in particular the G1-S transition or S phase progression. To this end, we disrupted LMO2 levels in mouse erythroleukemia (MEL) cells by RNA interference. We observed that LMO2-directed shRNAs #867 and 872 induced a two- and four-fold decrease in LMO2 protein levels, respectively (**Figure 3.5A**). Interestingly, the growth of MEL cells was significantly reduced by the use of *LMO2*-directed shRNAs (**Figure 3.5B**). Eventually, cells that grew after several passages in culture had levels of LMO2 that were comparable to control cells, despite the continuous presence of



**Figure 3.5 Disruption of LMO2 levels impairs S phase progression in murine erythroleukemia (MEL) cells**

A) Nuclear extracts of MEL cells infected with the pLKO.1 empty vector, a non-target shRNA or different LMO2-directed shRNAs were analysed by immunoblotting as indicated. B) The growth of MEL cells is decreased by *Lmo2*-directed RNA interference. MEL cells were seeded at  $5 \times 10^4$ /ml in culture and counted at the indicated time points. Input levels were set at 1. Data shown are the mean  $\pm$  SD of two independent experiments with two different batches of shRNA expressing MEL cells. C) G0/G1 MEL cells were purified and released in culture for different times, then stained with the Hoescht dye. One representative experiment is shown (n=3). D) Top: G0/G1 MEL cells were purified and released in culture for different times in the presence of  $\alpha$ - $^{32}\text{P}$ -dCTP. Total DNA was resolved on agarose gel and visualized by autoradiography. One representative experiment is shown (n=2). Bottom: Quantification of  $^{32}\text{P}$  incorporation shown on top. \*  $p < 0.0001$ . E) Decreased LMO2 expression does not globally affect the transcriptional level of replication genes. Total RNA was extracted from infected MEL cells and subjected to RT-PCR. RNA levels were normalized over the *Hprt* control gene and are indicated as a ratio of the expression in cells infected with a LMO2-directed shRNA compared to the expression in cells infected with the empty vector. *Peptidylprolyl isomerase A (Ppia)* expression is used as additional control. Data represent the mean of two independent experiments performed in duplicate.

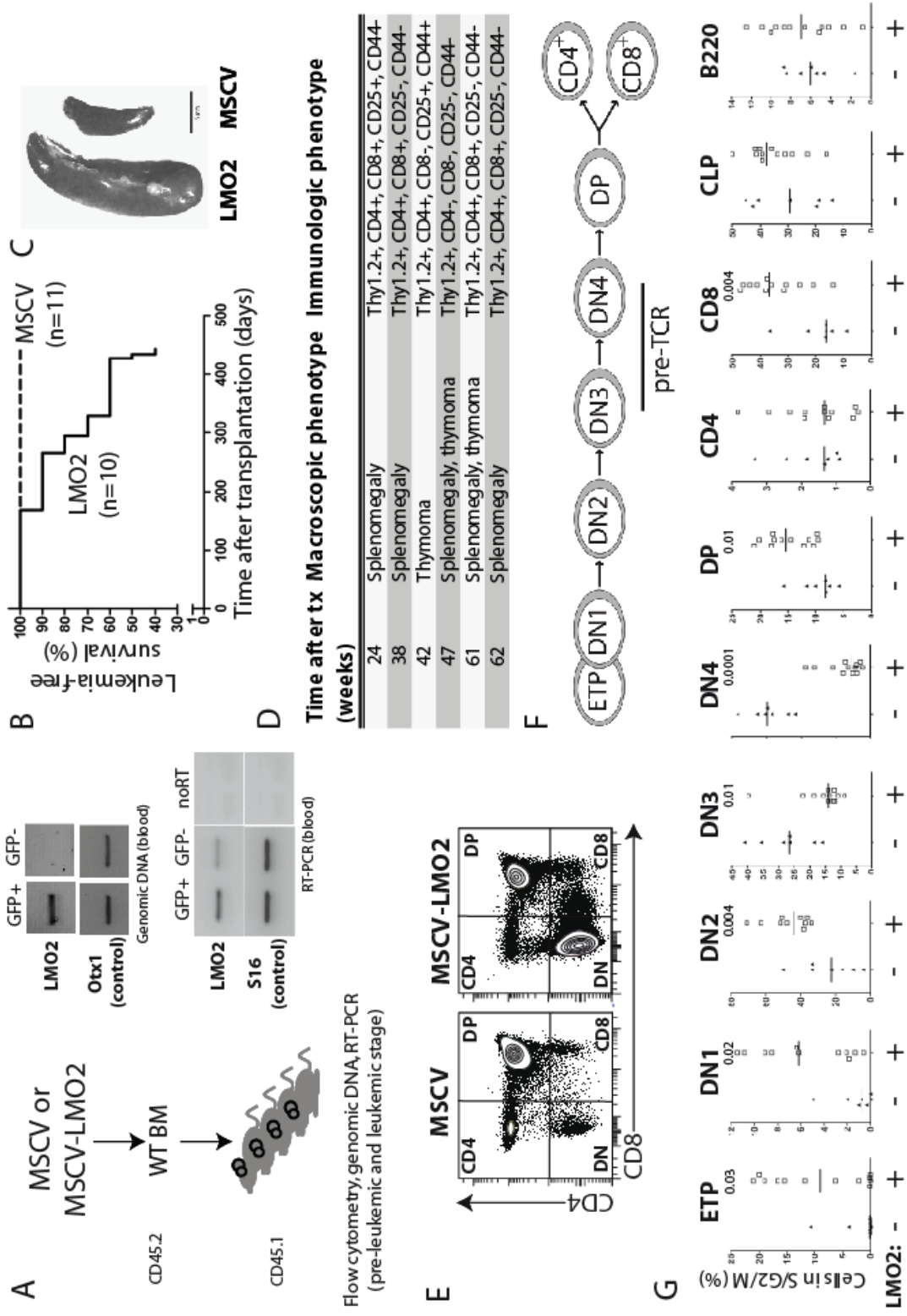
puromycin. Thus, there is selective pressure against low levels of LMO2. Therefore, shRNA delivery was reconducted *de novo* for each subsequent experiment, and LMO2 levels monitored. To evaluate the possibility that the growth defect caused by low LMO2 may be due to a cell cycle defect, G0/G1 cells purified by flow cytometry (**Suppl. Figure 3.10**) [473] were released in serum containing cultures. DNA contents were assessed by Hoescht staining and flow cytometry analysis to evaluate the kinetics of cell cycle progression (**Figure 3.5C**). After release, we observed that cells expressing LMO2 shRNAs (sh*LMO2*) initiated DNA synthesis with a delay compared to control cells expressing the wild-type vector, starting at 3h. Five hours after release, cells expressing the empty vector were mostly in late S and G2 phases, whereas cells with four-fold reduced LMO2 levels (sh867) were still in G1 and early S. After 16h, control cells have resumed a normal cell cycle profile whereas a large proportion of cells expressing sh*LMO2* was still in S phase. This delay was also observed with a different LMO2-directed shRNA (sh872), albeit with a more modest delay (**Figure 3.5C**, compare upper and lower panels), corresponding to a more modest decrease in LMO2 protein levels (**Figure 3.5A**). In contrast, a non-target shRNA did not affect cell cycle progression (**Figure 3.5C**, grey line, lower panel). Moreover, this cell cycle defect was not due to increased apoptosis in cells expressing the sh*LMO2* (**Suppl. Figure 3.11**). We therefore conclude that LMO2 levels control S phase progression in hematopoietic cells.

Since LMO2 associates with POLD1, PRIM1 and MCM6 on DNA and controls early S phase progression, we determined whether LMO2 affects *de novo* DNA synthesis. To this end, we assessed the kinetics of <sup>32</sup>P orthophosphate incorporation into synchronized MEL cells (**Figure 3.5D**, top), allowing for a quantification of the relative speed of DNA biosynthesis (**Figure 3.5D**, bottom). We observed that the speed of DNA synthesis is almost two-fold lower in cells in which LMO2 levels were decreased compared to control. Furthermore, the effect of LMO2 on DNA synthesis and the cell cycle was not due to decreased expression levels of genes coding for components of the DNA replication apparatus, as assessed by RT-PCR (**Figure 3.5E**). These results are in agreement with the absence of genes involved in DNA replication in the various genome-wide analysis of SCL-LMO2 target genes by

chromatin immunoprecipitation [327, 330, 339, 342]. In summary, our data indicate that LMO2 levels control the speed of DNA biosynthesis and S phase progression with an impact on cell growth.

### **3.4.6 LMO2 overexpression increases the percentage of thymocyte progenitors in S phase**

Ectopic LMO2 expression in the thymus of transgenic mice induces T-ALL with low penetrance [88, 154]. Here we used a different strategy to deliver LMO2 in hematopoietic stem cells via retroviral infection, as in the X-SCID gene therapy trial [143-145]. The MSCV vector used here induced a 2-fold increase in LMO2 levels by RT-PCR in transduced cells (**Figure 3.6A**). When transplanted into irradiated recipient mice, LMO2-expressing cells induced T-ALL in 60% of transplanted mice with a latency of 170 days to 450 days (**Figure 3.6B**), comparable to previous studies [89]. All mice developed T-ALL associated with splenomegaly and/or thymoma (**Figure 3.6C, D** and **Suppl. Figure 3.12**) whereas control mice remained healthy during this time frame. To assess the cellular effects of LMO2, the cell cycle status of different hematopoietic populations was analyzed during the preleukemic phase (**Figure 3.6E-G**). Interestingly, LMO2 overexpression did not lead to a global increase of the percentage of cells in S phase. Instead, the effect of LMO2 on the cell cycle status is cell type-dependent. In most bone marrow progenitors (data not shown) as well as in common lymphoid progenitors (CLP) and B cell subpopulations, LMO2 overexpression lead to a modest, non-significant shift in the proportion of cycling cells (S/G2/M) (**Figure 3.6G**). In contrast, in early thymic progenitors (ETP) and in the most immature DN1 as well as DN2 thymocyte subpopulations, LMO2 strongly increased the percentage of cycling cells. However, in the DN3/4 subpopulations, the percentage of cycling cells was reduced, as reported [37]. Indeed, the situation within these subpopulations is more complex, since LMO decreases preTCR expression [37, 114] and therefore pre-TCR-dependent cell proliferation and differentiation. Nonetheless, in cells that have progressed beyond this pre-TCR-dependent stages, i.e. DP and CD8<sup>low/+</sup> thymocytes, LMO2 overexpression was associated with an increased proportion of cycling cells. These results indicate that



**Figure 3.6 LMO2 overexpression specifically affects the cell cycle status of early thymic progenitors *in vivo***

A) Diagram showing the experimental strategy. Upper right panel: The presence of the transgene is detected by PCR in GFP<sup>+</sup> cells retrovirally-expressing LMO2. Lower right panel: Blood cells infected with MSCV-LMO2 show a 2-fold overexpression of LMO2 compared to GFP<sup>-</sup> non infected cells, as detected by RT-PCR. B) Kaplan-Meier curves showing the time of leukemia onset in mice transplanted with MSCV-LMO2 or MSCV transduced cells. C) Typical spleen from a leukemic mouse transplanted with LMO2-overexpressing cells, compared to control. D) Macroscopic and immunologic phenotype of LMO2-induced leukemias. E) Representative analysis of thymocyte subsets by flow cytometry. Note that LMO2 expression induces an expansion of the DN and the CD8<sup>low</sup> subsets. F) Thymocyte differentiation. The earliest thymocyte progenitor (ETP) settles in the thymus and progress through several CD4<sup>-</sup>CD8<sup>-</sup> double negative (DN) stages of development. At the DN3 and DN4 stages, thymocytes expressing the pre-TCR undergo extensive proliferation and progress to the CD4<sup>+</sup>CD8<sup>+</sup> double positive (DP) stages before becoming mature CD4 or CD8 lymphocytes that are released into the periphery. In mice, the immature CD8<sup>low</sup> population is a transitional stage between DN and DP cells. G) All thymocyte subsets as well as bone marrow common lymphoid progenitors (CLP) and B lymphocytes (B220) were analyzed for their cell cycle status by Hoescht staining. P values are indicated when the difference is significant between LMO2 and the vector alone (Student's t test).



the effect of LMO2 on the cell cycle is cell type specific. This activity is manifest in the earliest thymic progenitors *in vivo*, but not in bone marrow-derived CLP and B cells, even though retroviral gene delivery directs the expression of LMO2 in all hematopoietic cells (**Figure 3.6A**). This correlates with the T lineage tropism of LMO2 cell transformation.

## **3.5 Discussion**

In the present study, we provide evidence that the LMO2 oncogene interacts with MCM6, POLD1 and PRIM1, that LMO2 binds DNA replication origins and regulates G1-S transition.

### **3.5.1 A novel role for LMO2 in S phase regulation**

Our observations indicate that LMO2 specifically interacts with three basic components of the replication machinery and that this association occurs on DNA. Our observations are in agreement with previous reports describing the necessity of DNA templates for protein-protein interactions with the DNA replication machinery. For example, CDC6 is recruited to the MCM complex assembled on DNA [466]. Furthermore, the interaction between CYCLIN D1 and MCM7 [467], and between HOXC13 and ORC2 [438], occur only in chromatin extracts, and not in chromatin-depleted nuclear extracts. This contrasts with the MCM hexamers and ORC proteins that can assemble in the absence of DNA. Together, these observations point to the importance of the DNA template for the recruitment of protein partners by the MCM and ORC complexes.

CDC6 and CDT1 are present in the pre-replication complex but not the pre-IC (reviewed in [357]), while POLD1 and PRIM1 are recruited to the latter but not the former. In to the present study, we found MCM6, POLD1 and PRIM1 with LMO2 but not CDC6, suggesting that LMO2 associates to replication complexes when CDC6 and CDT1 already left replication origins. Furthermore, the direct interaction with DNA polymerase delta and DNA primase suggests that LMO2 is involved in the pre-initiation complex, comprising all the proteins needed for DNA replication.

In TF-1 cells, 70% of the LMO2 pool is associated with the SCL transcription factor [177]. The present study uncovers a novel role for LMO2 in DNA replication and several lines of evidence indicate that this role is independent of LMO2 association with SCL and GATA1. First, SCL does not associate with replication proteins in chromatin extracts (**Figure 3.1C**) or in two-hybrid assays (not shown). Second, LMO2 is found on DNA replication origins but not GATA1 nor SCL. Furthermore, SCL protein levels are not cell cycle-regulated like LMO2. Third, the transcription level of replication genes is not affected by decreased LMO2 levels. Several genome-wide studies support these observations. In global analysis of the transcriptome of primary thymocytes overexpressing LMO2, *Cdkn1a* was the only cell cycle gene that was affected and its levels were two-fold decreased [114]. However, the absence of *Cdkn1a* did not affect thymocyte numbers [465], nor the cell cycle of the various thymocyte subsets (Herblot and Hoang, *unpublished observations*). Moreover, genome-wide studies for SCL-LMO2 target genes in primary erythroid cells [327], multipotent cells [339] or Jurkat and Molt4 lymphoid T cell lines [330, 342] revealed that very few if any cell cycle genes are occupied by SCL or LMO2 in situ.

The SCL/LMO2 interaction is the critical event in the assembly of a transcription complex on DNA [79, 177], while GATA family members or E proteins are interchangeable. In contrast, LMO2 is not essential for the assembly of the pre-IC on DNA. Rather, LMO2 has a regulatory function in this case as the recruitment of LMO2 to the pre-IC controls the speed of DNA synthesis in hematopoietic cells.

### **3.5.2 LMO2 as a T-lineage oncogene**

LMO2 is expressed in hematopoietic stem cells and progenitors and therefore, can potentially interact with DNA replication proteins identified here in all cell types in which it is expressed. Indeed, RNA interference studies indicate a dose-dependent effect of LMO2 on cell cycle progression. Surprisingly, although hematopoietic cells are sensitive to decreased LMO2 levels, increasing LMO2 by overexpression studies using a retroviral vector that drives high expression in all hematopoietic cells [62, 474] did not affect most hematopoietic cells except thymic progenitors in which

LMO2 specifically influences the S phase. These observations suggest the importance of the thymic environment for this new function of LMO2. Coincidentally, LMO2 is a potent T-cell oncogene in human and in several mouse models (reviewed in [460]). Without excluding the possibility that increased cell survival could contribute to clonal expansion, we propose that the increased proliferation of T lymphocytes in the two T-ALL patients observed during the pre-leukemic phase [143] may be due to the effect of LMO2 on DNA synthesis in the T lineage.

DNA replication is a highly regulated process in all cells, and depends on the concerted action of multiple regulators. The susceptibility of T-lineage progenitors to the action of LMO2 shown here suggests that DNA replication in this lineage may be more prone to perturbations by oncogenic transcription factors such as LMO2. A misregulated S phase could lead to replication errors and ultimately cause genetic lesions facilitating the emergence of leukemia initiating cells. In support of this hypothesis, we and others have previously shown that *Notch1* gain of function mutations spontaneously occur in 60 to 80% of T-ALL arising in SCL-LMO1 transgenic mice, thus conferring leukemia-initiating potential to SCL-LMO transformed thymocytes [122, 124]. Thus, unlike *TLX1*-induced acute leukemias that typically develop with marked aneuploidy in transgenic mice [475], T-ALL from SCL-LMO1 or SCL-LMO2 transgenic mice do not exhibit gross chromosomal anomalies. Rather, mutations in the *Notch1* gene are small frameshift mutations that likely occur during DNA replication [122, 124]. We propose that this type of mutations could be a consequence of the functional interaction observed here between LMO2 and the DNA replication machinery.

### **3.5.3 Cross-talk between transcription and DNA replication**

Emerging evidence tends to favour a convergence between different DNA-dependent processes. The coupling between DNA replication and DNA repair, for example, is well documented [355]. On the other hand, the link between transcription and other DNA-dependent processes remains elusive. There is extensive evidence for transcription factors that can regulate the expression of genes involved in cell cycle

and DNA repair, but only a few examples document the direct, non-transcriptional implication of transcription factors in DNA replication. As discussed above, previous work demonstrated that HOXD13 and c-MYC, two other oncogenic transcription factors, are found at DNA replication origins and have been shown to directly regulate DNA replication [416, 439]. Conversely, DNA replication proteins could also be involved in transcription. In 293T cells, MCM2 and MCM5 have been shown to be necessary for POLII-mediated transcription [476]. The same group previously demonstrated a role for MCM5 in STAT1-mediated transcription [477]. Our observations argue in favour of a role for LMO2 in regulating DNA synthesis in primary hematopoietic cells, as discussed above.

Assembly of the replication apparatus occurs in two steps that are tightly controlled: the pre-replication complex is first assembled on DNA with inactive MCM2-7, followed by activation of MCM2-7 and recruitment of DNA polymerases into an active pre-IC. Thus, the three oncogenic proteins HOXD13, c-MYC and LMO2 appear to regulate DNA replication at different steps. The pre-replicative proteins CDC6 and/or CDT1 co-immunoprecipitate with c-MYC [416] and HOXD13 [439] but not with LMO2, suggesting that c-MYC and HOXD13 associate with the pre-RC and may have a role in the early steps of DNA replication. In contrast, LMO2 associates with the pre-initiation complex and this interaction takes place mostly at the G1/S transition. This difference is also reflected in the mechanism by which the oncogenes influence DNA replication. c-MYC regulates origin selection without affecting elongation speed, whereas LMO2 influences the speed of DNA synthesis and the kinetics of S phase progression. Finally, the association of c-MYC and HOXD13 with replication proteins could be observed in classical nuclear extracts, while the interaction of LMO2 with replication proteins is detected in chromatin extracts only, underscoring the importance of a DNA template in complex assembly as previously shown for other interactions involving replication proteins [438, 466, 467].

It is interesting to note that *LMO2* [460], *HOXD13* [450] and *c-MYC* [451] oncogenes are all misregulated in childhood leukemias as a result of chromosomal

translocations. Although leukemic cells exhibit multiple mutations and genetic anomalies, seventy per cent of chromosomal translocations in acute leukemias involve transcription factors [452], underscoring their importance as primary oncogenes. This could be due to their role as master transcription factors but also their implication in DNA replication as shown here and elsewhere [416, 439]. Transgenic mouse models of LMO2 revealed a significant differentiation blockade in the thymus during the preleukemic phase [88, 155]. While this could contribute to the oncogenic effect of LMO2, a major observation in patients that underwent X-SCID gene therapy was an uncontrolled T cell proliferation several months before the onset of leukemia [143]. This hyperproliferation is in line with the effect of LMO2 on DNA replication observed here *in vivo* during the pre-leukemic phase in mice. Combined, these observations suggest that the regulation of DNA replication by oncogenic transcription factors could be an important mechanism of cell transformation, in addition to their well documented transcriptional activities.

### 3.6 Experimental procedures

#### Yeast two hybrid

*LMO2* full-length cDNA was subcloned in the pGBKT7 vector and transformed in the Y187 yeast strain. The cDNA library from c-KIT<sup>+</sup> Lin<sup>-</sup> murine bone marrow cells was constructed in the pGADT7 vector using Matchmaker Library Construction and Screening Kit (BD Biosciences) and transformed in the AH109 yeast strain. The screen (selection of Ade<sup>+</sup>His<sup>+</sup>Leu<sup>+</sup>Trp<sup>+</sup> colonies) was performed by yeast mating. Positive clones were isolated and sequenced. To confirm the interactions and check for specificity, full length cDNAs for *Pold1*, *Prim1*, *Mcm2-5-6-10*, *Cdt1*, *Pcna* and *Caf1* were subcloned into pGADT7 and transformed in AH109 cells. After mating with Y187 cells expressing pGBKT7-LMO2 or controls, diploids were selected on media -Leu/-Trp, whereas interacting clones were Ade<sup>+</sup>His<sup>+</sup>Leu<sup>+</sup>Trp<sup>+</sup>X- $\alpha$ -Gal<sup>+</sup>.

#### Cell lines, shRNAs and mouse bone marrow transplantation

MEL cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) and passed every 2 days at the concentration of  $5 \times 10^4$  cells/ml. TF-1 cells were cultured

as described previously [478]. For transplantation experiments, bone marrow cells from 5-FU treated C57BL/6 mice were co-cultured for 48h on GP+E packaging cells stably expressing the viruses encoding LMO2 or control retroviruses (MSCV-GFP). Sublethally irradiated Pep3B mice were injected *i.v.* with  $2 \times 10^5$  transduced cells. Mice were sacrificed 7 month post-transplantation for flow cytometry analysis. For RNA interference experiments, VSV-G producing cells were transfected with plasmids encoding shRNAs against LMO2, with a non-target shRNA or with the empty pLKO.1 vector (Sigma). MEL cells were incubated with VSV-G supernatant for 48h, then selected with puromycin.

#### Flow cytometry and cell cycle analysis

For DNA content staining, TF-1 cells were incubated with the Hoescht dye (10 $\mu$ g/ml) in DMEM (10mM Hepes pH 7,4) supplemented with 2% FCS for 30 min at 37°C, then washed with PBS and sorted on FACS Aria (BD Biosciences) according to their cell cycle status. For synchronization experiments, MEL cells were purified according to cell size (FSC/SSC profile) since small cells are in G0/G1 (**Supplemental Figure 3.10**). Cells were seeded in 6-well plates at  $10^5$  per well and incubated at 37°C for the indicated times. Cells were then harvested, fixed and permeabilized with BD Cytofix/Cytoperm™ Fixation/Permeabilization Kit (BD Biosciences), then stained with 10 $\mu$ g/ml Hoescht for 10 min and washed with PBS. Post-transplantation analysis of bone marrow and thymus populations were described previously [37, 62, 122], using the following antibodies from BD Pharmingen: B220 (RA3-6B2), CD3 $\epsilon$  (145-2C11), CD4 (RM4-5), CD11b (M1/70), CD11c (HL3), CD25 (PC61), CD34 (RAM34), IgM (R6-60.2), NK1.1 (PK136), Sca1 (E13-161.7), TCR $\gamma\delta$  (GL3); from eBioscience CD8a (Ly-2), CD16/32 (93), CD44 (IM7), CD45.2 (104), CD48 (HM48-1), CD127 (A7R34), CD135 (A2F10), c-Kit (2B8), FC $\epsilon$ R1 (MAR-1), Gr-1 (RB6-8C5), TCR $\beta$  (H57-597), Ter-119, Thy1.2 (30-H12); and from BioLegend CD150 (TC15-12F12.2). After staining, cells were fixed, permeabilized and stained with Hoescht as described for MEL cells. FACS analysis was performed on an LSRII cytometer.

### <sup>32</sup>P orthophosphate labelling and quantification of *de novo* DNA synthesis

MEL cells in G0/G1 were obtained by flow cytometry purification based on cell size (**Supplemental Figure 3.10**), resuspended in DMEM (10mM Hepes pH 7,4) supplemented with 10% FCS at  $8 \times 10^5$ /ml and incubated with  $\alpha^{32}\text{P}$ -dCTP (100 $\mu\text{Ci}/\text{ml}$ , Perkin-Elmer). After the labelling period, cells were centrifuged, washed with PBS and lysed in DNA extraction buffer (80 mM Tris-HCl pH 8.0, 8 mM EDTA, 100 mM NaCl, 0,5% SDS). Lysates were incubated with proteinase K for 1h at 45°C, and DNA was extracted with phenol/chloroform followed by ethanol precipitation. DNA was pelleted, resuspended in water and resolved on an alkaline (NaOH) agarose gel. The gel was dried on a Biodyne membrane (Pall Corporation). <sup>32</sup>P incorporation was visualized by autoradiography and quantitated using the ImageQuant software.

### Protein extraction, IP, ChIP, IB, antibodies

TF-1 and MEL nuclear extracts were prepared as documented previously [79]. Chromatin extracts were performed by cross-linking TF-1 cells using formaldehyde at a final concentration of 0,5%. After 10 min., cross-linking was quenched by the addition of glycine to a final concentration of 0,125M. Cells were harvested, washed with PBS and protein extraction was performed by incubation with radio-immunoprecipitation (RIPA) buffer (10mM Tris-HCl pH8.0, 140mM NaCl, 1mM EDTA, 1% Triton X-100, 0,1% SDS, 0,1% deoxycholate). Extracts were sonicated. For IP, 2mg of protein were incubated overnight at 4°C with 10 $\mu\text{g}$  of antibody against LMO2 (R&D Systems, AF2726) in RIPA buffer. Protein complexes were immunoprecipitated by adding appropriately conjugated Pansorbin cells (Calbiochem) for 4h at 4°C, washed 3 times with RIPA buffer, and subjected to SDS-PAGE. After transfer on PVDF membranes (Millipore), proteins were visualized by immunoblotting using ECL Plus (GE Healthcare). The following antibodies were used for ChIP, IB and IF: mouse anti-CDC6 (Millipore, 05-550) rat anti-PRIM1 (Cell Signaling Technology, 4725S), rabbit anti-MCM5 (Bethyl Laboratories, A300-195A), mouse anti-tubulin (NeoMarkers, MS-582-P1). The rat anti-GATA1 (sc-265), rabbit anti-geminin (sc-13015), goat anti-laminB (sc-6217), goat anti-LDB1 (sc-11198), goat anti-MCM3 (sc-9850), mouse anti-MCM4 (sc-28317), mouse anti-MCM7 (sc-56429) and rabbit anti-PCNA (sc-7907) were all from Santa Cruz

Biotechnologies. The mouse anti-MCM2 (610700), mouse anti-MCM6 (611622), mouse anti-ORC2 (51-6875GR), mouse anti-POLD1 (610972) and mouse anti-SC35 (556363) were all from BD Biosciences. The BTL-73 mouse anti-SCL/Tal1 antiserum was generously provided by Dr D. Mathieu (Institut de Génétique Moléculaire, Montpellier, France).

### Microscopy

TF-1 cells were harvested, washed in PBS, fixed in 2% paraformaldehyde for 20 min at room temperature, washed, blocked and permeabilized with 0.1% NP40 in PBS containing 5% FCS (blocking buffer). Cells were then incubated with the specified antibodies in blocking buffer for 1 h and washed in 0.1% NP40 in PBS. Cells were then incubated with donkey FITC-conjugated anti-rabbit, Cy5-conjugated anti-goat and Cy3-conjugated anti-mouse antibodies (Jackson ImmunoResearch Laboratories) for 1h at room temperature and washed. The Hoescht dye was added (10 µg/ml) in PBS for 20 min. After washing, the cell pellet was deposited on the slide and mounted on a coverslide with Mowiol (Sigma-Aldrich). Images were acquired on an LSM 510 META confocal microscope (Zeiss) with a 63x magnification.

### RNA and ChIP analysis

For ChIP, the same procedure than for TF-1 IP was applied, except that we used 500 µg of extracts with 1 µg antibody. After precipitation, the immune pellet was recovered by addition of elution buffer (50mM Tris-HCl pH8.0, 10mM EDTA, 1% SDS), treatment with proteinase K and phenol/chloroforme extraction. For RT-PCR, total RNA was prepared from 50 000 cells as described previously [37]. Real-time quantitative PCR was done with PerfeCTa™ SYBR Green FastMix™ ROX (Quanta Biosciences) on a StepOnePlus Real-Time PCR System (Applied Biosystems).

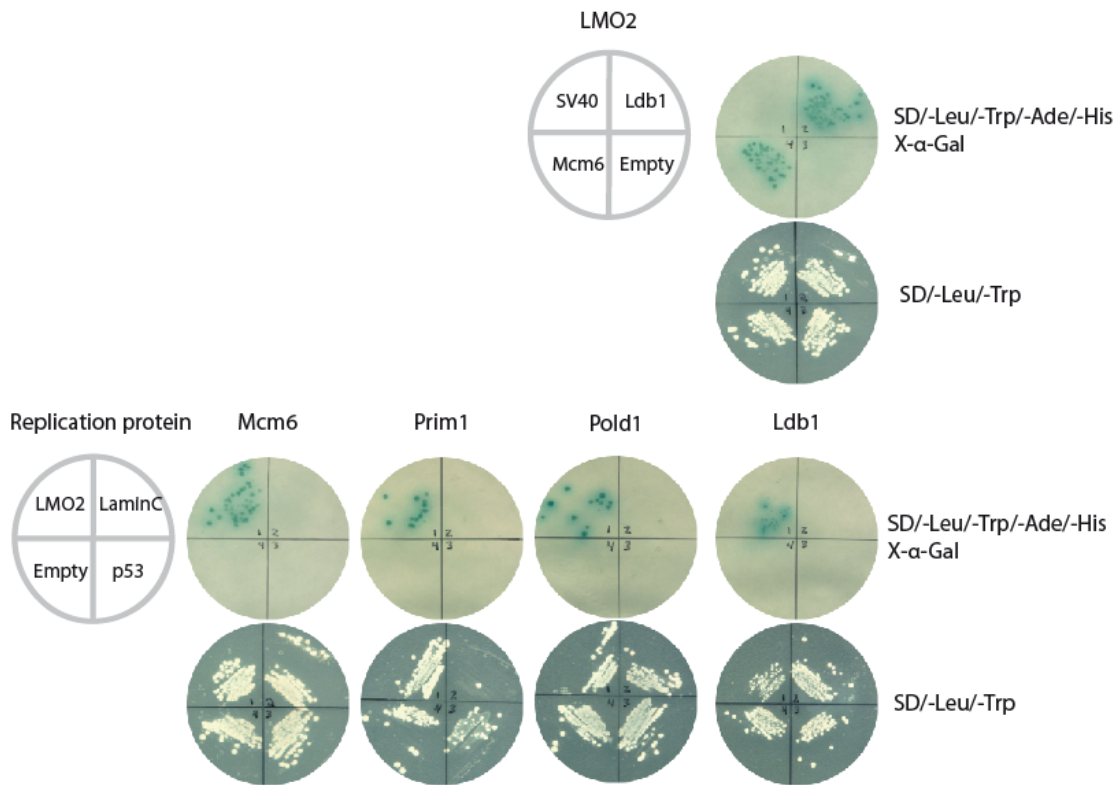
## **3.7 Acknowledgements**

We thank Dr Alain Verreault for critical discussion, Danièle Gagné (IRIC) for her assistance with cell sorting, Christian Charbonneau for his assistance with microscopy and imaging, Véronique Litalien for mice handling, Geneviève Boucher for bioinformatic analyses, Véronique Lisi and David Flaschner for assistance with



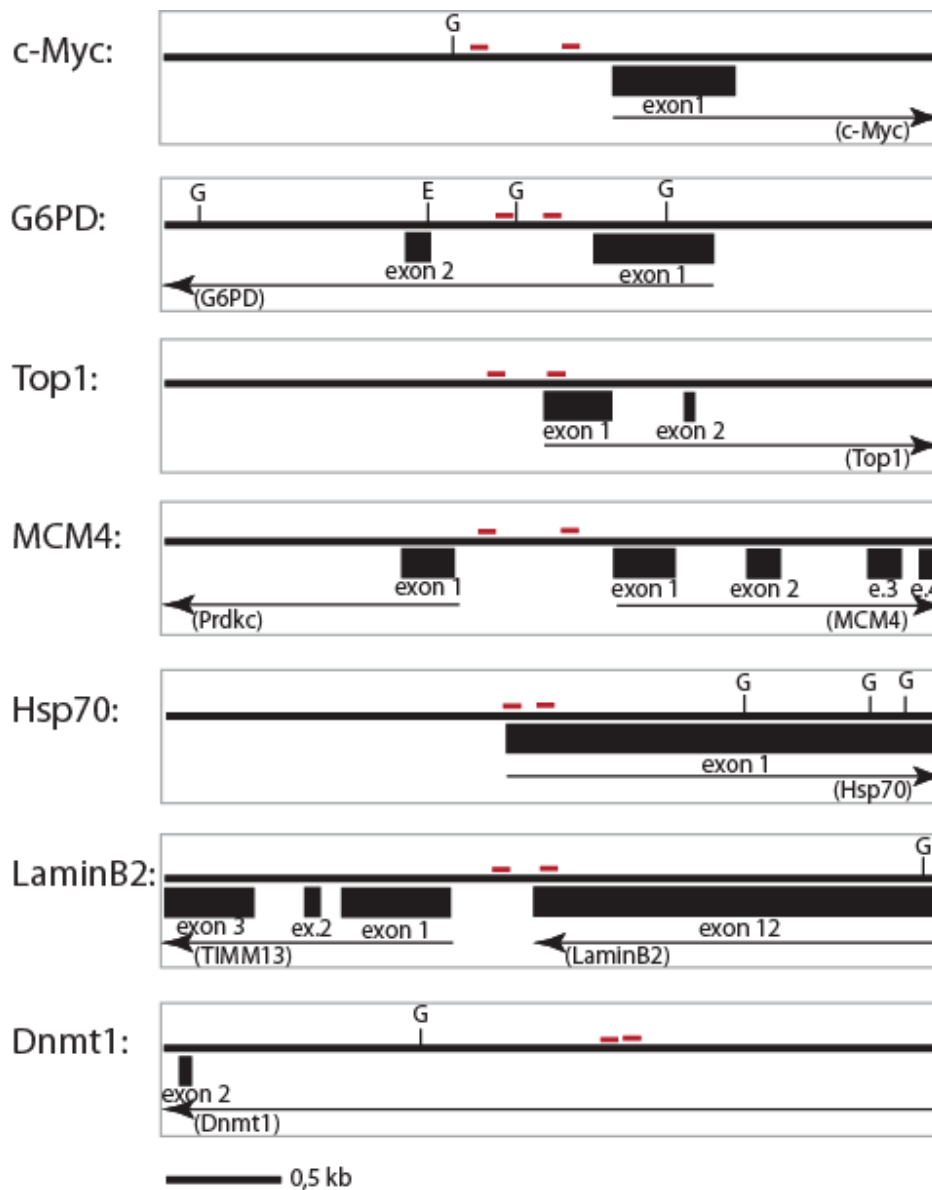
molecular biology techniques, Drs Jalila Chagraoui and Richard Martin for help with the retroviral gene transfer and the yeast-two hybrid technologies, respectively. This work was funded by the Canadian Institutes for Health Research (CIHR) and the Canadian Cancer Society Research institute to TH, a multiuser grant to support the flow cytometry and imaging service (CIHR) and a group grant from the Fonds de la Recherche en Santé du Québec to support in part IRIC infrastructure. MCS was supported by a Canada Graduate Scholarship Doctoral Award (CIHR) and a doctoral award from the Cole foundation.

### 3.8 Supplementary data



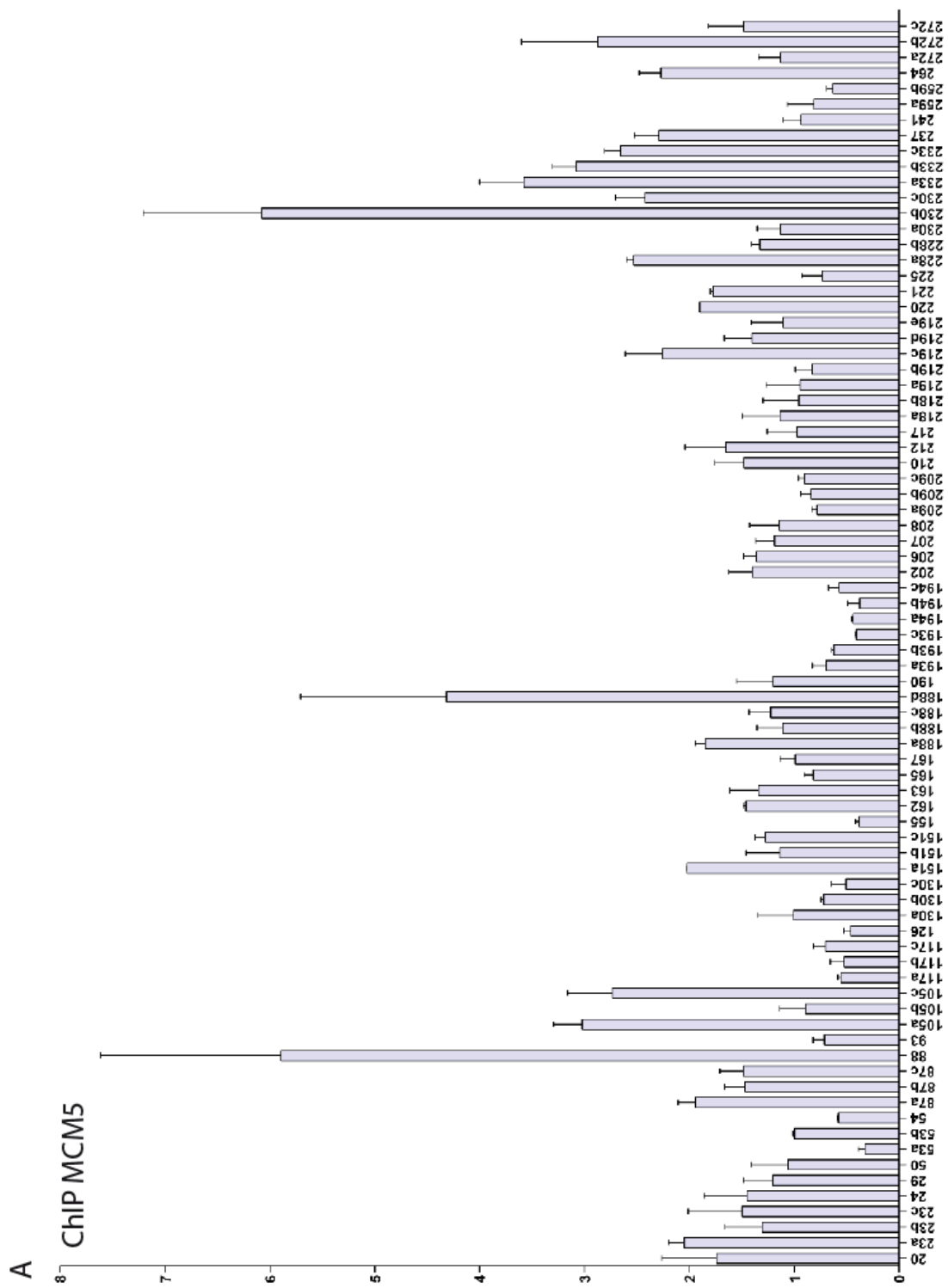
#### Supplemental Figure 3.7 LMO2 specifically interacts with replication proteins.

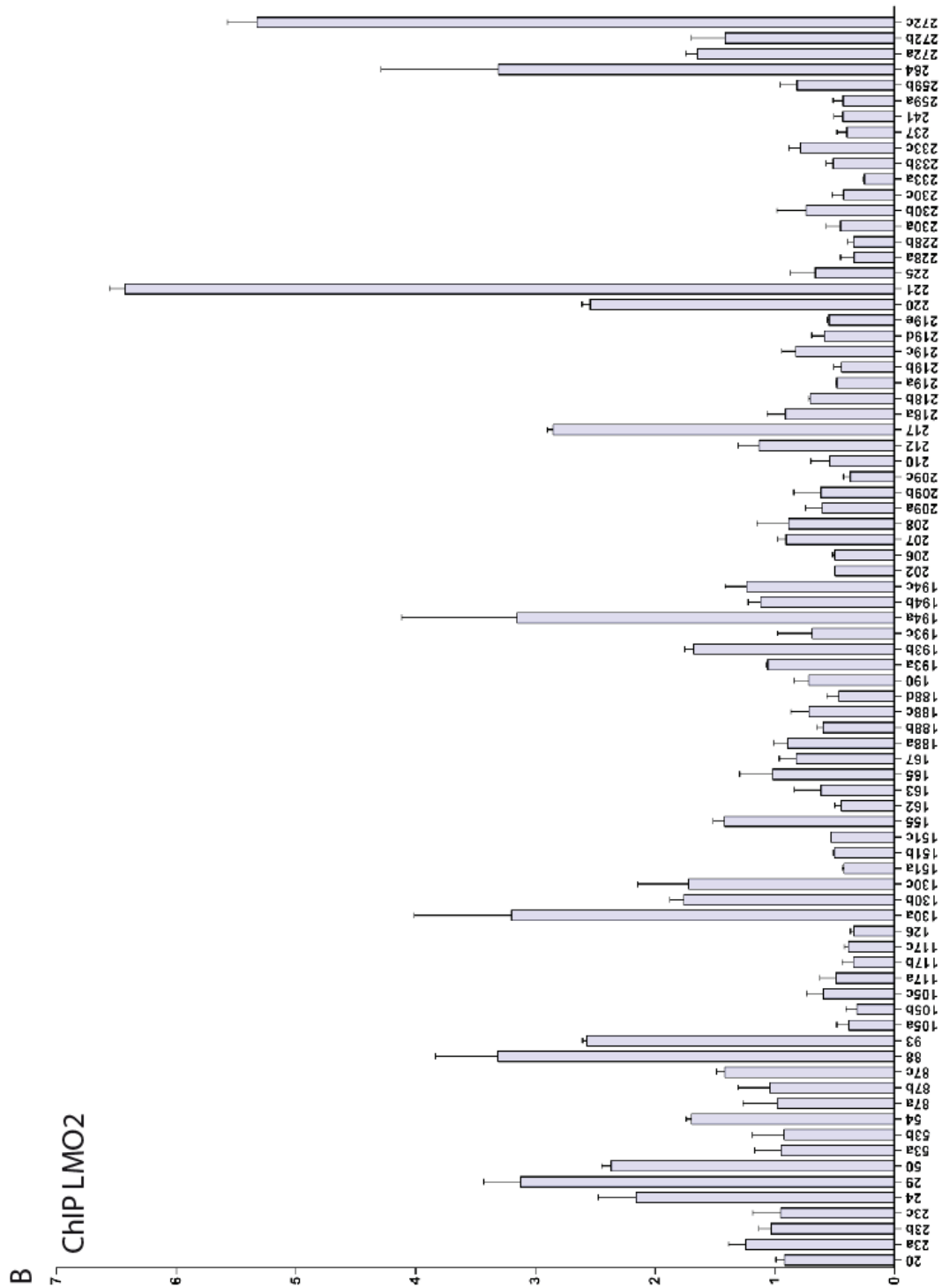
Yeast plates resulting from two hybrid experiments showed efficient yeast mating (lower panels) and selective interaction (upper panels) between LMO2 and MCM6, POL $\delta$ 1, PRIM1 and LDB1. No interaction was observed between LMO2 and SV40 large T antigen or the empty vector, and between replication proteins and lamin C, p53 or the empty vector.



**Supplemental Figure 3.8 Maps depicting the location of well characterized DNA replication origins and PCR probes.**

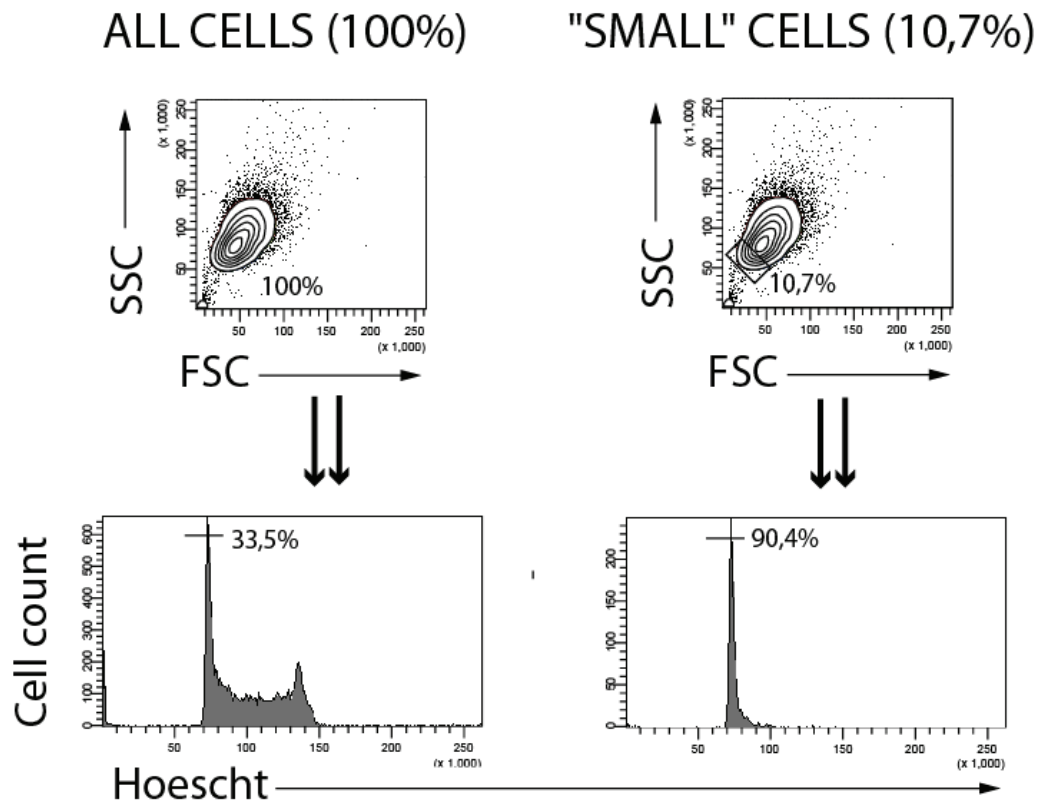
Gene exons are illustrated by black boxes. PCR primers are shown as red bars. G: GATA1/2 consensus sites; E: E47 consensus sites.





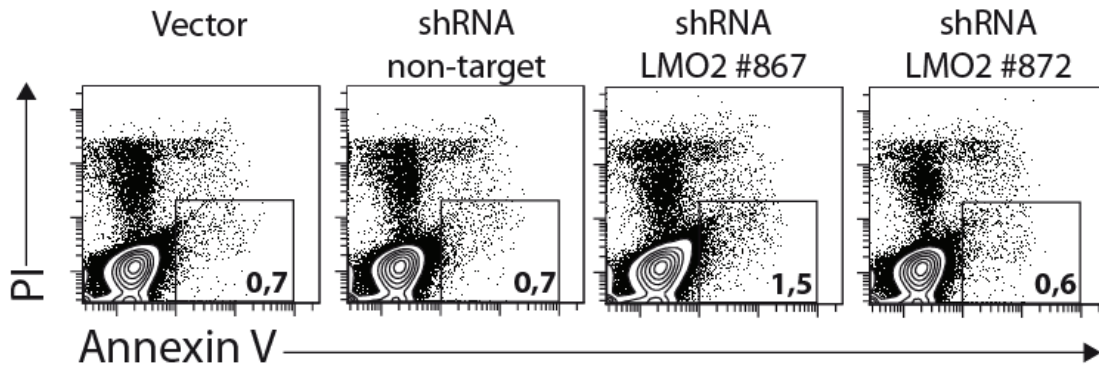
**Supplemental Figure 3.9 Chromatin immunoprecipitation of TF-1 cells with antibodies against MCM5 (A) or LMO2 (B).**

Histograms of the heatmap shown in Figure 3.4C. Data are the mean  $\pm$  SD of at least two independent experiments.



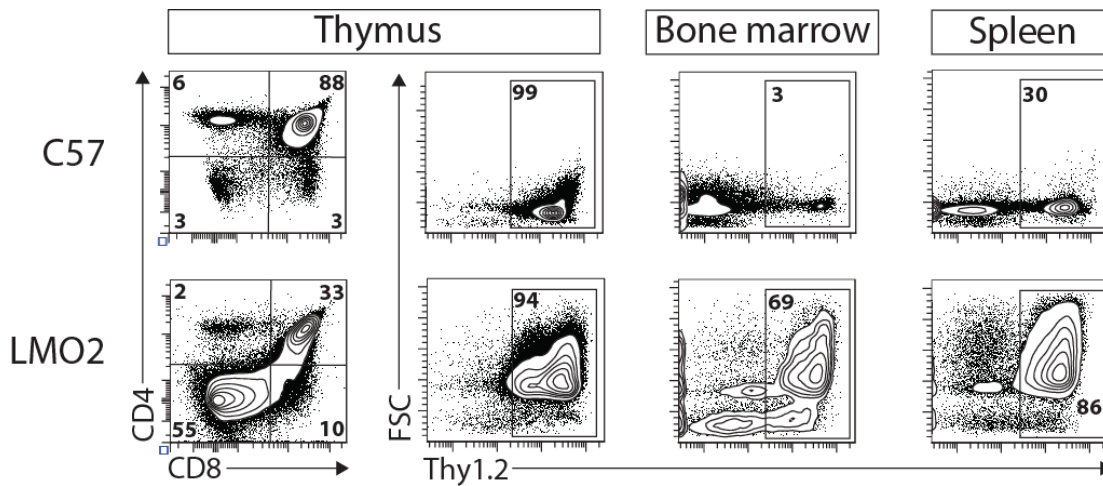
**Supplemental Figure 3.10 MEL cells can be synchronized in G<sub>0</sub>/G<sub>1</sub> by flow cytometry.**

Left panels: Cell cycle status of asynchronous MEL cells. Right panels: Cell cycle status of the 10% smaller cells determined by the FSC/SSC profile. More than 90% of the cells are in G<sub>0</sub>/G<sub>1</sub>.



**Supplemental Figure 3.11 Cells with reduced levels of LMO2 are not apoptotic.**

MEL cells infected as indicated were stained with the Annexin V apoptosis marker. Dead cells that stained with propidium iodide were excluded from the analysis.



**Supplemental Figure 3.12 Representative flow cytometry analysis of MSCV-LMO2 induced leukemia cells, compared to cells from a healthy mouse.**

In the thymus, an expansion of the double negative compartment is observed. Leukemic cells express Thy1, appear as blast cells (FSC profile) and invaded the bone marrow and the spleen.

## **4 Systematic mapping of protein-protein interactions for the oncoprotein LMO2**

Marie-Claude Sincennes<sup>1</sup>, Benoit Grondin<sup>1</sup>, Véronique Lisi<sup>1</sup>, Trang Hoang<sup>1,2</sup>

<sup>1</sup> Institute of Research in Immunology and Cancer (IRIC) – University of Montreal,  
Montréal, Québec, Canada H3C 3J7

<sup>2</sup> Pharmacology, Biochemistry and Molecular Biology Departments, University of  
Montréal, Québec, Canada H3C 3J7

This chapter is a work in progress and corresponds to a manuscript in preparation

## **4.1 Contribution des co-auteurs**

MCS : Design expérimental, réalisation des expériences, analyse des données, rédaction du manuscrit

BG : Design expérimental, réalisation des expériences (Figure 4), analyse des données

VL : Analyse des données (Figures 1B, 4C)

TH : Design expérimental, analyse des données, rédaction du manuscrit



## 4.2 Summary

The SCL and LMO2 transcription factors, which are protein interaction partners, are crucial for the generation of hematopoietic stem cells, and their aberrant expression in thymocytes lead to T cell acute lymphoblastic leukemia. While the transcriptome of SCL and LMO2 has been extensively studied, much less is known about the protein-protein interactions that govern their functions. Thus, identification of LMO2 and SCL protein partners could reveal new roles for these oncoproteins both in normal hematopoietic and leukemic cells. Here we performed an unbiased, systematic protein-protein interaction study for LMO2 and SCL specifically in hematopoietic stem and progenitor cells. While we identified only 8 interactions involving the SCL oncoprotein, we determined that LMO2 interacts with more than 130 different protein partners involved in various biological processes, particularly in transcription. The characterisation of SCL and LMO2 interactome will impact our knowledge of their molecular functions in hematopoiesis and leukemogenesis.

## 4.3 Introduction

The diversity of mature blood cells is generated from hematopoietic stem cells (HSC) that undergo lineage differentiation, a process controlled by the combined action of specific transcription factors. This is exemplified by the SCL (Stem Cell Leukemia) and LMO2 (LIM-Only 2) transcription factors, that associate together in a complex governing genetic programs essential for hematopoiesis ([320], reviewed in [461]). Indeed, targeted disruption of *Scl* or *Lmo2* in mouse models completely disrupts hematopoiesis [31, 51-55], underscoring their importance in the generation of hematopoietic stem cells. The SCL-LMO2 transcriptional complex is also required at different steps during hematopoietic lineage differentiation, particularly in erythropoiesis ([68, 79, 80], reviewed in [461]).

Many genome-wide studies described the transcriptional impact of SCL and LMO2 in different hematopoietic cells by the identification of their target genes [327, 330, 339]. Interestingly, SCL and LMO2 could act both as repressors and activators of transcription under different cellular contexts. For example, they associate on the

*Cdkn1a* promoter in different hematopoietic subpopulations [62, 329, 464], but *Cdkn1a* expression can be either activated [62, 329] or repressed [464], depending in part of the association of the co-repressor ETO2 to the complex [329]. This suggests that the association of SCL and LMO2 with different protein partners could dictate the transcriptional output of the complex.

In addition to their role in normal hematopoiesis, SCL and LMO2 are often aberrantly expressed in T cells from T-ALL patients, by the result of chromosomal translocations or genetic alterations [30, 135, 137, 138, 140, 141]. Moreover, in a X-SCID gene therapy trial, 4 patients developed T-ALL due to retroviral insertion in the *LMO2* locus [143-145]. In leukemic T cells, SCL and LMO2 also associate together [148, 317, 321] and perturb normal thymocyte transcriptional programs [37, 330, 335, 342]. Moreover, SCL and LMO2 collaborate together to accelerate leukemia onset in a mouse model [148]. This cooperation is explained in part by the different molecular roles of SCL and LMO2 within the transcriptional complex: SCL is responsible for tethering the complex to DNA via its bHLH domain, whereas LMO2 recruits LDB1 and GATA proteins via its LIM domains. Another possibility is that both SCL and LMO2 mediate protein-protein interactions important for transcriptional activity. Alternatively, SCL or LMO2 may possess molecular functions that are independent from each other, either in transcriptional complexes or in other cellular processes.

In this study, we aimed at identifying new SCL and LMO2-interacting proteins to dissect their molecular functions in hematopoietic cells. Different protein-protein interactions involving SCL and LMO2 have been reported in the literature [308, 329, 346, 347, 479-482], but their role in the regulation of SCL and LMO2 functions is not well described. However, the SCL-LMO2 or LDB1-LMO2 interactions are extensively characterised, possibly because these interactions are strong and frequent in the cell. Concordant with this, we previously demonstrated that 70% of the nuclear pool of LMO2 protein is found associated with SCL in erythroid cells [177]. However, other protein-protein interactions that appear more rarely, in specific cellular contexts, could be informative of the fine tuning of SCL and LMO2

functions. These subsets of interactions are hard to detect by protein purification approaches.

For these reasons, we identified new SCL and LMO2-interacting partners in hematopoietic stem/progenitor cells, by an unbiased, proteome-wide study using the yeast two hybrid technology. We identified a few SCL partners, but more than one hundred protein-protein interactions involving LMO2, many of which being observed in at least two different experimental conditions. LMO2 protein partners are associated with various cell processes including the regulation of transcription. These new interactions are informative of the relative roles of SCL and LMO2 in hematopoietic cells: SCL is mainly a DNA-binding protein that associates with a few protein partners, whereas LMO2 is responsible to nucleate protein complexes that regulate transcription and possibly other cellular functions.

#### 4.4 Results

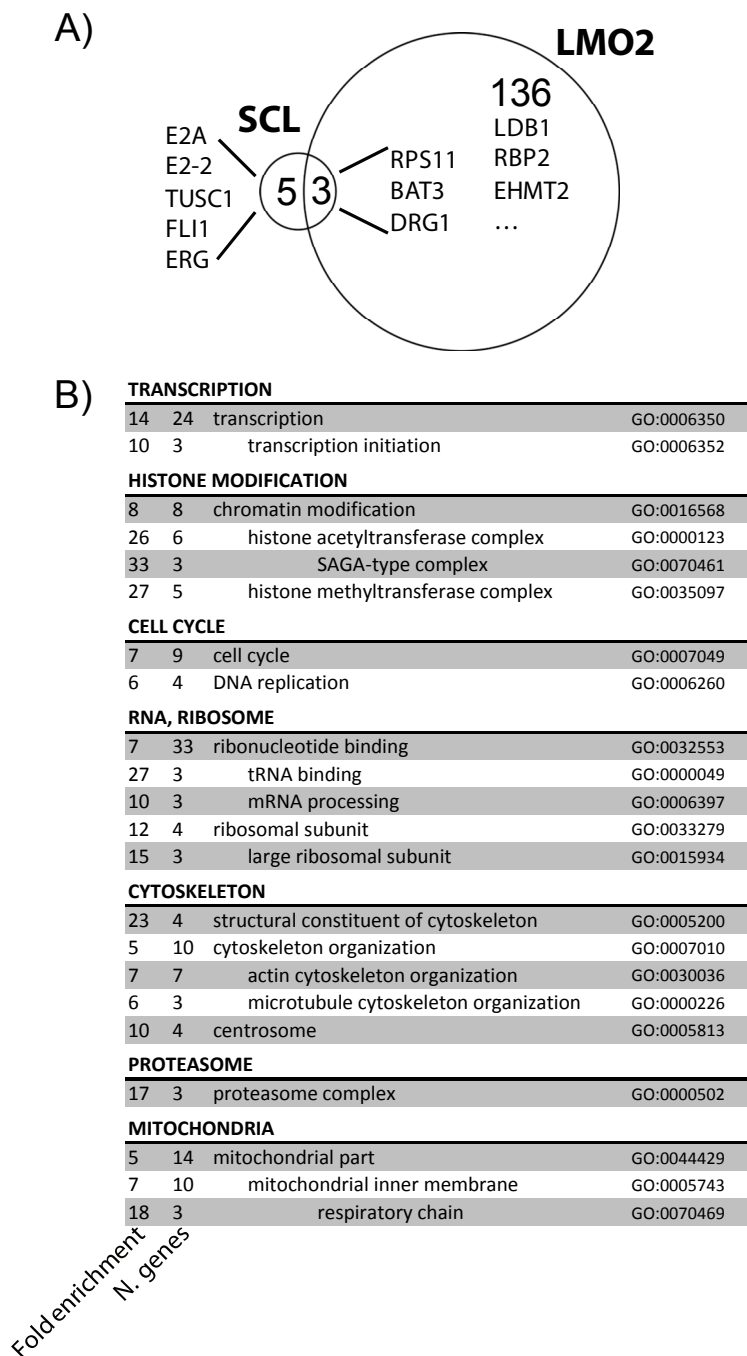
SCL and LMO2 are expressed in hematopoietic stem and progenitor cells. However, these cells are very rare, and are not easily amenable to large-scale proteomic approaches. In this context, the yeast two hybrid technique is a well proven unbiased method to uncover protein-protein interactions in pairs. To identify SCL and LMO2-involved interactions specifically in stem/progenitors, a cDNA library has been constructed from c-Kit<sup>+</sup> Lin<sup>-</sup> murine bone marrow cells (**Supplemental Figure 4.5**), in which each library gene is expressed as a fusion protein with the GAL4 activation domain. We obtained 3x10<sup>6</sup> independent clones, and the quality of the library was verified by digestion of plasmid DNA from 60 random clones (**Supplemental Figure 4.5**). The insertion rate of the library is of about 80%, and the insert length can reach 1.8 kb, with an average length of 0.75 kb.

To perform the yeast two hybrid screens, the cDNAs encoding full-length SCL or LMO2 have been subcloned into the pGBKT7 expression vector, and are expressed in the Y187 yeast strain as a fusion with the GAL4 DNA binding domain. The vector contains a tryptophane selection marker. After yeast mating, diploid yeasts expressing interacting proteins have been selected on media devoid of adenine and histidine,

since the AH109 strain contains *ADE2*, *HIS3* and *MEL1* (encoding  $\alpha$ -galactosidase) GAL4-responsive elements. The SCL and LMO2 interaction screens were performed in saturating conditions and all positive clones growing on selection media have been sequenced.

For SCL, the positive clones encoded 8 different genes (**Figure 4.1A**), three of these (E2A, E2-2 [323-326], DRG1 [479, 482]) being known SCL-interacting partners. In addition, we identified FLI1 and ERG, which have recently been shown to associate together with SCL and LMO2 in a significant proportion of SCL target genes [339]. In striking contrast with SCL, we identified 139 protein-protein interactions involving LMO2 (**Figure 4.1A and Supplemental Table III**). This high number of clones is not due to the activation of reporter genes by the bait in absence of protein interaction, since LMO2-transformed yeasts did not grow when plated on the selection media (not shown). From the positive clones, LDB1, EHMT2 and RBP2 were identified, which are known interacting partners of LMO2 [346, 462, 483]. Three proteins (RPS11, BAT3, DRG1) were found in common from both screens, and for one of these (DRG1), the specific interaction with SCL was confirmed *in vitro* and *in vivo* [482], excluding the possibility that this reflects a false positive event from the screen. Interestingly, four SCL-specific interacting partners out of five are transcription factors, whereas protein partners shared by SCL and LMO2 are involved in other cellular processes. These results suggest that the SCL interactome is modest and more restricted to transcriptional activities, compared to that of LMO2. Since the SCL-LMO2 transcriptional multi-protein complex is highly dynamic and changes in composition depending on the cell type and on the target gene, these results also suggest that LMO2 is more prone to bring diversity to the complex than SCL.

We then used bioinformatics tools to highlight Gene Ontology (GO) terms that were significantly enriched in the list of SCL and LMO2 partners, compared to 10 000 lists of random genes. Not surprisingly, SCL and LMO2 partner lists were enriched in GO terms associated with transcription and chromatin modification. For LMO2, GO terms associated with other cell processes were enriched compared to random sets of



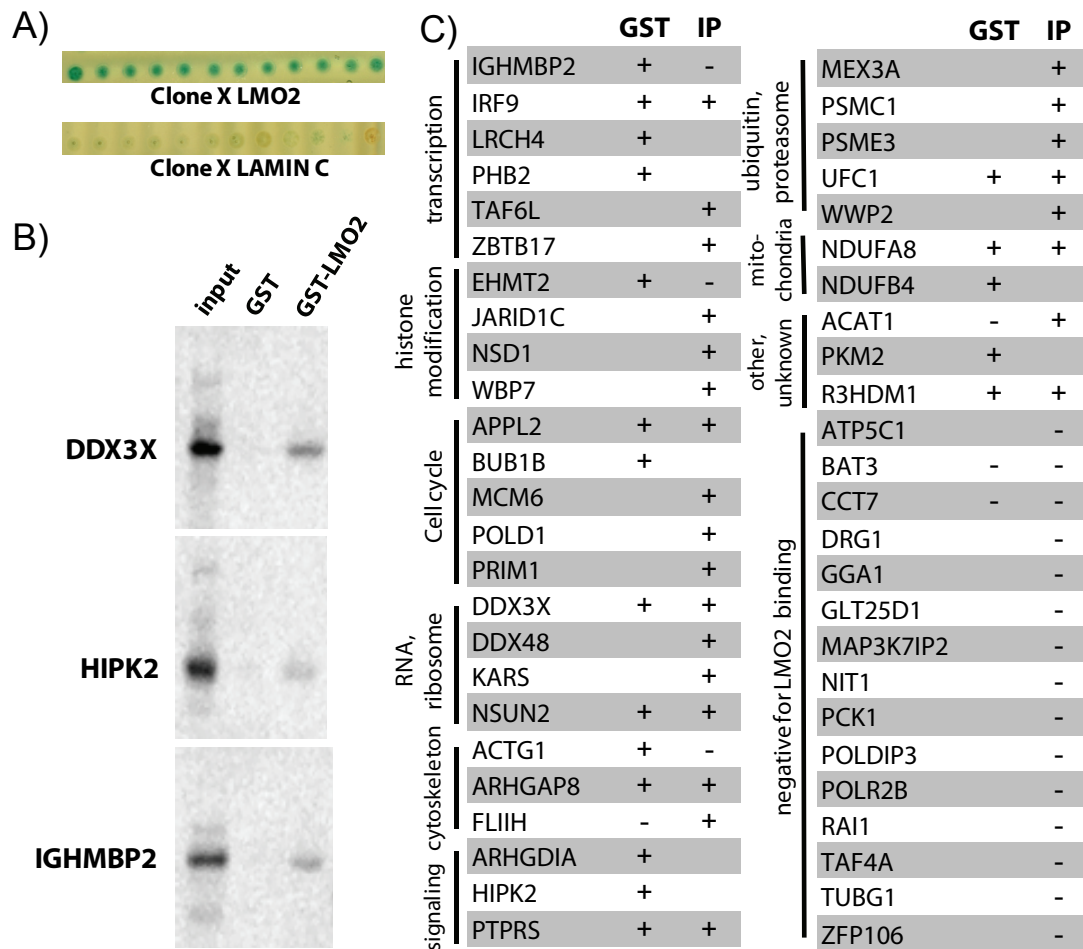
**Figure 4.1 Identification of SCL and LMO2 protein partners.**

A) Venn diagram representing the number of partners revealed by yeast two hybrid with SCL and LMO2. SCL interacting partners and a few LMO2 partners are also indicated. B) Gene Ontology (GO) terms enrichment analysis of SCL and LMO2-interacting proteins. The representation of each GO term has been measured within the 8 and the 139 SCL and LMO2 partners, respectively, compared to  $10^4$  lists of 139 random genes. The number of genes associated with each enriched GO term, as well as the fold enrichment (LMO2 binding partners vs random genes) are indicated.

genes (**Figure 4.1B**). For example, GO terms associated with cell cycle, particularly with DNA replication, were identified from the analysis, according to the new role for LMO2 in DNA replication that we recently characterized (Sincennes MC et al, *to be submitted*). Interestingly, LMO2 partners were associated with diverse cellular functions, in the nucleus as well as in other organelles, suggesting that LMO2 could have multiple functions in different cellular localisations.

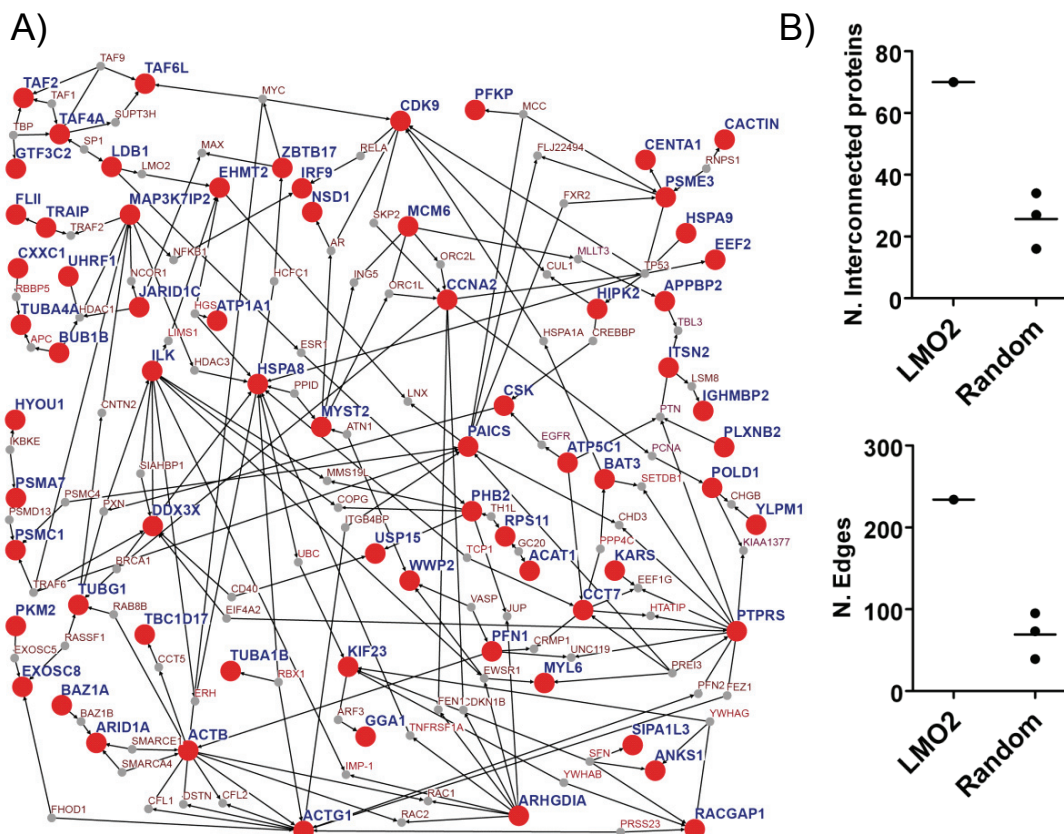
As a secondary screen, we tested all individual positive clones for their interaction with LMO2, with LaminC or with the empty pGBKT7 vector, on selection media (SD/-Leu/-Trp/-His/-Ade/+X- $\alpha$ -gal) (**Figure 4.2A**). All clones interacted with LMO2, but not with laminC or the empty vector, underscoring the specificity of the interactions. To determine if these interactions could be observed *in vitro*, we used two different techniques: LMO2-GST pulldown, and *in vitro* immunoprecipitation assays with an anti-LMO2 antibody. Both techniques involved  $^{35}\text{S}$ -radiolabelled proteins which were incubated with LMO2-GST or with nuclear extracts of pcDNA3-LMO2 transfected 293T cells (**Figure 4.2B, C**). There was an overall concordance between these two techniques which allowed us to confirm 35 interactions out of 50 proteins in all seven functional categories shown in **Figure 4.1B**. In 5 cases, we observed different outcomes between the two techniques, possibly because the interaction interface was masked by the LMO2 antibody or by GST fusion. Fifteen interactions were not revealed by *in vitro* binding assays with LMO2, which could be due to the experimental conditions that were not optimal for LMO2 binding to these specific proteins. Overall, 70% of protein interactions revealed by the yeast-two hybrid technique were independently confirmed *in vitro*, suggesting that the proportion of false positive within LMO2 interactors is relatively low.

We next asked the question whether the proteins that we identified using LMO2 as bait were reported to be part of protein-protein interaction networks (**Figure 4.3**). We built a protein-protein interaction network using the Osprey program. In parallel, we also tried to build similar networks between random proteins. Strikingly, 70 of 133 proteins recognized in the Osprey database were involved in a complex network with



**Figure 4.2 Protein interactions involving LMO2 could be detected by different techniques.**

A) Representative yeast two hybrid assay performed with all positive clones obtained from the LMO2 screen. All clones have been tested for interaction with LMO2 (top), LAMIN C (bottom) or the empty pGBKT7 vector (not shown). B) Representative GST pull-down results for 3 different proteins as indicated. C) Interaction (+ or -) between LMO2 and the indicated proteins have been tested by GST pull-down or *in vitro* immunoprecipitation as specified. Proteins are classified according to their main known function. Proteins negative for LMO2 binding (not classified by function) are also indicated.



**Figure 4.3 LMO2 partners are involved in a complex protein-protein interaction network.**

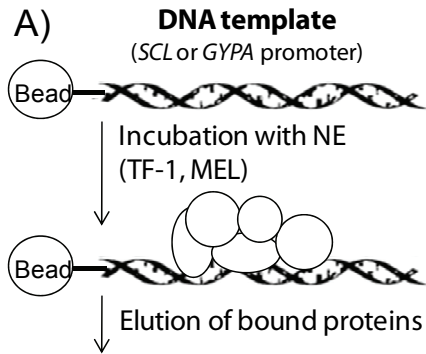
A) Protein-protein interaction network between LMO2 protein partners ( $n=133$  present in Osprey database, represented by large red dots and blue characters), either direct or mediated by one intermediate partner (represented by small grey dots), built from the Osprey Network Visualisation System. Three independent networks have also been built from 3 different sets of random proteins ( $n=133$  present in Osprey database) (not shown). B) Features of the Osprey network obtained for LMO2 partners, compared to 3 independent networks built from random sets of proteins.



234 edges. In comparison, the networks built from 133 Osprey-recognized random proteins did not reach a similar level of complexity with 2.5-fold less interconnected proteins and 3.4-fold lower numbers of edges (**Figure 4.3B and Supplemental Figure 4.6**). This suggests that LMO2 could be a central node that interconnects several protein networks.

Since transcription was the most enriched GO term associated both with SCL and with LMO2 partners, we then asked whether these newly-identified interactions could take place on DNA. To this end, we identified proteins associated to two SCL/LMO2 target genes by mass spectrometry. The SCL/LMO2 target genes used are *SCL* [339] and *GLYCOPHORIN A (GYPA)* [79]. The *SCL* promoter contains two GATA binding sites, whereas the *GYPA* promoter contains two GATA and one E-box binding sites. We performed a DNA capture assay, in which protein complexes bound to promoter templates are isolated from MEL or TF-1 nuclear extracts (**Figure 4.4A**). In order to control the capture of protein complexes that specifically assemble via the GATA and E-box binding sites, we performed simultaneous DNA capture with the *SCL* and *GYPA* promoters mutated in these elements. Captured protein complexes were silver stained and subjected to mass spectroscopy for identification of individual proteins. MS outputs for which equal or close to equal peptide numbers were obtained from both the wt and mutated templates were not considered to be captured in a GATA / E box dependent manner. As expected [79, 329], our DNA capture assay our GATA / E box dependent MS output lists show members of the SCL/LMO2/GATA transcriptional complex (**Figure 4.4B**).

GO terms associated with hematopoiesis, transcription and chromatin modification were enriched in the list of proteins identified, consistent with the procedure that specifically selects promoter-bound proteins from hematopoietic cells (**Figure 4.4C**). Strikingly, many SCL and LMO2 interactors identified by two hybrid were present on their target genes, specifically on intact, non mutated, promoters (**Figure 4.4D**). This suggests that LMO2 could be responsible for their tethering to promoters. Interestingly, many of these proteins are not known to regulate transcription, raising



**Identification of bound proteins by mass spectrometry**

**C)**

**TRANSCRIPTION**

12	22	transcription activator activity	GO:0016563
10	21	transcription repressor activity	GO:0016564
171	6	integrator complex	GO:0032039
38	3	transcription factor TF1C complex	GO:0033276

**HISTONE MODIFICATION**

16	24	chromatin modification	GO:0016568
32	9	histone methyltransferase complex	GO:0035097
27	4	MLL1 complex	GO:0071339
24	8	histone acetyltransferase complex	GO:0000123
49	4	STAGA complex	GO:0030914
24	3	histone deacetylation	GO:0016575
58	5	NuRD complex	GO:0016581

**HEMOPOIESIS**

23	3	regulation of erythrocyte differentiation	GO:0045646
11	4	T cell differentiation	GO:0030217

**NUCLEAR EXPORT**

12	5	regulation of nucleocytoplasmic transport	GO:0046822
70	3	regulation of protein export from nucleus	GO:0046825

**CHROMOSOME**

10	34	chromosome organization	GO:0051276
15	4	condensed chromosome	GO:0000793
111	3	sister chromatid cohesion	GO:0007062

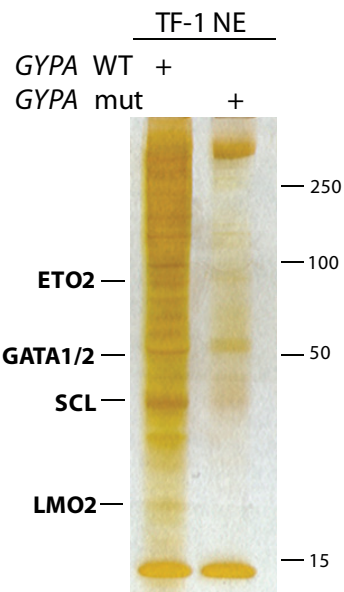
**RNA PROCESSING**

9	28	RNA processing	GO:0006396
27	13	mRNA processing	GO:0006397
10	9	spliceosomal complex	GO:0005681
108	4	U5 snRNP	GO:0005682

**CELL CYCLE**

8	12	cell division	GO:0051301
10	9	DNA replication	GO:0006260
78	4	DNA unwinding during replication	GO:0006268
13	5	cell cycle checkpoint	GO:0000075
15	3	mitotic cell cycle checkpoint	GO:0007093

**B)**



**D)**

	SCL		GYPA
	MEL	TF-1	TF-1
ACTB	13, 9	4, 0	9, 0
ACTG1*	1, 1	0, 7	0, 5
ATP5C1		1, 2	
BAZ1A	3, 0		9, 0
CCDC109A			2, 0
CDK9	2, 1		
E2A*		6, 11	8, 5
E2-2*		2, 7	1, 1
FIP1L1	2, 5	1, 5	9, 2
HSPA8	6, 0	1, 0	
KIF23			5, 0
LDB1*	14, 13	11, 16	12, 7
NFIC		2, 2	
PHB2*	1, 8		0, 5
PIGS			2, 0
POLDIP3		4, 0	7, 2
RPL17			2, 0
RPS11	3, 0	2, 4	
TFAP4		2, 0	5, 9
TUBA4A	2, 2	0, 4	1, 2
ZBTB17*		0, 1	1, 0

**Figure 4.4 Identification of LMO2 and SCL partners associated on their target genes.**

A) Diagram representing the experimental procedure for the DNA binding assay. B) Representative silver stained protein gel after capture of protein complexes on wild type (wt) or mutated (mut) *Glycophorin A* promoter (*GYP A*) with TF-1 nuclear extracts. The relative positions of proteins from the LMO2/SCL/GATA transcriptional complex, identified by mass spectrometry, are indicated. C) Gene Ontology (GO) terms enrichment analysis of proteins identified from the DNA capture assay, compared to  $10^4$  lists of random genes. The fold enrichment (first number), as well as number of genes associated with each enriched GO term (second number), are indicated. D) LMO2 protein partners identified by yeast two hybrid and DNA capture assays. Only the proteins identified by DNA capture in a GATA or GATA/E-box dependent manner were considered. The number of peptides identified from two experiments is indicated (peptide number exp.1, peptide number exp.2). \* indicates that the interaction has been confirmed *in vitro* by GST pulldown or immunoprecipitation with SCL or LMO2 as a bait.

the question of their purpose on transcriptional promoters. Combining these results with the ones observed with GST pulldowns and immunoprecipitations, 54 different LMO2 interactions have been observed in at least one different context in addition to the yeast two hybrid screen.

## **4.5 Discussion**

In the present study, we identified new protein-protein interactions involving SCL and LMO2, some of which that may be relevant in the context of normal hematopoietic or leukemic cells. The cDNA library that we constructed from c-Kit<sup>+</sup> Lin<sup>-</sup> hematopoietic stem and progenitor cells is a useful tool to any project aiming at characterizing protein-protein interactions in hematopoietic stem and progenitor cells.

### **4.5.1 Comparison of SCL and LMO2-interacting partners**

From the yeast two hybrid screen, more than 130 new LMO2 partners have been identified. This high number of identified interactions was not expected, since LMO2 is a very small protein. However, LMO2 protein contains essentially two LIM domains, which role is to mediate protein-protein interactions [185]. LIM domains are very versatile in their binding properties, and were found to bind several different protein motifs [185]. Other LIM-only proteins have the potential to mediate multiple different protein-protein interactions: for example, FHL2 (four and a half LIM domains 2) has more than 50 different protein partners reported to date, with very diverse functions in the nucleus as well as in the cytoplasm (reviewed in [484]). Like other members of its family, LMO2 seems to have the potential to interact with many different protein partners. In comparison, a similar yeast two hybrid screen with SCL as bait revealed only 8 different partners, suggesting that the library do not account for the high number of clones observed for the LMO2 screen. Thus, LMO2 partners are unlikely to be non-specific.

SCL protein partners identified in this study are mainly transcription factors, underscoring the primary role of SCL in the regulation of gene expression. Like LIM domains, HLH domains are known to mediate, probably in a more restricted manner,

protein-protein interactions. Particularly, HLH domains are involved in heterodimerization between bHLH transcription factors. Accordingly, we identified known heterodimerization partners of SCL, i.e. E2A and E2-2 [323-326]. In addition, two transcription factors from the ETS family, FLI1 and ERG, have been identified. Recently, a genome-wide study revealed that in hematopoietic stem/progenitor cells, an important proportion of SCL and LMO2 target genes are co-occupied by FLI1 and by ERG [339]. The identification of FLI1 and ERG by two different unbiased studies underscores the importance of these proteins in the regulation of SCL and LMO2 transcriptional function.

Interestingly, all interactions involving LMO2 have been confirmed by independent yeast two hybrid assay, whereas GST pulldown and immunoprecipitation assays confirmed 70% of the tested interactions *in vitro*. This verification rate is comparable to another large-scale study which employed the yeast two hybrid methodology for the identification of protein-protein interactions between transcription factors [483]. Moreover, 21 protein partners of SCL and LMO2 were detected on the *SCL* and/or *GYP A* promoters. Together, these results suggest that many of the newly identified LMO2-involved interactions are likely to be relevant in a biological context.

#### **4.5.2 The LMO2 interactome: more than transcription**

About one third of newly identified LMO2 protein partners are involved in the regulation or transcription and chromatin remodelling, in agreement with the known function of LMO2 as a transcription factor. Accordingly, several LMO2 partners were also detected on the *SCL* and/or *GYP A* promoters, two LMO2 target genes. The DNA capture technique coupled to mass spectrometry identified proteins that are almost exclusively involved in DNA-dependent processes (**Figure 4.4C**), according to the nuclear extraction performed prior to the assay. In contrast, the yeast two hybrid technique identified partners associated with broader cellular functions (**Figure 4.1B**), but is restricted to interactions involving LMO2. These differences underscore the complementarity of both methods in the identification of protein-protein interactions. Surprisingly, the combination of DNA capture and yeast two

hybrid methodologies pointed to many proteins that are not known to be involved in transcriptional regulation. Using DNA capture alone, these candidates would probably have been discarded from further analyses. The interaction of these candidates with LMO2 as assessed by yeast two hybrid raise the intriguing possibility that these proteins participate actively in the regulation of LMO2 target genes.

Unexpectedly, a large proportion of LMO2 protein partners are involved in non-transcriptional functions, since LMO2 new protein partners are significantly associated with GO terms representing ribosomal, mitochondrial and cytoskeletal functions. A proportion of these proteins could be involved in transcriptional regulation, or alternatively, new cellular functions could be attributed to LMO2. As a proof of principle, we recently characterized a new role for LMO2 in the regulation of DNA replication, by direct interaction with MCM6, DNA polymerase delta and DNA primase identified from the yeast two hybrid screen (Sincennes MC et al, *to be submitted*). Since LMO2 can be detected in the cytoplasm [44], it is not surprising to identify potential partners within this cell compartment. Particularly, a large proportion of LIM domain proteins are involved in cytoskeletal functions [185], suggesting that this function could be a common feature of LIM domain proteins including LMO2. LMO2 being an adaptor protein, it is likely that LMO2 could be part of different protein complexes, regulating diverse functions in addition to its role in transcription and in DNA replication.

### **4.5.3 LMO2 protein interactions in normal and leukemic stem cells**

Recently, LMO2 was shown to induce self-renewal of thymocyte progenitors (double-negative, DN T cells) [114]. In normal hematopoiesis, the ability to self-renew is restricted to hematopoietic stem cells (HSC), but in leukemia and in different cancers, a small population of cells, called cancer stem cells, possess the ability to self-renew and can generate tumors were transplanted into an immunodeficient host (reviewed in [485, 486]). In LMO-induced leukemias, the leukemia-initiating cell is a committed T cell progenitor whose biological properties have been modified by aberrant LMO expression [114, 122]. LMO2 is also required at the onset of normal hematopoiesis [31, 55], probably in the generation of

hematopoietic stem cells. Together, these results suggest that LMO2 is important in the regulation of stem cell biology, and that misexpression of LMO2 could confer stem cell-like properties to a more mature cell. By studying LMO2 function in normal hematopoietic stem and progenitor cells, we could reveal important roles for LMO2 both in normal hematopoiesis and in leukemogenesis. In addition, it would be interesting to assess the interaction of these newly identified protein partners with other members of the LMO family, which share homology with LMO2 and act as oncogenes in different tissues [135, 278, 282, 289, 296, 297].

## 4.6 Experimental procedures

### Cell culture and cell sorting

MEL cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) and passed every 2 days at the concentration of  $5 \times 10^4$  cells/ml. TF-1 cells were cultured as described previously [478]. For library construction, primary murine bone marrow cells were depleted of lineage positive ( $\text{Lin}^+$ ) cells through immunomagnetic bead cell separation as previously described [62]. The remaining  $\text{Lin}^+$  cells were further excluded by flow cytometry using fluorescein isothiocyanate (FITC)-conjugated antibodies against lineage markers CD3, B220, Gr1, CD11b, Fc $\epsilon$ RI and TER119. c-Kit $^+$  Lin $^-$  cells (or TER119 $^{+/-}$  cells for fetal livers) were sorted on FACSaria (BD Biosciences).

### cDNA library construction and yeast two hybrid screening

RNA extraction from purified bone marrow cells was performed as described previously [37]. Poly A $^+$  mRNA was purified using Oligotex beads (Qiagen). RT-PCR was performed according to the Library Construction and Screening Kit User Manual (Clontech), in separate reactions containing either random or dT primers. The library was constructed with 50% amplified cDNA from random primers and 50% from dT primers. The Sma1-linearized pGADT7-Rec vector and the amplified cDNA were transformed in the AH109 yeast strain and the cDNA was incorporated into the vector by homologous recombination in yeast. The vector possessing a leucine selection marker, the library-transformed yeasts grow on media without leucine. The

3X10<sup>6</sup> independent colonies were pooled and aliquots were frozen. Yeast two hybrid screening was performed by yeast mating between the AH109 (library) and Y187 (bait) strains, and colonies growing on selective media (SD/-Leu/-Trp/-Ade/-His/+3-amino-1,2,4-triazole) were isolated. Yeast DNA was extracted and transformed into bacteria. Bacterial plasmidic DNA was extracted and sequenced.

#### GST pulldown, immunoprecipitation

GST pulldown experiments were performed as previously described [79]. For *in vitro* immunoprecipitations, 10µl of (<sup>35</sup>S)methionine-*in vitro* translated proteins were incubated with 10µg nuclear extracts of pcDNA3-LMO2 transfected 293T cells and with 0,1µg of antibody against LMO2 (R&D Systems, AF2726) in IP buffer (20mM Tris-HCl pH8.0, 10% glycerol, 6,25mM MgCl<sub>2</sub>, 0,1% NP40, 2,5mM DTT, 0,1mM EDTA, 100mM NaCl), overnight at 4°C. Protein complexes were immunoprecipitated by adding appropriately conjugated Pansorbin cells (Calbiochem) for 4h at 4°C, washed 3 times with IP buffer, and subjected to SDS-PAGE. After transfer on PVDF membranes, <sup>35</sup>S-radiolabelled proteins were visualized by autoradiography.

#### DNA capture assay and mass spectrometry

The -84 *GYP A* and -562 *SCL* promoter (human promoter from -562 to +7) fragments, wild type or mutant for the GATA sites and E-box, were biotinylated and subsequently immobilized on streptavidin-coated magnetic beads (Dynabeads M-280) as described in [329, 487]. DNA capture assays with nuclear extracts from TF-1 or MEL cells were performed as previously described [329, 487]. DNA bound proteins were eluted by boiling the beads in SDS-PAGE sample buffer, and resolved on polyacrylamide gels. Silver stained gel bands were digested by trypsin and analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analyses as previously described [488].

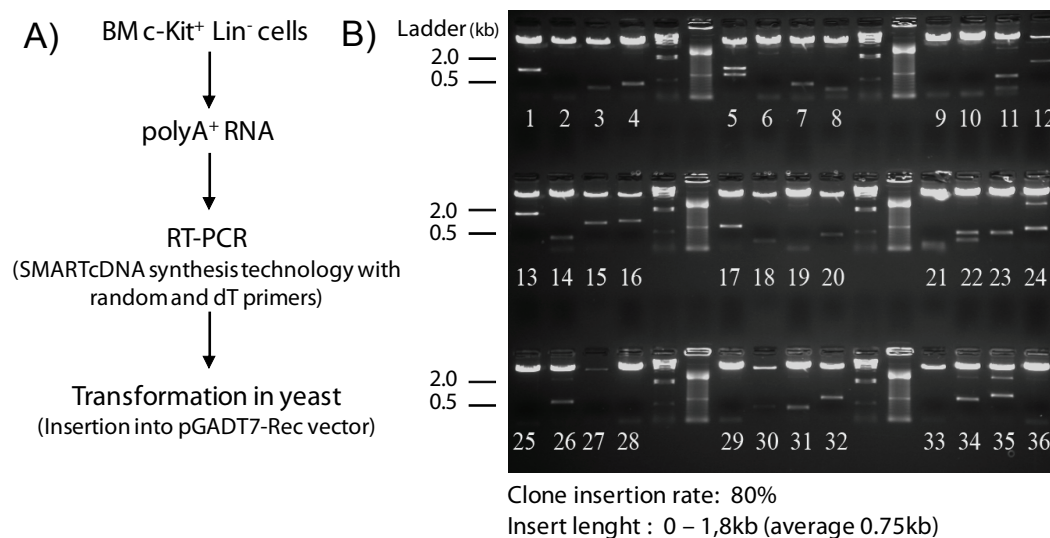
## **4.7 Acknowledgements**

We thank Dr Richard Martin for coaching on cDNA library construction, Danièle Gagné (IRIC) for assistance with cell sorting and the IRIC proteomics core facility



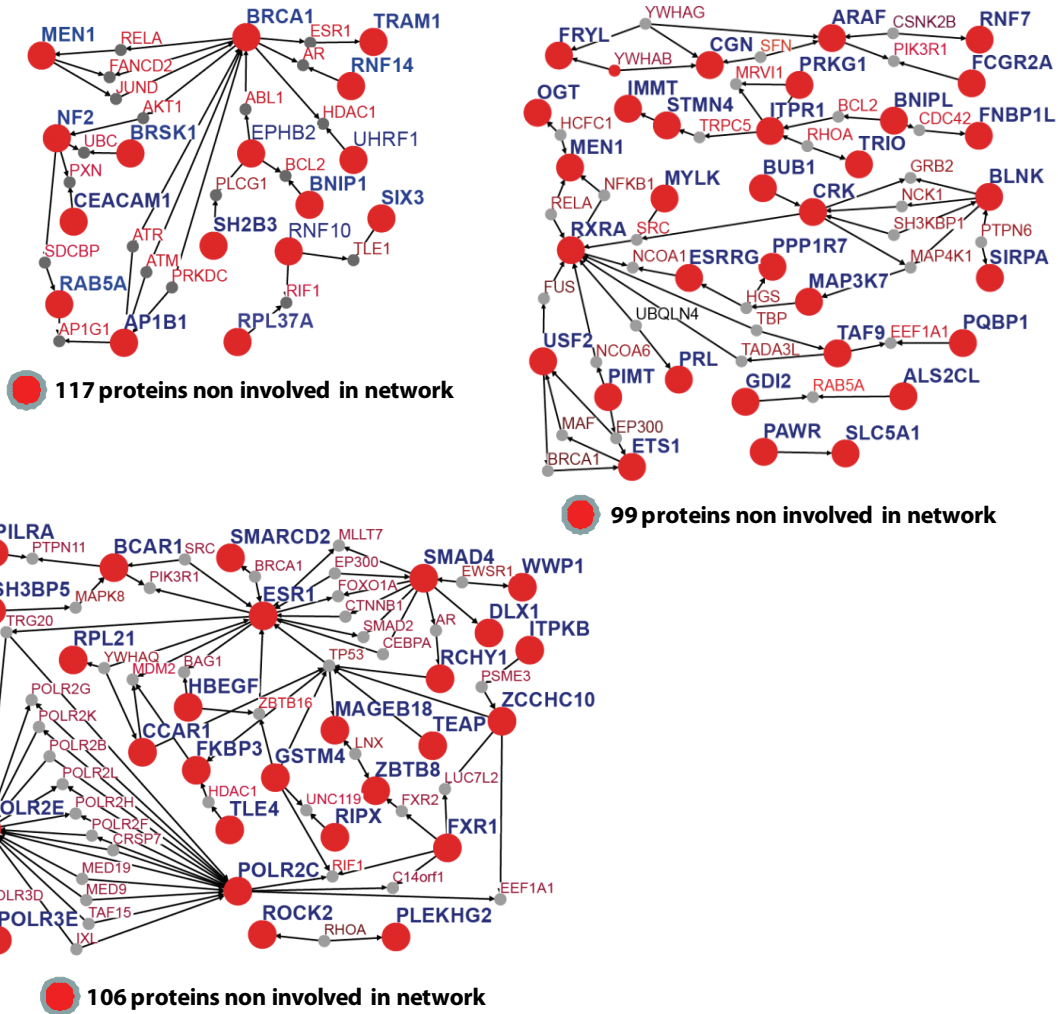
(Dr Pierre Thibault) for mass spectrometry analysis. This work was funded by the Canadian Institutes for Health Research (CIHR) and the Canadian Cancer Society Research institute to TH and a group grant from the Fonds de la Recherche en Santé du Québec to support in part IRIC infrastructure. MCS was supported by a Canada Graduate Scholarship Doctoral Award (CIHR) and a doctoral award from the Cole foundation.

## 4.8 Supplementary data



### Supplemental Figure 4.5 Construction of a c-Kit<sup>+</sup> Lin<sup>-</sup> bone marrow stem/progenitor cell cDNA library

A) Diagram of the library generation protocol. B) Enzymatic digestion (BamHI/EcoRI) of 36 random library plasmidic clones allowing for the release of the cDNA insert (different fragment length) from the pGADT7-Rec vector (8kb). The insertion rate and insert length have been calculated from 60 random library clones.



### Supplemental Figure 4.6 Protein-protein interaction networks involving 133 Osprey-available random proteins

Three different networks have been generated with 3 independent sets of random genes. Only the proteins involved in networks are shown. Proteins from the random list of genes are represented by large red dots and blue characters, while intermediate protein partners are represented by small grey dots.

Gene name	Symbol (s)	Gene name	Symbol (s)
LIM domain binding 1	Ldb1	RKEN cDNA 2700008N14 gene, TBC1 domain family, member 9B	2700008N14RK, Tbc108R
Junonji, AT rich interactive domain 1C (Rbp2 like)	Jaid1c, Smcx	prohibitin 2	Phb2
developmentally regulated GTP binding protein 1	Drg1, Nedd3	tensin like C1 domain-containing phosphatase	Tenc1
cyclin-dependent kinase 9 (CDC2-related kinase)	Cdk9	tubulin, gamma 1	Tubg1
proteasome (prosome, macropain) 28 subunit, 3 (K1 antigen)	Psmc3, K1	profilin 1	Pfn1
MYST histone acetyltransferase 2	Myst2, Hbo1	actin, beta, cytoplasmic	Actb
myeloperoxidase	Mpo	Rac GTPase-activating protein 1	Racgap1
lysyl-tRNA synthetase	Kars	interferon dependent positive acting transcription factor 3 gamma	Isgf3g, Irf9
ribosomal protein L17	Rpl17	exosome component 8	Exosc8
YLP motif containing 1	Ylpm1, zap3	heat shock protein 8	Hspa8
Np95	Uhfr1, np95	chaperonin subunit 7 (eta)	cc17
polymerase (DNA directed), delta 1, catalytic subunit	Pold1	HLA-B-associated transcript 3	Bat3
Amyloid beta precursor protein binding protein2	Appb2, PAT1	guanine nucleotide binding protein, beta 1	Gnb1
lipic acid synthetase	Lias, lip1	Rho GTPase activating protein 8	Arhgap8, Prf5
Rab6 interacting protein 1	Rab6ip1, Dnm5a	(threonyl-tRNA synthetase-like 2	Tars12
retinoic acid induced 1	Rai1, Gt1	ATP synthase, H+ transporting, mitochondrial F0 complex, subunit d	Atp5h
bromodomain adjacent to zinc finger domain 1A	Baz1a, Acf1	integrin linked kinase	Ilk
cytoskeleton associated protein 2-like	Ckap2l	golgi associated, gamma adaptin ear containing, ARF binding protein 1	Gga1
TAF6-like RNA polymerase II, p300/CBP-associated factor (PCAF)-associated factor	Taf6l, Paf65A	c-src tyrosine kinase	Csk
centaurin, alpha 1	Centaur1, Adap1	kinesin family member 23	Kif23, MKLP1
hypoxia up-regulated 1	Hyou1, Grp170	elastase 2, neutrophil	Ela2
RKEN cDNA 1810013C15 gene, ATG2 autophagy related 2 homolog A	1810013C15RK, Ag2a	WW, C2 and coiled-coil domain containing 2	Wwc2
beta-1,4-N-acetyl-galactosaminyl transferase 1	B4galnt1, ggm2	ankyrin repeat and SAM domain containing 1	Anks1
small optic lobes homolog	Solh	similar to signal-induced proliferation-associated 1 like 3; RIKEN cDNA 2610511M17	Sipa13
ATP synthase, H+ transporting, mitochondrial F1 complex, gamma polypeptide 1	Atp5c1	UDP-N-acetylglucosamine pyrophosphorylase 1	Uap1
ribosomal protein S11	Rps11	transmembrane protein 70	Tmem70
minichromosome maintenance deficient 6 (MIS5 homolog S. pombe) (S. cerevisiae)	Mcm6	acetyl-Coenzyme A acetyltransferase 1	Acat1, thil
NADH dehydrogenase (ubiquinone) 1 alpha subcomplex 8	Ndufa8	WW domain containing E3 ubiquitin protein ligase 2	Wwp2
homeodomain interacting protein kinase 2	Hpk2	copper chaperone for superoxide dismutase	Ccs
EGL nine homolog 2 (C. elegans)	Egln2	tubulin, alpha 4A	Tuba4a
SET domain containing (lysine methyltransferase) 8	Setd8	ATPase, Na+/K+ transporting, alpha 1 polypeptide	Atp1a1
general transcription factor IIC, polypeptide 2 beta	Gtf3c2	phosphofructokinase, platelet	Pfkp
cyclin A2	Ccna2, CCn1	actin, gamma, cytoplasmic 1	Actg1
WW domain binding protein 7	Wbp7, MlI2	tubulin, alpha 2	Tuba1b, tuba2
zinc finger protein 106, SH3 domain binding protein 3	Zfp106, sirn	phosphatidylinositol glycan, class S	Pigs
binding uninhibited by benzimidazoles 1 homolog beta (S. cerevisiae)	Bub1b, bubr1	NADH dehydrogenase (ubiquinone) 1 beta subcomplex 4	Ndufb4
protease (prosome, macropain) 26S subunit, ATPase 1	Psmc1, S4	euchromatic histone lysine N-methyltransferase 2	Ehmt2, G9a
MAP kinase kinase kinase 7 interacting protein 2	Map3k7ip2, TAB2	TAF2 RNA polymerase II, TATA box binding protein (TBP)-associated factor	Taf2
leucine-rich repeats and calponin homology domain containing 4	Lrch4, SAP25	similar to ring finger protein 26	Rnf26
WD repeat domain 46	Wdr46, Bing4	nebulin-related anchoring protein	Nrap

Gene name	Symbol (s)	Gene name	Symbol (s)
myosin, light polypeptide 6, alkali, smooth muscle and non-muscle	Myi6	SEC53-like (S. cerevisiae)	Sec63
RIKEN cDNA 2510012.108 gene	2510012.108Rik, cacth	polymerase (DNA-directed), delta interacting protein 3	Poldip3
capping protein (actin filament) muscle Z-line alpha 2	Capza2, cappa2	heat shock protein 9A	Hspa9, mortalin
Myc-induced SUN domain containing protein	Misu, Nsun2	similar to ATP/GTP binding protein 1	Agtpbp1, Nna1
plirlysin metalloprotease 1	Plirm1	RIKEN cDNA 2700083E18 gene	2700083E18Rik, Mex3a
yippee-like 3 (Drosophila)	Ypel3, Suap	ubiquitin associated and SH3 domain containing, A	Ubash3a, Sis-2
flightless 1 homolog (Drosophila)	Fliih, Flii	NADH dehydrogenase (ubiquinone) Fe-S protein 1	Ndufs1
integral membrane protein 2B	Itrn2b	protein tyrosine phosphatase, receptor type, S	Ptprs, PTPsigma
N-acetylneuraminic acid synthase (sialic acid synthase)	Nans	adenylosuccinate synthetase like 1	Adssl1
DNA primase, p49 subunit	Prim1	Rho GDP dissociation inhibitor (GDI) alpha	Arhgdia
Phosphoenolpyruvate carboxykinase	Pck1, PEPCK	nitrilase 1	Nrl1
AT rich interactive domain 1A (Swi1 like)	Arnt1a, Osa1	DEAD (Asp-Glu-Ala-Asp) box polypeptide 49 ; eukaryotic translation initiation factor 4A, isoform 3	Ddx48, eif4a3
DEAD/H (Asp-Glu-Ala-Asp/Hs) box polypeptide 3 X-linked	Ddx3x, Fin14	RIO kinase 2 (yeast)	RioK2
zinc finger and BTB domain containing 17	Zbib17, Miz1	similar to sushi domain containing 1	Susd1
nuclear receptor-binding SET-domain protein 1	Nsd1	hexosaminidase A	Hexa, hex-1
ubiquitin-fold modifier conjugating enzyme 1	Ufc1	eukaryotic translation elongation factor 2	Eef2
ras homolog gene family, member T2	Rhot2, Miro2	superkiller virulicidic activity 2-like	Skiv2l
R3H domain 1 (binds single-stranded nucleic acids)	R3hdm1	glycosyltransferase 25 domain containing 1	Glt25a1
M-phase phosphoprotein 9	Mphosph9	glucosaminyl (N-acetyl) transferase 2, l-branching enzyme	Gent2
TBC1 domain family, member 17	Tbc1d17	RIKEN cDNA 9130227C08Rik gene	9130227C08Rik, khynv
TRAF-interacting protein	Traip	COX11 homolog, cytochrome c oxidase assembly protein	cox11
ubiquitin specific peptidase 15	Usp15	CXXC finger 1 (PHD domain)	Cxxc1, Cgbbp
		phosphoribosylaminoimidazole carboxylase	Paics
TAF4A RNA polymerase II, TATA box binding protein (TBP)-associated factor	Taf4a	plexin B2	Plexnb2
proteasome (prosome, macropain) subunit, alpha type 7	Pasma7	FIP1 like 1 (S. cerevisiae)	Fip1l1
immunoglobulin mu binding protein 2	Ighmbp2	pyruvate kinase, muscle	Pkm2
DEAD (Asp-Glu-Ala-Asp) box polypeptide 39	Ddx39, Bat1	sequence BC017158, protein unknown	BC024080
ankyrin repeat domain 49	Ankrd49, Gbif	collet-coil domain containing 109A	Ccdc109a
		intersectin 2	Itsn2

Table III List of binding partners identified from the LMO2 yeast two hybrid screen

## **5. Discussion**

### **5.1 Librairie d'ADNc de progéniteurs hématopoïétiques : un outil unique et puissant**

Dans le cancer, une petite population de cellules réacquiert certaines propriétés des cellules souches normales et sont appelées cellules souches leucémiques. Par extrapolation, l'étude des cellules souches peut donc permettre de révéler des mécanismes de transformation par les oncogènes. Toutefois, les cellules souches hématopoïétiques se prêtent difficilement aux techniques classiques de biologie moléculaire. Une librairie de cDNA faite à partir de ces cellules représente une façon efficace d'étudier la biologie des cellules souches hématopoïétiques, en identifiant les réseaux d'interactions protéiques qui s'y trouvent.

#### **5.1.1 Cellules souches et progéniteurs hématopoïétiques comme modèle biologique**

L'impact fonctionnel de LMO2, autant dans son rôle transcriptionnel que dans la réplication de l'ADN, dépend du type cellulaire dans lequel il est exprimé. Par exemple, le complexe transcriptionnel SCL/LMO2 se lie au promoteur de *Cdkn1a* et, selon le contexte cellulaire, peut soit activer (dans les érythrocytes [329] et les cellules souches [62]) ou réprimer (dans les mégakaryocytes [464] et les thymocytes [114]) la transcription de ce gène. De la même façon, lorsque LMO2 est surexprimé dans toutes les cellules hématopoïétiques, son activité sur le cycle cellulaire s'observe uniquement dans les progéniteurs thymocytaires. Ainsi, le choix du type cellulaire est d'une importance capitale pour l'étude de la fonction d'une protéine donnée. On peut supposer que LMO2 interagit avec des partenaires différents dans la lignée érythroïde, mégakaryocytaire et thymocytaire, et que ces interactions dictent la fonction de LMO2.

Dans l'hématopoïèse normale, nos observations (non publiées) indiquent que LMO2 est plus fortement exprimé dans les cellules souches hématopoïétiques et dans les progéniteurs multi/oligopotents, suggérant une fonction pour LMO2 dans la

régulation des propriétés des cellules souches hématopoïétiques. D'ailleurs, LMO2 est requis pour la spécification des cellules souches hématopoïétiques durant le développement [31, 55]. De plus, l'expression ectopique de LMO2 dans les cellules T confèrent à celles-ci des propriétés d'auto-renouvellement retrouvées dans les cellules souches hématopoïétiques [114]. L'expression ectopique de LMO2 dans les thymocytes change le transcriptome de ces cellules, en réprimant l'expression de gènes associés à la différenciation des cellules T tout en activant la transcription de gènes normalement exprimés dans les cellules souches hématopoïétiques. Ainsi, LMO2 pourrait transformer les thymocytes en cellules souches leucémiques. Il est donc possible que certains partenaires protéiques de LMO2 soient présents autant dans les cellules souches hématopoïétiques normales que dans les cellules T, influençant ainsi la fonction de LMO2 dans ces cellules. L'identification de partenaires protéiques de LMO2 dans les cellules souches et progéniteurs hématopoïétiques peut nous permettre d'assigner de nouvelles fonctions pour LMO2 dans la régulation de la biologie des cellules souches hématopoïétiques, qui pourraient aussi s'appliquer dans la leucémogénèse. C'est pourquoi nous avons décidé d'identifier les partenaires protéiques de LMO2 dans les cellules souches hématopoïétiques.

L'étude moléculaire des cellules souches hématopoïétiques représente un défi de taille. En effet, les cellules souches hématopoïétiques représentent une population très rare dans la moelle osseuse. Selon les différentes combinaisons de marqueurs extracellulaires utilisés pour la purification, on peut obtenir entre 2 000 et 40 000 cellules souches par souris (la pureté de la population étant inversement proportionnelle au nombre de cellules purifiées). Même si on purifie à la fois les cellules souches et les progéniteurs hématopoïétiques, on ne pourrait obtenir une quantité raisonnable de cellules pour permettre la purification de complexes protéiques par des méthodes biochimiques classiques (immunoprécipitation, purification par affinité).

Normalement, les cellules souches hématopoïétiques ne survivent pas en culture. Toutefois, la transduction de certains gènes, particulièrement *HoxB4*, dans les cellules

souches permet à celles-ci de se multiplier en culture. *HoxB4*, lorsque transduit dans les cellules souches hématopoïétiques par infection rétrovirale, amène leur expansion autant *in vitro* qu'*in vivo*, tout en retenant leur capacité à se différencier normalement vers les lignées myéloïdes et lymphoïdes, sans transformer les cellules [489, 490]. Toutefois, HOXB4 est un facteur de transcription qui se lie sur le promoteur de plusieurs gènes dans les cellules hématopoïétiques [491], et sa surexpression modifie de façon significative le transcriptome de ces cellules, notamment en activant l'expression de *Lmo2*. Ces résultats soutiennent que la surexpression de *HoxB4* modifie la biologie des cellules souches hématopoïétiques, entre autres en augmentant leur fréquence de division et en modifiant l'expression de plusieurs gènes. Ainsi, l'expansion des cellules souches en culture n'est pas nécessairement un choix approprié pour étudier les propriétés intrinsèques des cellules souches dans des conditions normales.

Il existe très peu de modèles de lignées cellulaires qui soient représentatifs des cellules souches hématopoïétiques. D'abord, une caractéristique importante des cellules souches hématopoïétiques est leur faible taux de division, dû à leur état de quiescence. Ainsi, une lignée cellulaire que l'on maintient en culture doit forcément proliférer et se retrouver dans les phases actives du cycle cellulaire, ce qui ne représente pas la majorité des cellules souches hématopoïétiques. Certaines lignées cellulaires (EML, HPC-7, FDCPmix) sont tout de même considérées comme des modèles de cellules souches/progéniteurs, de par leur caractère multipotent. Toutefois, ces cellules sont souvent transduites avec des gènes permettant leur immortalisation, ce qui change leur transcriptome et leurs propriétés biologiques. Ainsi, les lignées cellulaires multipotentes ne sont pas les meilleurs modèles pour étudier la biologie des cellules souches hématopoïétiques.

En résumé, il est très difficile d'obtenir une grande quantité de cellules souches et de progéniteurs hématopoïétiques sans modifier la biologie particulière à ces cellules. Tout dépendant de la question biologique qui est posée, on peut choisir d'utiliser des cellules primaires en petite quantité, d'amplifier les cellules souches en culture ou encore d'utiliser un modèle cellulaire. Évidemment, la combinaison de différentes

approches peut s'avérer un choix judicieux pour valider les résultats obtenus avec une première méthodologie.

### **5.1.2 Librairies de cDNA commercialement disponibles et leurs limitations**

Les interactions protéiques ne peuvent être fonctionnelles que si elles ont lieu dans le tissu qui nous intéresse. Pour cette raison, nous avons décidé de conserver les propriétés normales des cellules souches et des progéniteurs hématopoïétiques en purifiant ces cellules primaires à partir de moelle osseuse de souris. Nous avons créé une librairie de cDNA à partir de ces cellules, qui est représentative du transcriptome des cellules normales, dans des conditions d'homéostasie. De plus, le système de double hybride permet de détecter des interactions plus directes que ce qui est possible d'obtenir avec d'autres techniques de purification protéique, comme l'immunoprécipitation ou la purification par affinité.

La technique de double hybride est souvent critiquée de par son bruit de fond important (nombre élevé de faux positifs). Toutefois, dans le système choisi au laboratoire, la souche de levure utilisée est modifiée pour réduire l'incidence de faux positifs. Cette souche contient trois différents gènes rapporteurs (*ADE2*, *HIS3*, *MEL1*). De plus, l'ajout d'un inhibiteur compétitif de la protéine His3p, le 3-amino-1,2,4-triazole, permet de contrôler la stringence du criblage et de sélectionner uniquement les interactions de haute affinité. Dans ces conditions, mon criblage avec LMO2 comme appât a identifié des interactions protéiques qui ont été confirmées à 100% par des tests indépendants de double hybride, et au moins à 70% par des techniques indépendantes (IP, GST), indiquant que peu de faux positifs se retrouvent dans ma liste de partenaires potentiels.

Il n'existe pas de librairie de cDNA de cellules souches et progéniteurs hématopoïétiques disponible commercialement. Sur le marché, des librairies faites à partir de moelle osseuse totale sont disponibles, ce qui comprend des dizaines de populations hématopoïétiques différentes, dont une minorité possède des propriétés souches. Cet outil unique offre donc des perspectives avantageuses à tout projet



visant à révéler des interactions protéiques gouvernant la biologie des cellules souches et des progéniteurs. De plus, la librairie que j'ai construite se compare, de par ses caractéristiques (nombre de clones indépendants, taille des inserts, taux d'insertion), aux librairies de cDNA disponibles commercialement. Notre librairie possède aussi des caractéristiques qui la distinguent de plusieurs librairies commerciales. D'abord, elle a été construite à partir d'ARNm purifié, et non à partir d'ARN total incluant l'ARN ribosomal et l'ARN de transfert. De plus, le cDNA a été synthétisé avec une combinaison d'oligos dT et d'oligos dégénérés, ce qui génère une librairie plus diversifiée et représentant autant les extrémités 5' et 3' de l'ARNm. La librairie de cDNA créée par notre laboratoire est donc un outil puissant pour la détection d'interactions protéiques.

### **5.1.3 Utilisation de la librairie de cDNA dans l'étude des cellules souches hématopoïétiques**

La librairie de cDNA a été construite à partir de l'ARNm de cellules de moelle osseuse dépourvues des marqueurs de lignage et exprimant le marqueur c-Kit. Cela signifie que les cellules choisies pour créer la librairie sont purifiées selon des critères rigoureux et constituent une population de cellules hautement enrichie en cellules souches et progéniteurs hématopoïétiques. Il n'existe pas de librairie équivalente disponible commercialement pour permettre l'étude des cellules souches. Ainsi, cette librairie unique constitue un atout majeur pour quiconque souhaite étudier les interactions protéine-protéine qui gouvernent la biologie des cellules souches. Dans un futur rapproché, cette librairie pourrait être utilisée par notre équipe ou par d'autres groupes pour effectuer des criblages additionnels de double hybride de levure, en choisissant comme appât d'autres protéines essentielles aux propriétés des cellules souches.

Par exemple, le facteur de transcription c-MYC est un régulateur majeur de la biologie des cellules souches. Alors que l'inactivation de *c-Myc* ne compromet pas la génération de cellules souches hématopoïétiques, celles-ci ne réussissent pas à se différencier en cellules matures, causant un défaut grave d'hématopoïèse définitive

[492]. L'inactivation de *c-Myc* dans la moelle osseuse résulte en la perte progressive de toutes les cellules hématopoïétiques matures [493, 494]. À l'opposé, les cellules souches hématopoïétiques s'accumulent suite à la délétion de *c-Myc* [493, 494]. Certaines voies de signalisation régulent aussi la biologie des cellules souches hématopoïétiques, comme par exemple la voie NOTCH. La corrélation positive entre le nombre d'ostéoblastes et le nombre de cellules souches hématopoïétiques dépend de l'activation de la voie NOTCH [495]. La surexpression d'une forme activée de NOTCH1 dans les cellules souches hématopoïétiques amène l'expansion du réservoir de cellules souches autant *in vitro* qu'*in vivo* [133, 134]. Inversement, l'inhibition de la voie NOTCH par différentes méthodes (surexpression d'un dominant négatif de CBF-1 ou de MAML1, ainsi que traitement avec un inhibiteur de la  $\gamma$ -sécrétase), amène une accélération de la différenciation des cellules souches en culture [496]. La librairie que nous avons créée pourrait être utile dans l'identification de partenaires protéiques de c-MYC ou encore de NOTCH1.

De plus, puisque les cellules souches cancéreuses partagent essentiellement les mêmes propriétés que les cellules souches normales (**Section 1.2.2**), la librairie pourra servir à identifier de nouveaux partenaires protéiques pour des oncogènes puissants impliqués dans les fonctions des cellules souches. D'ailleurs, notre équipe a effectué un criblage d'interaction pour la protéine de fusion leucémogénique NUP98-HOXA9 à l'origine de leucémies myéloïdes aiguës [497] (**Annexe 3**), et nous avons identifié 62 différents partenaires protéiques. Les partenaires ainsi identifiés offrent de nouvelles perspectives dans l'étude de cette protéine de fusion oncogénique. En comparant l'interactome de NUP98-HOXA9 avec celui de LMO2, nous avons identifié 11 protéines en commun (**Annexe 3**). De façon intéressante, la plupart de ces protéines ne jouent pas un rôle connu dans la régulation transcriptionnelle. Cinq de ces interactions ont été confirmées *in vitro* ou par capture de l'ADN (**Chapitre 4**), excluant la possibilité que ces interactions représentent des faux positifs générés par la méthodologie de double hybride. Ces 11 partenaires protéiques communs pourraient être importants pour l'activité leucémogénique de ces facteurs de transcription. Un candidat intéressant est le suppresseur de tumeurs HIPK2 [498]. Cette protéine kinase phosphoryle P53 [499, 500], PML [501] et AML1 [502], et est

parfois mutée dans la leucémie myéloïde aigue [503]. L'interaction entre HIPK2 et LMO2 ou NUP98-HOXA9 mérite d'être confirmée par immunoprécipitation. De plus, il serait intéressant de déterminer si LMO2 et NUP98-HOXA9 sont phosphorylées par HIPK2 *in vitro* et/ou *in vivo*. L'implication de HIPK2 (ou d'autres candidats) dans la leucémogénèse induite par LMO2 ou NUP98-HOXA9 pourrait être testée en transplantation chez la souris, en surexprimant les facteurs de transcription oncogéniques tout en diminuant les niveaux de HIPK2 à l'aide de shRNAs. Ainsi, on pourra déterminer si l'apparition de la leucémie est accélérée ou retardée par la diminution de HIPK2.

Le criblage d'interaction effectué pour SCL et LMO2 prouve que l'emploi de cette librairie est efficace dans l'identification de nouveaux partenaires protéiques. Ainsi, il serait intéressant d'approfondir l'étude de la fonction transcriptionnelle du complexe SCL-LMO2, par exemple en effectuant de nouveaux criblages avec les autres membres du complexe transcriptionnel (E2A, LDB1, ETO2...) comme appât. Les interactions nouvellement identifiées pourraient être confirmées *in vitro* et *in vivo*, et comparées aux nouvelles interactions identifiées par spectrométrie de masse par la technique de capture sur l'ADN décrite dans le **Chapitre 4 (Figure 4.4)**. D'ailleurs, cette technique de capture a permis d'identifier de nouvelles protéines associées aux promoteurs des gènes cibles du complexe SCL-LMO2. Plusieurs de ces protéines n'ont pas de fonction connue dans le processus de transcription. Ainsi, il sera intéressant de pousser l'étude de la régulation transcriptionnelle du complexe SCL-LMO2, qui varie selon le type cellulaire et selon le gène cible (**Section 1.3.4**). Pour ma part, je me suis consacrée à l'étude des nouveaux partenaires de LMO2.

## 5.2 Nouveaux partenaires potentiels de LMO2

Un criblage de double hybride avec LMO2 comme appât a permis d'identifier 139 protéines différentes. Parmi celles-ci, trois partenaires connus de LMO2 ont été identifiés (LDB1, EHMT2 et RBP2). Il semble que LMO2 possède plusieurs partenaires protéiques différents, avec des fonctions diversifiées. Certaines interactions protéiques sont conservées avec d'autres protéines à domaines LIM,

suggérant que les différents membres de cette famille de protéines puissent réguler des fonctions cellulaires semblables.

### **5.2.1 Identification de partenaires connus de LMO2 par double hybride**

Des criblages de double hybride avec LMO2 comme appât ont précédemment été rapportés dans la littérature [345-347]. Toutefois, ces études ont révélé très peu de partenaires protéiques de LMO2. Différentes raisons peuvent expliquer le faible nombre de partenaires identifiés : le tissu d'origine de la librairie utilisée ne permet pas toujours l'expression de partenaires fonctionnels de LMO2 (foie de rat [347], lymphocytes T humains [346], leucémie dérivée de souris *Lmo2<sup>Tg</sup>* [345]) ; la qualité de la librairie n'était peut-être pas optimale ; le nombre de clones criblés était plus faible (environ 10 fois moins de clones criblés que dans notre étude), ce qui peut empêcher l'identification de certains partenaires.

En contrepartie, bien que mon criblage ait détecté l'interaction entre LMO2 et trois de ses partenaires connus (LDB1 [462], RBP2 [346] et EHMT2 [483]), les autres partenaires connus de LMO2 n'ont pas été identifiés. Certaines interactions précédemment rapportées n'ont pas été identifiées dans un contexte hématopoïétique (AF6 [347], HEN1/2 [308, 309], BEX2 [308]), laissant la possibilité que les protéines impliquées ne soient pas exprimées dans les cellules souches et progéniteurs hématopoïétiques. Les autres interactions peuvent avoir été manquées en raison d'une faible abondance du transcrit dans la librairie de cDNA, ou encore à cause de la stringence du criblage qui a sélectionné uniquement des interactions de forte affinité.

### **5.2.2 LMO2 : petite protéine, grand nombre de partenaires**

LMO2 est une toute petite protéine (18kDa) contenant essentiellement deux domaines LIM. Pourtant, le criblage de double hybride a révélé que 139 partenaires différents peuvent s'associer à LMO2. Le taux de faux positifs de ce criblage semble plutôt faible, puisque 70% des interactions testées ont été confirmées *in vitro*, ce qui signifie que la plupart des interactions détectées pourraient avoir un intérêt physiologique. De

plus, un criblage identique avec SCL comme appât a révélé uniquement huit partenaires différents, montrant que le nombre élevé de partenaires identifiés avec LMO2 n'est pas causé par les paramètres de criblage ou par la librairie utilisée. Comment une si petite protéine peut posséder autant de partenaires ?

Premièrement, il est clair que LMO2 ne s'associe pas à toutes ces protéines au même moment. Certaines interactions peuvent avoir lieu dans un type cellulaire particulier, ou même dans une organelle particulière. Effectivement, bien que plus de 60% des nouveaux partenaires identifiés soient des protéines nucléaires, certaines protéines sont cytoplasmiques et même mitochondriales. Il n'est pas exclu que les protéines cytoplasmiques et mitochondriales puissent être transloquées au noyau et s'associer à LMO2. À l'inverse, LMO2 pourrait aussi se retrouver dans le cytoplasme et la mitochondrie. En effet, LMO2 ne possède pas de signal de localisation nucléaire, ni aucun autre motif déterminant sa localisation cellulaire. Cela signifie que LMO2 puisse être transporté dans le noyau ou les autres organelles par le biais d'interactions protéiques. Même si, dans la plupart des types cellulaires étudiés, LMO2 est majoritairement une protéine nucléaire, son expression est tout de même détectée dans le cytoplasme [44]. D'ailleurs, LMO3 et LMO4 peuvent être détectés hors du noyau, où ils interagissent avec des partenaires cytoplasmiques [504, 505]. L'expression de LMO2 dans les autres organelles n'a jamais été investiguée. LMO2 pourrait donc s'associer à différentes protéines en fonction de sa localisation cellulaire. De plus, certaines interactions pourraient avoir lieu uniquement dans un sous-type cellulaire particulier (hématopoïétique ou non, puisque LMO2 est exprimé de façon plutôt ubiquitaire), ou dans des conditions spécifiques.

Deuxièmement, le domaine LIM est une plateforme d'interactions protéiques qui a le potentiel de lier plusieurs motifs différents. Il n'est donc pas surprenant que les partenaires potentiels de LMO2 soient nombreux et variés. D'ailleurs, d'autres protéines à domaines LIM, dont FHL2, ont aussi le potentiel d'interagir avec un nombre impressionnant de partenaires différents. Comme LMO2, la protéine FHL2 est plutôt petite (279 acides aminés, 32kDa) et possède uniquement des domaines LIM (quatre domaines entiers ainsi qu'un doigt de zinc additionnel). FHL2 est

principalement une protéine cardiaque, mais son expression est détectée dans plusieurs autres organes. FHL2 peut être retrouvé autant dans le noyau que le cytoplasme [228, 246, 248]. FHL2 est aussi détecté dans la mitochondrie en essais de transfection transitoire [506]. Jusqu'à maintenant, plus de 50 partenaires différents pour FHL2 ont été rapportés, avec des fonctions très diversifiées (revu dans [484]). Par exemple, comme LMO2, FHL2 peut agir autant comme activateur ou répresseur de la transcription, tout dépendant des nombreux partenaires avec lesquels il s'associe [241, 247, 248, 250, 252, 253, 507, 508]. FHL2 régule aussi l'adhésion cellulaire en s'associant aux intégrines [199, 260]. Par ailleurs, FHL2 s'associe à des enzymes métaboliques telles la phosphofructokinase [258], à des protéines de signalisation comme ERK2 [509], de même qu'à des canaux potassiques [510] et à des métalloprotéases [259]. Ainsi, FHL2 est une protéine très polyvalente qui régule plusieurs fonctions cellulaires en s'associant à des partenaires diverses. LMO2, une protéine de la même famille, pourrait posséder des propriétés similaires.

### **5.2.3 Les protéines à domaines LIM pourraient réguler des processus cellulaires semblables**

Le parallèle entre LMO2 et FHL2 suggère que les différentes protéines à domaines LIM pourraient s'associer à des partenaires protéiques semblables afin de moduler les mêmes fonctions protéiques. Par exemple, plusieurs protéines à domaines LIM, incluant LMO2, ont un rôle bien caractérisé dans la transcription. Plusieurs interactions ont été observées entre les protéines à domaines LIM et les bHLH : LMO2/SCL [316, 317] ; LMO2-4/HEN1-2 [289, 308, 309] ; LMO4/neurogenin2 [307] ; LMX1/E47 [511] ; FHL2/HAND1 [246] ; FHL3/MyoD [256] ; CRP3/MyoD, myogenin [235]. De la même façon, des protéines à domaines LIM de différentes sous-familles (LMO2/4 [314, 315, 318], CRP1/2 [512]) peuvent interagir avec les membres de la famille GATA. On peut donc émettre l'hypothèse que les domaines LIM possèdent des potentiels d'interaction protéique semblables. Ainsi, trois nouveaux partenaires de LMO2 sont connus pour interagir, parfois *via* les mêmes domaines protéiques, avec d'autres protéines à domaines LIM, tel que résumé dans le **Tableau IV** (page suivante).

<b>Partenaire protéique</b>	<b>Résidus (aa) du partenaire responsables de l'interaction</b>	<b>Protéine à domaine LIM</b>
HIPK2	391-574	LMO2
	1-565	ZYXIN
	812-907	FHL2
ILK	9-207	LMO2
	1-163	PINCH
CSK	269-451	LMO2
	ND	PAXILLIN

**Tableau IV. Partenaires potentiels de LMO2 interagissant avec d'autres protéines à domaines LIM**

Premièrement, HIPK2, une kinase qui phosphoryle p53, interagit avec FHL2 et ZYXIN, en plus d'être un partenaire potentiel de LMO2. L'interaction FHL2-HIPK2 stimule l'activité transcriptionnelle de p53 [513], alors que l'interaction ZYXIN-HIPK2 stabilise la protéine HIPK2, avec un impact sur la réponse aux dommages à l'ADN dépendante de p53 [215]. Deuxièmement, ILK, une kinase impliquée entre autres dans l'adhésion cellulaire, interagit avec PINCH [202]. Enfin, CSK est une tyrosine kinase qui phosphoryle PAXILLIN [514].

Par ses nouvelles interactions protéiques, LMO2 est associé à des nouvelles fonctions cellulaires, dont certaines sont partagées par d'autres protéines à domaines LIM. Par exemple, plusieurs nouveaux partenaires de LMO2 sont directement impliqués dans les processus d'adhésion focale et de régulation du cytosquelette, pour lesquels plusieurs protéines à domaines LIM jouent un rôle important. De la même façon, bien qu'un rôle dans la réplication de l'ADN n'ait pas été décrit pour d'autres protéines à domaines LIM, FHL2 interagit avec MCM7 en essais de double hybride [515].

#### **5.2.4 Possibilités de fonctions non-transcriptionnelles pour LMO2**

Dans les cellules érythroïdes, environ 70% du pool nucléaire de LMO2 est co-précipité avec son partenaire transcriptionnel SCL [177], suggérant que la fonction principale de LMO2 dans ces cellules soit au sein de complexes transcriptionnels.

Cependant, une quantité non négligeable de LMO2 n'est pas associée à SCL, suggérant que LMO2 puisse posséder des fonctions indépendantes. De plus, il est possible que cette proportion varie selon le type cellulaire analysé, et que la fonction transcriptionnelle de LMO2 soit moins importante dans d'autres cellules.

Dans la littérature, la seule fonction connue de LMO2 est son rôle dans l'échafaudage de complexes transcriptionnels. Il est donc peu surprenant que près du tiers (31%) des partenaires potentiels de LMO2 identifiés par double hybride soient impliqués dans la transcription et/ou le remodelage de la chromatine. De façon surprenante, une proportion importante des protéines identifiées sont impliquées dans d'autres processus cellulaires. Il est possible que parmi ces protéines, certaines possèdent un rôle transcriptionnel qui demeure à être caractérisé. Par exemple, la protéine lysyl-tRNA-synthetase a été identifiée comme partenaire potentiel de LMO2. Bien que le rôle le mieux caractérisé de cette protéine soit dans le chargement des ARN de transfert importants pour la traduction, un rôle transcriptionnel a été décrit pour cette protéine, de par son interaction avec les facteurs de transcription MITF [516] et USF2 [517] dans les mastocytes.

Cependant, plusieurs partenaires potentiels de LMO2 peuvent être regroupés selon des fonctions cellulaires communes (analyse d'enrichissement par « *GO terms* »), ce qui suggère que LMO2 pourrait réguler ces fonctions non-transcriptionnelles. Nous en avons fait la preuve en montrant que LMO2 régule la réplication de l'ADN dans les cellules hématopoïétiques. D'ailleurs, une récente analyse génomique de la liaison de LMO2 à l'ADN par ChIP-seq a identifié près de 10 000 sites de liaison, dont seulement 263 sont situés dans les promoteurs des gènes [339]. Ces observations supportent l'hypothèse selon laquelle LMO2 pourrait posséder d'autres fonctions dépendantes de l'ADN, en plus de son rôle transcriptionnel. La liste impressionnante des partenaires de LMO2 présentée dans cette thèse ouvre plusieurs perspectives quant aux différentes fonctions possibles pour cette protéine oncogénique. L'implication potentielle de LMO2 dans certaines fonctions cellulaires est discutée brièvement dans les sections suivantes.



#### **5.2.4.1 Lien potentiel entre LMO2 et les protéines liant l'ARN**

La transcription est souvent couplée spatio-temporellement à la maturation de l'ARNm dans le noyau et possiblement à sa traduction dans le cytoplasme [518, 519]. Ainsi, LMO2 pourrait favoriser l'expression de ses gènes cibles en recrutant des complexes protéiques impliquant des protéines qui s'associent à l'ARN de même que des protéines ribosomales. L'interaction entre LMO2 et les protéines ribosomales, ou d'autres protéines liant l'ARN, pourrait être vérifiée par co-immunoprécipitation. De plus, il serait intéressant de déterminer si LMO2 s'associe aux polysomes, c'est-à-dire aux ribosomes liés à l'ARNm, par exemple en purifiant les polysomes par gradient de sucrose et en analysant la présence de LMO2 par immunobuvardage. Par ailleurs, des recherches récentes ont démontré l'importance des longs ARN non codants intergéniques (lincRNAs) dans la régulation de l'expression génique [520-523]. En s'associant aux protéines liant l'ARN, LMO2 pourrait intervenir dans le recrutement de tels ARNs aux promoteurs de ses différents gènes cibles et moduler la transcription de cette façon, en plus de sa participation dans des complexes transcriptionnels. Afin de déterminer si LMO2 lie des ARN non codants, on peut faire une immunoprécipitation avec un anticorps contre LMO2, suivi d'une purification de l'ARN lié à LMO2 et d'une analyse sur des micropuces à ARN. Ces différentes techniques permettraient d'éclaircir le rôle de LMO2 dans les processus dépendants de l'ARN.

#### **5.2.4.2 Lien potentiel entre LMO2 et les protéines du cytosquelette**

La possibilité que LMO2 régule l'adhésion cellulaire ou les fonctions du cytosquelette a précédemment été évoquée dans cette thèse, parce que plusieurs partenaires potentiels de LMO2 sont impliqués dans ces fonctions, et puisque d'autres protéines à domaines LIM sont connues pour leur rôle dans ces processus. L'adhésion des cellules souches hématopoïétiques à leur niche est un déterminant majeur dans la biologie des cellules souches, entre autres en régulant leur quiescence. LMO2 pourrait réguler l'adhésion et/ou la motilité des cellules souches et des progéniteurs hématopoïétiques de par son interaction avec des protéines de cytosquelette. Dans un premier temps, il faudrait confirmer les interactions identifiées en double hybride

entre LMO2 et les protéines du cytosquelette. Plusieurs expériences complexes permettraient de déterminer si LMO2 influence l'adhésion des cellules souches à leur niche. Par exemple, on pourrait utiliser les souris *Lmo2<sup>fl/fl</sup>* générées par McCormack et al [81] pour générer des souris inactivées pour *Lmo2* spécifiquement dans les cellules souches hématopoïétiques. L'adhésion, la mobilisation ou la déplétion du réservoir de cellules souches hématopoïétiques pourrait être visualisée *in vivo* par microscopie. Plus simplement, la présence de LMO2 dans les plaques d'adhésion focale pourrait aussi être investiguée par microscopie dans des lignées hématopoïétiques ou des cellules primaires.

#### **5.2.4.3 Lien potentiel entre LMO2 et les protéines mitochondriales**

LMO2 s'associe aussi à plusieurs protéines mitochondriales, dont certaines participent à la chaîne respiratoire. Le lien entre LMO2, la mitochondrie et les cellules hématopoïétiques n'a pas du tout été élucidé. Toutefois, les cellules souches hématopoïétiques résident dans un milieu hypoxique, et les niveaux d'oxygène contrôlent, du moins en partie, les propriétés des cellules souches hématopoïétiques. Par exemple, les cellules souches en milieu hypoxique sont plus compétitives quant à leur pouvoir de repopulation d'un hôte immunodéficient suite à une transplantation [22, 24]. De plus, la capacité de reconstitution des cellules souches hématopoïétiques est compromise lorsque ces cellules accumulent des ROS (*reactive oxygen species*), par exemple suite à l'inactivation d'ATM (*ataxia telangiectasia mutated*) [23] ou des membres de la famille de facteurs de transcription FoxO [25, 524]. Ainsi, la chaîne respiratoire de la mitochondrie pourrait réguler l'activité des cellules souches hématopoïétiques. Afin de déterminer si LMO2 s'associe aux protéines mitochondriales, des expériences de co-immunoprécipitation pourraient être effectuées, et la présence de LMO2 dans les mitochondries pourrait être détectée par microscopie. Notre équipe a déjà démontré que les souris *Scf<sup>tg</sup>Lmo1<sup>tg</sup>* présentaient une accumulation de ROS dans les thymocytes (Tremblay M et Hoang T, données non publiées). Il reste à déterminer si la régulation des niveaux d'oxygène dans la cellule par SCL et LMO se fait par régulation transcriptionnelle, ou possiblement *via* une association de LMO2 aux protéines mitochondriales. Il serait aussi intéressant de

déterminer si LMO2 influence le niveau de ROS dans les cellules souches hématopoïétiques, en plus de son effet sur les thymocytes.

En résumé, il est possible que le rôle de LMO2 dans l'hématopoïèse et la leucémogénèse ne soit pas totalement dû à sa participation dans des complexes transcriptionnels. Le nouveau rôle pour LMO2 dans la réplication de l'ADN en est la preuve.

### **5.3 LMO2 comme régulateur de la réplication de l'ADN**

Mes résultats indiquent que LMO2 régule directement la réplication de l'ADN dans les cellules hématopoïétiques. Le mécanisme précis selon lequel LMO2 influence ce processus n'a toutefois pas été complètement élucidé. LMO2 pourrait intervenir au niveau de l'assemblage du complexe répliatif de pré-initiation, ou encore influencer le choix des origines de réplication à être activées. Une autre alternative est que LMO2 modifie l'état de la chromatine aux origines de réplication. Ces possibilités sont discutées dans les prochaines sections.

#### **5.3.1 LMO2 pourrait faciliter l'assemblage du complexe de pré-initiation de la réplication de l'ADN**

Mon criblage de double hybride indique que LMO2 interagit directement avec MCM6, PRIM1 et POLD1. Alors que MCM6 se retrouve à la fois dans le complexe de pré-réplication et dans le complexe de pré-initiation de la réplication, POLD1 et PRIM1 sont des enzymes faisant partie spécifiquement du complexe de pré-initiation. Il est donc possible que LMO2 puisse intervenir dans la transition entre les complexes de pré-réplication et de pré-initiation. LMO2 pourrait se lier aux origines de réplication en s'associant à MCM6, et servir par la suite de plateforme de recrutement pour d'autres protéines comme PRIM1 et POLD1. Bien que l'interaction avec MCM6 stabilise le niveau protéique de LMO2, l'effet de LMO2 sur la stabilité du complexe de pré-initiation n'a pas été caractérisé. Une façon simple de tester cette hypothèse serait d'employer des shRNAs contre LMO2 dans les cellules hématopoïétiques, et de mesurer le recrutement de POLD1 et de PRIM1 aux origines

par CHIP ou par immunobuvardage sur des extraits protéiques de chromatine. D'autre part, alors que cette nouvelle activité de LMO2 dans la réplication de l'ADN semble indépendante de son association avec SCL et GATA, l'implication de LDB1 n'a pas vraiment été investiguée. Des shRNAs contre LDB1 pourraient être employés pour déterminer si LDB1 est nécessaire à la fonction de LMO2 dans la réplication de l'ADN.

### **5.3.2 LMO2 pourrait influencer le choix des origines de réplication au moment de leur activation**

Au niveau mécanistique, bien que LMO2 influence la vitesse d'élongation des fourches de réplication, il est possible que LMO2 influence aussi le nombre d'origines actives dans la cellule. En effet, chez les mammifères, une minorité d'origines de réplication est constitutivement active, à chaque cycle cellulaire, alors qu'une large proportion d'origines est dormante ou flexible [358, 392]. Les facteurs régulant l'activité des origines n'ont pas encore été identifiés, et il est possible que certains facteurs de transcription puissent jouer ce rôle. De par sa faible expression et sa courte demi-vie [177], LMO2 pourrait être une protéine limitante qui influencerait l'activation des origines dans le contexte spécifique de la différenciation hématopoïétique. Ainsi, il aurait été intéressant de quantifier le nombre d'origines actives dans des cellules exprimant différents niveaux protéiques de LMO2 (ex. cellules MEL exprimant ou non des *shLMO2*). D'autre part, certaines origines sont activées en début de phase S (« *early origins* » ou origines précoces), alors que d'autres sont activées plus tardivement. Puisque les origines précoces sont souvent retrouvées près des régions activement transcrites, on pourrait émettre l'hypothèse que les facteurs de transcription comme LMO2 influenceraient de façon préférentielle l'activité de ce sous-type particulier d'origines. Afin de déterminer quel type d'origine est régulé par LMO2, il faudrait d'abord connaître l'ordre d'activation des origines dans les cellules hématopoïétiques, ce qui est complexe. Il serait possible de corrélérer le recrutement de LMO2 à quelques origines avec l'ordre d'activation de ces origines, comme cela a été fait pour la protéine RB [434]. Néanmoins, ce type

d'expérience permettrait uniquement d'établir une corrélation, et non de valider fonctionnellement l'effet de LMO2 sur ces origines.

### **5.3.3 LMO2 pourrait modifier l'état de la chromatine aux origines de réplication**

L'état de la chromatine aux origines de réplication influence fort probablement leur activité. Certaines modifications d'histones permettent un relâchement de la chromatine, ce qui favorise le recrutement de protéines sur l'ADN. Plusieurs origines de réplication contiennent des marques de chromatine active. Il est possible que l'effet de LMO2 sur la réplication de l'ADN soit dû à sa capacité de modifier le statut de la chromatine aux origines de réplication, ce qui favoriserait l'assemblage des complexes de pré-initiation sur l'ADN. D'ailleurs, le criblage de double hybride nous a permis d'identifier plusieurs modificateurs de la chromatine en tant que partenaires potentiels de LMO2, dont certains sont connus pour réguler la réplication de l'ADN (HBO1, SETD8, ACF1, G9A).

Par exemple, chez les mammifères, les origines de réplication sont acétylées sur l'histone H4 par l'enzyme HBO1 (MYST2) [403]. En s'associant aux protéines ORC1 [525] et MCM2 [526], HBO1 favorise l'assemblage des protéines MCM aux origines de réplication [527], et l'activité acétyltransférase de HBO1 est essentielle à cette fonction [403]. D'autre part, un rôle pour HBO1 dans l'érythropoïèse a récemment été décrit [528]. L'histone méthyltransférase SETD8 (PR-SET7) a aussi été associée à un rôle dans la réplication de l'ADN. SETD8 s'associe à PCNA [529, 530] de même qu'aux foci de réplication sur l'ADN [531], et sa déplétion perturbe la progression en phase S des cellules [529-531], notamment en modifiant la vitesse d'élongation des fourches de réplication et le nombre d'origines actives [531]. La mono-méthylation de l'histone H4 par SETD8 facilite le recrutement du complexe de pré-initiation aux origines de réplication [404]. Un autre exemple est ACF1 (BAZ1A), membre du « *chromatin assembly complex* », qui est requis pour la réplication de l'hétérochromatine. La déplétion de ACF1 cause un délai dans la progression tardive en phase S, causé par un défaut de réplication de

l'hétérochromatine [532]. De façon plus indirecte, l'histone méthyltransférase G9A (EHMT2) est aussi colocalisée aux sites de réplication sur l'ADN, et est retrouvée en complexe protéique avec PCNA durant la phase S [533].

Le lien entre LMO2, la réplication de l'ADN et toutes ces enzymes de remodelage de la chromatine n'a pas été adressé jusqu'à présent. Dans un premier temps, il faudrait vérifier l'interaction entre LMO2 et ces protéines par immunoprécipitation. Il serait aussi possible de déterminer si les origines liées par LMO2 par ChIP sont aussi liées par ces protéines, ou encore si les origines liées par LMO2 présentent des modifications d'histones propres à ces enzymes de remodelage. Par la suite, il serait intéressant de déterminer si LMO2 favorise le recrutement des enzymes de remodelage aux origines ou si, inversement, ce sont les enzymes de remodelage qui favorisent le recrutement de LMO2. Cela pourrait être testé à l'aide de shRNAs contre LMO2 et contre les enzymes de remodelage, en mesurant le recrutement des protéines aux origines par ChIP.

#### **5.3.4 Autres possibilités**

En plus des protéines MCM6, DNA polymérase et DNA primase, le criblage de double hybride a aussi permis l'identification de protéines additionnelles dont un rôle dans la réplication de l'ADN a été décrit. Par exemple, la protéine NP95 (UHRF1) est colocalisée avec PCNA en début de phase S [534, 535], et son expression est nécessaire à la fois pour l'entrée en phase S [536] et pour la réplication de l'hétérochromatine [537]. NP95 recrute l'enzyme DNMT1 sur l'ADN hémi-méthylé et est nécessaire à la maintenance des marques de méthylation sur l'ADN durant la réplication [538]. De façon intéressante, NP95 est spécifiquement exprimée durant la phase S dans des thymocytes normaux, mais son expression demeure anormalement élevée dans les lymphomes de souris [539]. La protéine POLDIP3 (PDIP46) a été identifiée dans mon criblage. Bien qu'un rôle pour cette protéine dans la réplication de l'ADN n'ait pas été caractérisé, POLDIP3 a d'abord été identifiée comme partenaire de l'ADN polymérase delta dans un criblage de double hybride [540], suggérant qu'elle puisse influencer la fonction de la polymérase durant la réplication.

Bien entendu, le lien entre LMO2, ces partenaires potentiels et la réplication de l'ADN reste à clarifier. L'interaction entre LMO2 et ces protéines devrait être vérifiée par immunoprécipitation. La présence de ces protéines aux origines de réplication pourrait être déterminée par CHIP. L'impact de ces protéines sur la réplication de l'ADN dans les cellules hématopoïétiques pourrait être mesuré à l'aide de shRNAs contre ces protéines.

## **5.4 Lien entre transcription, réplication et cancer**

Les facteurs de transcription oncogéniques perturbent souvent les programmes transcriptionnels qui gouvernent le cycle cellulaire et la réplication de l'ADN. De plus, tel que décrit dans les **Chapitres 2 et 3**, certains de ces facteurs de transcription, dont LMO2, régulent directement la réplication de l'ADN. Ainsi, certains oncogènes induisent un stress répliatif dans la cellule. Une réplication aberrante de l'ADN peut entraîner de l'instabilité génomique, un état favorisant l'apparition de mutations qui ultimement peuvent transformer la cellule. D'ailleurs, l'expression de certains gènes de réplication de l'ADN est dérégulée dans différents cancers.

### **5.4.1 Les facteurs de transcription qui influencent la réplication de l'ADN**

Bien entendu, la plupart des facteurs de transcription qui influencent la réplication de l'ADN le font d'abord *via* leur effet transcriptionnel, par exemple en régulant l'expression des protéines de réplication et de cycle cellulaire. Toutefois, il semble que certains facteurs de transcription influencent plus directement la réplication de l'ADN, en s'associant aux protéines de réplication, tel que discuté dans le **Chapitre 2**. Plus particulièrement, les exemples de c-MYC, HOXD13 (**Chapitre 2**) et LMO2 (**Chapitre 3**) sont les plus convaincants. De façon intéressante, ces facteurs de transcription ne semblent pas influencer la réplication de l'ADN par les mêmes mécanismes. Alors que c-MYC [416] et HOXD13 [439] s'associent au complexe de pré-réplication (comprenant les protéines CDC6 et CDT1), LMO2 ne s'associe pas à CDC6 mais semble plutôt s'associer au complexe de pré-initiation de la réplication, comprenant l'ADN polymérase et la primase. Cela suggère que c-MYC et HOXD13

pourraient jouer un rôle dans la mise en place des protéines de réplication aux origines, alors que LMO2 pourrait être plus impliqué dans l'activation des origines. D'ailleurs, la surexpression de HOXD13 favorise l'association des protéines du complexe de pré-réplication aux origines [439]. Dans le même ordre d'idées, la déplétion de *c-Myc* [416], contrairement à celle de *Lmo2*, n'influence pas la vitesse d'élongation des fourches de réplication. Cela met en évidence le fait que les facteurs de transcription peuvent réguler la réplication de l'ADN par des mécanismes distincts. Ceci dit, est-ce que les facteurs de transcription régulent directement la réplication de l'ADN ?

Il est bien connu que les facteurs de transcription peuvent recruter des enzymes de remodelage de la chromatine aux promoteurs de leurs gènes cibles. Ces mêmes enzymes sont aussi recrutées aux origines de réplication. On ne peut pas exclure la possibilité que les facteurs de transcription influencent la réplication de l'ADN en recrutant des enzymes de remodelage de la chromatine, ce qui permettrait de relâcher la chromatine aux origines pour permettre l'accessibilité de la machinerie de réplication sur l'ADN. D'ailleurs, tel que mentionné précédemment, mon criblage de double hybride a identifié deux partenaires potentiels de LMO2 (HBO1 et SETD8) qui modifient la chromatine et influencent la réplication de l'ADN [403, 527, 529-531]. Ainsi, il est possible que l'association entre les facteurs de transcription et les enzymes de remodelage de la chromatine puisse être importante pour la régulation de la réplication de l'ADN. Ceci dit, l'association de c-MYC [416], HOXD13 [439] et LMO2 avec plusieurs protéines de réplication de l'ADN suggère que ces facteurs de transcription puissent aussi être plus directement impliqués dans l'assemblage du complexe de réplication ou dans l'activation des origines. Ces différents mécanismes ne sont pas mutuellement exclusifs.

Tel que proposé dans le **Chapitre 2**, il est aussi probable que les facteurs de transcription influencent seulement un sous-type particulier d'origines. D'ailleurs, nos expériences de ChIP démontrent que LMO2 est recruté sur moins de 50% des origines de réplication testées. Les facteurs de transcription ne sont pas essentiels à la réplication de l'ADN, mais pourraient faciliter l'activation des origines dans des



contextes spécifiques. Puisque le nombre d'origines activées est inversement proportionnel à la longueur de la phase S de la cellule, on peut émettre l'hypothèse que certains facteurs de transcription pourraient faciliter l'activation d'origines additionnelles lorsque le cycle cellulaire doit être court, par exemple lors d'un stress, ou lorsqu'une prolifération cellulaire rapide est nécessaire. Particulièrement, l'hématopoïèse est un processus dynamique qui requiert parfois une division rapide et massive des cellules. Le fait que LMO2, c-MYC et HOXD13 soient tous des facteurs de transcription exprimés dans le système hématopoïétique n'est peut-être pas un hasard. Afin de déterminer si ces facteurs de transcription influencent la réplication à des origines particulières, il faudrait d'abord connaître l'inventaire des origines de réplication chez les mammifères, ce qui semble être une tâche très ardue, et ensuite classer les origines par sous-type : constitutives, dormantes, précoces, tardives, etc. En résumé, l'implication des facteurs de transcription dans la réplication de l'ADN est un domaine émergent où beaucoup de questions doivent être élucidées. Puisque plusieurs facteurs de transcription, dont LMO2 [30, 135], c-MYC [451] et HOXD13 [450], sont des oncogènes puissants, leur rôle dans la réplication de l'ADN doit être mieux décrit, puisque la réplication de l'ADN est souvent perturbée dans différents cancers, tel que discuté dans la section suivante.

#### **5.4.2 Réplication de l'ADN et instabilité génomique**

Une réplication de l'ADN aberrante peut causer de l'instabilité génomique qui favorise l'émergence de cancers. Plusieurs oncogènes, lorsque surexprimés, perturbent la réplication de l'ADN, ce qui favorise la transformation de la cellule.

Plusieurs différents oncogènes (*Mos*, *Cdc6*, *Cyclin E*, *Ras*), lorsque surexprimés, induisent une terminaison prématurée des fourches de réplication, ce qui engendre la réponse aux dommages à l'ADN et subséquentement cause la sénescence des cellules [412, 413]. La présence d'oncogènes amène une augmentation dans le nombre d'origines de réplication actives et engendre une asymétrie dans la progression des fourches de réplication de part et d'autre de l'origine, ce qui pourrait causer une instabilité et une terminaison prématurée de la réplication [412]. Dans des cellules où

la réponse aux dommages à l'ADN est inhibée (déplétion ou inhibition de ATM, déplétion de CHK2), la sénescence est supprimée et la transformation cellulaire est plus efficace [412, 413]. Ainsi, les oncogènes amènent une réplication aberrante de l'ADN, contre laquelle la cellule a développé des mécanismes de protection, entre autres en activant la réponse aux dommages à l'ADN qui engendre la sénescence ou l'apoptose. Lorsque ces mécanismes de protection sont défectueux ou absents, une réplication aberrante peut provoquer de l'instabilité génomique, permettant l'accumulation de mutations favorisant la transformation cellulaire [453, 541, 542].

On peut donc penser que des altérations génétiques qui activent l'action des protéines de réplication pourraient être retrouvées dans différents types de cancers. Chez l'humain, *CDC6* et *CDT1* sont surexprimés dans environ 50% des cas de cancer du poumon [543], et la surexpression de ces gènes induit une ré-réplication qui provoque de l'instabilité génomique en l'absence de P53 [544]. Des anticorps contre *CDC6* et *MCM5* reconnaissent spécifiquement, et avec une grande sensibilité, des cellules de cancer de l'utérus, comparativement à des frottis de cellules utérines saines [545]. Les protéines de la famille MCM sont fréquemment employées comme biomarqueurs dans différents cancers, et leur expression est souvent corrélée à un mauvais pronostic [546].

Au même titre qu'une réplication excessive, une réplication déficiente semble aussi générer de l'instabilité génomique. En effet, lorsque trop peu de complexes de pré-réplication sont formés, moins d'origines sont activées, ce qui peut engendrer une réplication incomplète durant la phase S et ainsi causer des cassures d'ADN et des altérations chromosomiques [454, 455, 547-550]. En accord avec ces observations, un allèle hypomorphe de *Mcm4* prédispose les cellules aux bris chromosomiques et favorise l'apparition d'adénocarcinomes chez la souris [456]. De la même façon, la déplétion conditionnelle de *Mcm2* compromet la survie à long terme des souris en favorisant l'émergence de cancers, particulièrement de lymphomes [457]. De façon plus indirecte, la surexpression de *CYCLIN E*, un oncogène associé à plusieurs cancers, compromet l'assemblage des protéines MCM aux origines de réplication [551], ce qui pourrait expliquer l'instabilité génomique associée à l'expression de

*CYCLINE* [552]. Ainsi, la duplication de l'ADN à chaque division cellulaire doit être étroitement surveillée, et lorsque celle-ci se fait dans des conditions sous-optimales, l'apparition d'erreurs ou de dommages peut mener à des altérations génétiques favorisant l'émergence du cancer.

Dans le contexte de cette thèse, il serait intéressant de déterminer si la surexpression de LMO2 dans les cellules hématopoïtiques cause des dommages à l'ADN. Ceci pourrait être vérifié par microscopie ou par immunobuvardage contre des protéines de réponse aux dommages à l'ADN : phosphorylation de  $\gamma$ H2AX, ATM, ATR, CHK1 ou CHK2. Il est possible que la sénescence des thymocytes observée lors de la surexpression de LMO et de SCL (Tremblay M et Hoang T, données non publiées) soit causée par l'induction de dommages à l'ADN par ces oncogènes. Dans un contexte où l'intégrité génomique est perturbée, la leucémie induite par LMO et SCL apparaît plus rapidement (Tremblay M et Hoang T, données non publiées). Il est donc possible qu'en plus de son effet transcriptionnel, LMO2 induise la leucémie en perturbant la réplication de l'ADN et en causant de l'instabilité génomique, favorisant l'apparition de mutations additionnelles qui accélèrent la transformation de la cellule.

## Conclusion

L'objectif de cette thèse était d'attribuer de nouvelles fonctions pour le facteur de transcription oncogénique LMO2, à partir de l'identification de nouveaux partenaires protéiques. Ces nouvelles fonctions peuvent être pertinentes autant dans le contexte de l'hématopoïèse normale que dans la leucémie induite par LMO2. À cet effet, la librairie de cDNA de cellules souches et progéniteurs hématopoïétiques constitue un outil spécifique et pertinent pour quiconque veut identifier des interactions protéine-protéine gouvernant la biologie des cellules souches.

Mes recherches ont établi que LMO2 avait le potentiel d'interagir avec de nombreux partenaires protéiques aux fonctions très diversifiées. Plusieurs de ces interactions peuvent être informatives quant au rôle de LMO2 dans la régulation de ses différents gènes cibles. En outre, des fonctions non-transcriptionnelles ont été proposées pour LMO2, mettant en évidence son caractère multifonctionnel. Un de ces nouveaux rôles a été caractérisé plus en profondeur, soit l'implication de LMO2 dans la réplication de l'ADN. L'étude des nouvelles fonctions potentielles de LMO2 pourrait contribuer à notre compréhension du mode de fonctionnement de cet oncogène, et possiblement à celui des autres protéines LMO qui pourraient partager les mêmes partenaires protéiques. De plus, nous avons établi que certains partenaires de LMO2 s'associent à d'autres oncogènes, comme SCL ou NUP98-HOXA9, suggérant que différents facteurs de transcription oncogéniques pourraient transformer la cellule par des mécanismes moléculaires communs.

Ainsi, l'ensemble des travaux présentés dans cette thèse met en lumière l'importance des interactions protéine-protéine dans l'identification de nouvelles fonctions pour des protéines connues, par exemple pour des facteurs de transcription comme LMO2.

## Bibliographie

1. Weissman, I.L., *Stem cells: units of development, units of regeneration, and units in evolution*. Cell, 2000. **100**(1): p. 157-68.
2. Mikkola, H.K. and S.H. Orkin, *The journey of developing hematopoietic stem cells*. Development, 2006. **133**(19): p. 3733-44.
3. Boisset, J.C., et al., *In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium*. Nature, 2010. **464**(7285): p. 116-20.
4. Cheshier, S.H., et al., *In vivo proliferation and cell cycle kinetics of long-term self-renewing hematopoietic stem cells*. Proc Natl Acad Sci U S A, 1999. **96**(6): p. 3120-5.
5. Kiel, M.J., et al., *Haematopoietic stem cells do not asymmetrically segregate chromosomes or retain BrdU*. Nature, 2007. **449**(7159): p. 238-42.
6. Wilson, A., et al., *Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair*. Cell, 2008. **135**(6): p. 1118-29.
7. Takizawa, H., et al., *Dynamic variation in cycling of hematopoietic stem cells in steady state and inflammation*. J Exp Med, 2011. **208**(2): p. 273-84.
8. Foudi, A., et al., *Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells*. Nat Biotechnol, 2009. **27**(1): p. 84-90.
9. Morrison, S.J. and J. Kimble, *Asymmetric and symmetric stem-cell divisions in development and cancer*. Nature, 2006. **441**(7097): p. 1068-74.
10. Wu, M., et al., *Imaging hematopoietic precursor division in real time*. Cell Stem Cell, 2007. **1**(5): p. 541-54.
11. Yamashita, Y.M., D.L. Jones, and M.T. Fuller, *Orientation of asymmetric stem cell division by the APC tumor suppressor and centrosome*. Science, 2003. **301**(5639): p. 1547-50.
12. Beckmann, J., et al., *Asymmetric cell division within the human hematopoietic stem and progenitor cell compartment: identification of asymmetrically segregating proteins*. Blood, 2007. **109**(12): p. 5494-501.
13. Walker, M.R., K.K. Patel, and T.S. Stappenbeck, *The stem cell niche*. J Pathol, 2009. **217**(2): p. 169-80.
14. Zhang, J., et al., *Identification of the haematopoietic stem cell niche and control of the niche size*. Nature, 2003. **425**(6960): p. 836-41.
15. Nilsson, S.K., H.M. Johnston, and J.A. Coverdale, *Spatial localization of transplanted hemopoietic stem cells: inferences for the localization of stem cell niches*. Blood, 2001. **97**(8): p. 2293-9.
16. Kiel, M.J., et al., *SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells*. Cell, 2005. **121**(7): p. 1109-21.
17. Askenasy, N., et al., *Transplanted hematopoietic cells seed in clusters in recipient bone marrow in vivo*. Stem Cells, 2002. **20**(4): p. 301-10.
18. Sipkins, D.A., et al., *In vivo imaging of specialized bone marrow endothelial microdomains for tumour engraftment*. Nature, 2005. **435**(7044): p. 969-73.

19. Levesque, J.P., F.M. Helwani, and I.G. Winkler, *The endosteal 'osteoblastic' niche and its role in hematopoietic stem cell homing and mobilization*. *Leukemia*, 2010. **24**(12): p. 1979-92.
20. Arai, F., et al., *Tie2/angiopoietin-1 signaling regulates hematopoietic stem cell quiescence in the bone marrow niche*. *Cell*, 2004. **118**(2): p. 149-61.
21. Ceradini, D.J., et al., *Progenitor cell trafficking is regulated by hypoxic gradients through HIF-1 induction of SDF-1*. *Nat Med*, 2004. **10**(8): p. 858-64.
22. Danet, G.H., et al., *Expansion of human SCID-repopulating cells under hypoxic conditions*. *J Clin Invest*, 2003. **112**(1): p. 126-35.
23. Ito, K., et al., *Regulation of oxidative stress by ATM is required for self-renewal of haematopoietic stem cells*. *Nature*, 2004. **431**(7011): p. 997-1002.
24. Winkler, I.G., et al., *Positioning of bone marrow hematopoietic and stromal cells relative to blood flow in vivo: serially reconstituting hematopoietic stem cells reside in distinct nonperfused niches*. *Blood*, 2010. **116**(3): p. 375-85.
25. Tothova, Z., et al., *FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress*. *Cell*, 2007. **128**(2): p. 325-39.
26. Kollet, O., et al., *Osteoclasts degrade endosteal components and promote mobilization of hematopoietic progenitor cells*. *Nat Med*, 2006. **12**(6): p. 657-64.
27. Mendez-Ferrer, S., et al., *Haematopoietic stem cell release is regulated by circadian oscillations*. *Nature*, 2008. **452**(7186): p. 442-7.
28. Katayama, Y., et al., *Signals from the sympathetic nervous system regulate hematopoietic stem cell egress from bone marrow*. *Cell*, 2006. **124**(2): p. 407-21.
29. Manaia, A., et al., *Lmo2 and GATA-3 associated expression in intraembryonic hemogenic sites*. *Development*, 2000. **127**(3): p. 643-53.
30. Royer-Pokora, B., U. Loos, and W.D. Ludwig, *TTG-2, a new gene encoding a cysteine-rich protein with the LIM motif, is overexpressed in acute T-cell leukaemia with the t(11;14)(p13;q11)*. *Oncogene*, 1991. **6**(10): p. 1887-93.
31. Warren, A.J., et al., *The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development*. *Cell*, 1994. **78**(1): p. 45-57.
32. Begley, C.G., et al., *The gene SCL is expressed during early hematopoiesis and encodes a differentiation-related DNA-binding motif*. *Proc Natl Acad Sci U S A*, 1989. **86**(24): p. 10128-32.
33. Kallianpur, A.R., J.E. Jordan, and S.J. Brandt, *The SCL/TAL-1 gene is expressed in progenitors of both the hematopoietic and vascular systems during embryogenesis*. *Blood*, 1994. **83**(5): p. 1200-8.
34. Elefanty, A.G., et al., *SCL expression in the mouse embryo detected with a targeted lacZ reporter gene demonstrates its localization to hematopoietic, vascular, and neural tissues*. *Blood*, 1999. **94**(11): p. 3754-63.
35. Akashi, K., et al., *Transcriptional accessibility for genes of multiple tissues and hematopoietic lineages is hierarchically controlled during early hematopoiesis*. *Blood*, 2003. **101**(2): p. 383-9.
36. Ramalho-Santos, M., et al., *"Stemness": transcriptional profiling of embryonic and adult stem cells*. *Science*, 2002. **298**(5593): p. 597-600.

37. Herblot, S., et al., *SCL and LMO1 alter thymocyte differentiation: inhibition of E2A-HEB function and pre-T alpha chain expression*. Nat Immunol, 2000. **1**(2): p. 138-44.
38. Green, A.R., et al., *SCL is coexpressed with GATA-1 in hemopoietic cells but is also expressed in developing brain*. Oncogene, 1992. **7**(4): p. 653-60.
39. Hoang, T., et al., *Opposing effects of the basic helix-loop-helix transcription factor SCL on erythroid and monocytic differentiation*. Blood, 1996. **87**(1): p. 102-11.
40. Natkunam, Y., et al., *The oncoprotein LMO2 is expressed in normal germinal-center B cells and in human B-cell lymphomas*. Blood, 2007. **109**(4): p. 1636-42.
41. Yamada, Y., et al., *The oncogenic LIM-only transcription factor Lmo2 regulates angiogenesis but not vasculogenesis in mice*. Proc Natl Acad Sci U S A, 2000. **97**(1): p. 320-4.
42. Hinks, G.L., et al., *Expression of LIM protein genes Lmo1, Lmo2, and Lmo3 in adult mouse hippocampus and other forebrain regions: differential regulation by seizure activity*. J Neurosci, 1997. **17**(14): p. 5549-59.
43. Herberth, B., et al., *SCL, GATA-2 and Lmo2 expression in neurogenesis*. Int J Dev Neurosci, 2005. **23**(5): p. 449-63.
44. Neale, G.A., et al., *Expression of the proto-oncogene rhombotin-2 is identical to the acute phase response protein metallothionein, suggesting multiple functions*. Cell Growth Differ, 1995. **6**(5): p. 587-96.
45. van Eekelen, J.A., et al., *Expression pattern of the stem cell leukaemia gene in the CNS of the embryonic and adult mouse*. Neuroscience, 2003. **122**(2): p. 421-36.
46. Foroni, L., et al., *The rhombotin gene family encode related LIM-domain proteins whose differing expression suggests multiple roles in mouse development*. J Mol Biol, 1992. **226**(3): p. 747-61.
47. Ciau-Uitz, A., F. Liu, and R. Patient, *Genetic control of hematopoietic development in Xenopus and zebrafish*. Int J Dev Biol, 2010. **54**(6-7): p. 1139-49.
48. Mead, P.E., et al., *Primitive erythropoiesis in the Xenopus embryo: the synergistic role of LMO-2, SCL and GATA-binding proteins*. Development, 2001. **128**(12): p. 2301-8.
49. Patterson, L.J., M. Gering, and R. Patient, *Scl is required for dorsal aorta as well as blood formation in zebrafish embryos*. Blood, 2005. **105**(9): p. 3502-11.
50. Patterson, L.J., et al., *The transcription factors Scl and Lmo2 act together during development of the hemangioblast in zebrafish*. Blood, 2007. **109**(6): p. 2389-98.
51. Robb, L., et al., *Absence of yolk sac hematopoiesis from mice with a targeted disruption of the scl gene*. Proc Natl Acad Sci U S A, 1995. **92**(15): p. 7075-9.
52. Shivdasani, R.A., E.L. Mayer, and S.H. Orkin, *Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL*. Nature, 1995. **373**(6513): p. 432-4.
53. Robb, L., et al., *The scl gene product is required for the generation of all hematopoietic lineages in the adult mouse*. Embo J, 1996. **15**(16): p. 4123-9.

54. Porcher, C., et al., *The T cell leukemia oncoprotein SCL/tal-1 is essential for development of all hematopoietic lineages*. Cell, 1996. **86**(1): p. 47-57.
55. Yamada, Y., et al., *The T cell leukemia LIM protein Lmo2 is necessary for adult mouse hematopoiesis*. Proc Natl Acad Sci U S A, 1998. **95**(7): p. 3890-5.
56. Schlaeger, T.M., et al., *Tie2Cre-mediated gene ablation defines the stem-cell leukemia gene (SCL/tal1)-dependent window during hematopoietic stem-cell development*. Blood, 2005. **105**(10): p. 3871-4.
57. Lancrin, C., et al., *Blood cell generation from the hemangioblast*. J Mol Med, 2010. **88**(2): p. 167-72.
58. Visvader, J.E., Y. Fujiwara, and S.H. Orkin, *Unsuspected role for the T-cell leukemia protein SCL/tal-1 in vascular development*. Genes Dev, 1998. **12**(4): p. 473-9.
59. Yamada, Y., et al., *The LIM-domain protein Lmo2 is a key regulator of tumour angiogenesis: a new anti-angiogenesis drug target*. Oncogene, 2002. **21**(9): p. 1309-15.
60. Lazrak, M., et al., *The bHLH TAL-1/SCL regulates endothelial cell migration and morphogenesis*. J Cell Sci, 2004. **117**(Pt 7): p. 1161-71.
61. Herblot, S., P.D. Aplan, and T. Hoang, *Gradient of E2A activity in B-cell development*. Mol Cell Biol, 2002. **22**(3): p. 886-900.
62. Lacombe, J., et al., *Scl regulates the quiescence and the long-term competence of hematopoietic stem cells*. Blood, 2010. **115**(4): p. 792-803.
63. Mikkola, H.K., et al., *Haematopoietic stem cells retain long-term repopulating activity and multipotency in the absence of stem-cell leukaemia SCL/tal-1 gene*. Nature, 2003. **421**(6922): p. 547-51.
64. Reynaud, D., et al., *SCL/TAL1 expression level regulates human hematopoietic stem cell self-renewal and engraftment*. Blood, 2005. **106**(7): p. 2318-28.
65. Curtis, D.J., et al., *SCL is required for normal function of short-term repopulating hematopoietic stem cells*. Blood, 2004. **103**(9): p. 3342-8.
66. Brunet de la Grange, P., et al., *Low SCL/TAL1 expression reveals its major role in adult hematopoietic myeloid progenitors and stem cells*. Blood, 2006. **108**(9): p. 2998-3004.
67. Souroullas, G.P., et al., *Adult hematopoietic stem and progenitor cells require either Lyl1 or Scl for survival*. Cell Stem Cell, 2009. **4**(2): p. 180-6.
68. Hall, M.A., et al., *Functional but abnormal adult erythropoiesis in the absence of the stem cell leukemia gene*. Mol Cell Biol, 2005. **25**(15): p. 6355-62.
69. Hall, M.A., et al., *The critical regulator of embryonic hematopoiesis, SCL, is vital in the adult for megakaryopoiesis, erythropoiesis, and lineage choice in CFU-S12*. Proc Natl Acad Sci U S A, 2003. **100**(3): p. 992-7.
70. McCormack, M.P., et al., *A critical role for the transcription factor Scl in platelet production during stress thrombopoiesis*. Blood, 2006. **108**(7): p. 2248-56.
71. Salmon, J.M., et al., *Aberrant mast-cell differentiation in mice lacking the stem-cell leukemia gene*. Blood, 2007. **110**(10): p. 3573-81.



72. Dey, S., et al., *The TAL1/SCL transcription factor regulates cell cycle progression and proliferation in differentiating murine bone marrow monocyte precursors*. Mol Cell Biol, 2010. **30**(9): p. 2181-92.
73. Aplan, P.D., et al., *The SCL gene product: a positive regulator of erythroid differentiation*. Embo J, 1992. **11**(11): p. 4073-81.
74. Hansson, A., et al., *The Lim-only protein LMO2 acts as a positive regulator of erythroid differentiation*. Biochem Biophys Res Commun, 2007. **364**(3): p. 675-81.
75. Visvader, J.E., et al., *The LIM-domain binding protein Ldb1 and its partner LMO2 act as negative regulators of erythroid differentiation*. Proc Natl Acad Sci U S A, 1997. **94**(25): p. 13707-12.
76. Terano, T., et al., *Transcriptional control of fetal liver hematopoiesis: dominant negative effect of the overexpression of the LIM domain mutants of LMO2*. Exp Hematol, 2005. **33**(6): p. 641-51.
77. Anguita, E., et al., *Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2*. Embo J, 2004. **23**(14): p. 2841-52.
78. Song, S.H., C. Hou, and A. Dean, *A positive role for NLI/Ldb1 in long-range beta-globin locus control region function*. Mol Cell, 2007. **28**(5): p. 810-22.
79. Lahlil, R., et al., *SCL assembles a multifactorial complex that determines glycophorin A expression*. Mol Cell Biol, 2004. **24**(4): p. 1439-52.
80. Xu, Z., et al., *Identification of a TAL1 target gene reveals a positive role for the LIM domain-binding protein Ldb1 in erythroid gene expression and differentiation*. Mol Cell Biol, 2003. **23**(21): p. 7585-99.
81. McCormack, M.P., et al., *The LMO2 T-cell oncogene is activated via chromosomal translocations or retroviral insertion during gene therapy but has no mandatory role in normal T-cell development*. Mol Cell Biol, 2003. **23**(24): p. 9003-13.
82. Grutz, G., A. Forster, and T.H. Rabbitts, *Identification of the LMO4 gene encoding an interaction partner of the LIM-binding protein LDB1/NLII: a candidate for displacement by LMO proteins in T cell acute leukaemia*. Oncogene, 1998. **17**(21): p. 2799-803.
83. Michell, A.C., et al., *A novel role for transcription factor Lmo4 in thymus development through genetic interaction with Cited2*. Dev Dyn, 2010. **239**(7): p. 1988-94.
84. Goardon, N., et al., *Ectopic expression of TAL-1 protein in Ly-6E.1-htal-1 transgenic mice induces defects in B- and T-lymphoid differentiation*. Blood, 2002. **100**(2): p. 491-500.
85. Hahn, W.C., et al., *Creation of human tumour cells with defined genetic elements*. Nature, 1999. **400**(6743): p. 464-8.
86. Vogelstein, B. and K.W. Kinzler, *The multistep nature of cancer*. Trends Genet, 1993. **9**(4): p. 138-41.
87. Rubnitz, J.E. and C.H. Pui, *Recent advances in the treatment and understanding of childhood acute lymphoblastic leukaemia*. Cancer Treat Rev, 2003. **29**(1): p. 31-44.

88. Neale, G.A., J.E. Rehg, and R.M. Goorha, *Ectopic expression of rhombotin-2 causes selective expansion of CD4-CD8- lymphocytes in the thymus and T-cell tumors in transgenic mice*. *Blood*, 1995. **86**(8): p. 3060-71.
89. Treanor, L.M., et al., *Functional interactions between Lmo2, the Arf tumor suppressor, and Notch1 in murine T-cell malignancies*. *Blood*, 2011. **117**(20): p. 5453-62.
90. Petzer, A.L. and E. Gunsilius, *Hematopoietic stem cells in chronic myeloid leukemia*. *Arch Med Res*, 2003. **34**(6): p. 496-506.
91. Marculescu, R., et al., *V(D)J-mediated translocations in lymphoid neoplasms: a functional assessment of genomic instability by cryptic sites*. *J Exp Med*, 2002. **195**(1): p. 85-98.
92. Biernaux, C., et al., *Detection of major bcr-abl gene expression at a very low level in blood cells of some healthy individuals*. *Blood*, 1995. **86**(8): p. 3118-22.
93. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer*. *Cell*, 2000. **100**(1): p. 57-70.
94. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation*. *Cell*, 2011. **144**(5): p. 646-74.
95. Soussi, T., *p53 alterations in human cancer: more questions than answers*. *Oncogene*, 2007. **26**(15): p. 2145-56.
96. Zhou, B.P. and M.C. Hung, *Dysregulation of cellular signaling by HER2/neu in breast cancer*. *Semin Oncol*, 2003. **30**(5 Suppl 16): p. 38-48.
97. Karnoub, A.E. and R.A. Weinberg, *Ras oncogenes: split personalities*. *Nat Rev Mol Cell Biol*, 2008. **9**(7): p. 517-31.
98. Bachman, K.E. and B.H. Park, *Dual nature of TGF-beta signaling: tumor suppressor vs. tumor promoter*. *Curr Opin Oncol*, 2005. **17**(1): p. 49-54.
99. Stevaux, O. and N.J. Dyson, *A revised picture of the E2F transcriptional network and RB function*. *Curr Opin Cell Biol*, 2002. **14**(6): p. 684-91.
100. Canepa, E.T., et al., *INK4 proteins, a family of mammalian CDK inhibitors with novel biological functions*. *IUBMB Life*, 2007. **59**(7): p. 419-26.
101. Giacinti, C. and A. Giordano, *RB and cell cycle progression*. *Oncogene*, 2006. **25**(38): p. 5220-7.
102. Chowdhury, I., B. Tharakan, and G.K. Bhat, *Current concepts in apoptosis: the physiological suicide program revisited*. *Cell Mol Biol Lett*, 2006. **11**(4): p. 506-25.
103. Miyashita, T. and J.C. Reed, *Tumor suppressor p53 is a direct transcriptional activator of the human bax gene*. *Cell*, 1995. **80**(2): p. 293-9.
104. Cory, S., D.C. Huang, and J.M. Adams, *The Bcl-2 family: roles in cell survival and oncogenesis*. *Oncogene*, 2003. **22**(53): p. 8590-607.
105. Serrano, M., et al., *Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a*. *Cell*, 1997. **88**(5): p. 593-602.
106. Dhomen, N., et al., *Oncogenic Braf induces melanocyte senescence and melanoma in mice*. *Cancer Cell*, 2009. **15**(4): p. 294-303.
107. Lazzarini Denchi, E., et al., *Deregulated E2F activity induces hyperplasia and senescence-like features in the mouse pituitary gland*. *Mol Cell Biol*, 2005. **25**(7): p. 2660-72.

108. Shay, J.W. and S. Bacchetti, *A survey of telomerase activity in human cancer*. Eur J Cancer, 1997. **33**(5): p. 787-91.
109. Bonnet, D. and J.E. Dick, *Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell*. Nat Med, 1997. **3**(7): p. 730-7.
110. Al-Hajj, M., et al., *Prospective identification of tumorigenic breast cancer cells*. Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3983-8.
111. O'Brien, C.A., et al., *A human colon cancer cell capable of initiating tumour growth in immunodeficient mice*. Nature, 2007. **445**(7123): p. 106-10.
112. Singh, S.K., et al., *Identification of human brain tumour initiating cells*. Nature, 2004. **432**(7015): p. 396-401.
113. Clevers, H., *The cancer stem cell: premises, promises and challenges*. Nat Med, 2011. **17**(3): p. 313-9.
114. McCormack, M.P., et al., *The Lmo2 oncogene initiates leukemia in mice by inducing thymocyte self-renewal*. Science, 2010. **327**(5967): p. 879-83.
115. De Keersmaecker, K., P. Marynen, and J. Cools, *Genetic insights in the pathogenesis of T-cell acute lymphoblastic leukemia*. Haematologica, 2005. **90**(8): p. 1116-27.
116. Sullivan, C., et al., *Targeted therapy of chronic myeloid leukemia*. Biochem Pharmacol, 2010. **80**(5): p. 584-91.
117. Jurcic, J.G., S.L. Soignet, and A.P. Maslak, *Diagnosis and treatment of acute promyelocytic leukemia*. Curr Oncol Rep, 2007. **9**(5): p. 337-44.
118. Graux, C., et al., *Cytogenetics and molecular genetics of T-cell acute lymphoblastic leukemia: from thymocyte to lymphoblast*. Leukemia, 2006. **20**(9): p. 1496-510.
119. Teitell, M.A. and P.P. Pandolfi, *Molecular genetics of acute lymphoblastic leukemia*. Annu Rev Pathol, 2009. **4**: p. 175-98.
120. Martens, J.H. and H.G. Stunnenberg, *The molecular signature of oncofusion proteins in acute myeloid leukemia*. FEBS Lett, 2010. **584**(12): p. 2662-9.
121. Ferrando, A.A., et al., *Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia*. Cancer Cell, 2002. **1**(1): p. 75-87.
122. Tremblay, M., et al., *Modeling T-cell acute lymphoblastic leukemia induced by the SCL and LMO1 oncogenes*. Genes Dev, 2010. **24**(11): p. 1093-105.
123. Ellisen, L.W., et al., *TAN-1, the human homolog of the Drosophila notch gene, is broken by chromosomal translocations in T lymphoblastic neoplasms*. Cell, 1991. **66**(4): p. 649-61.
124. Weng, A.P., et al., *Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia*. Science, 2004. **306**(5694): p. 269-71.
125. Lee, S.Y., et al., *Mutations of the Notch1 gene in T-cell acute lymphoblastic leukemia: analysis in adults and children*. Leukemia, 2005. **19**(10): p. 1841-3.
126. Mansour, M.R., et al., *High incidence of Notch-1 mutations in adult patients with T-cell acute lymphoblastic leukemia*. Leukemia, 2006. **20**(3): p. 537-9.
127. O'Neil, J., et al., *Activating Notch1 mutations in mouse models of T-ALL*. Blood, 2006. **107**(2): p. 781-5.

128. Maillard, I., T. Fang, and W.S. Pear, *Regulation of lymphoid development, differentiation, and function by the Notch pathway*. *Annu Rev Immunol*, 2005. **23**: p. 945-74.
129. Pear, W.S., et al., *Exclusive development of T cell neoplasms in mice transplanted with bone marrow expressing activated Notch alleles*. *J Exp Med*, 1996. **183**(5): p. 2283-91.
130. Pui, J.C., et al., *Notch1 expression in early lymphopoiesis influences B versus T lineage determination*. *Immunity*, 1999. **11**(3): p. 299-308.
131. Deftos, M.L., et al., *Notch1 signaling promotes the maturation of CD4 and CD8 SP thymocytes*. *Immunity*, 2000. **13**(1): p. 73-84.
132. Robey, E., et al., *An activated form of Notch influences the choice between CD4 and CD8 T cell lineages*. *Cell*, 1996. **87**(3): p. 483-92.
133. Stier, S., et al., *Notch1 activation increases hematopoietic stem cell self-renewal in vivo and favors lymphoid over myeloid lineage outcome*. *Blood*, 2002. **99**(7): p. 2369-78.
134. Varnum-Finney, B., et al., *Pluripotent, cytokine-dependent, hematopoietic stem cells are immortalized by constitutive Notch1 signaling*. *Nat Med*, 2000. **6**(11): p. 1278-81.
135. Boehm, T., et al., *The rhombotin family of cysteine-rich LIM-domain oncogenes: distinct members are involved in T-cell translocations to human chromosomes 11p15 and 11p13*. *Proc Natl Acad Sci U S A*, 1991. **88**(10): p. 4367-71.
136. Begley, C.G., et al., *Chromosomal translocation in a human leukemic stem-cell line disrupts the T-cell antigen receptor delta-chain diversity region and results in a previously unreported fusion transcript*. *Proc Natl Acad Sci U S A*, 1989. **86**(6): p. 2031-5.
137. Chen, Q., et al., *The tal gene undergoes chromosome translocation in T cell leukemia and potentially encodes a helix-loop-helix protein*. *Embo J*, 1990. **9**(2): p. 415-24.
138. Finger, L.R., et al., *Involvement of the TCL5 gene on human chromosome 1 in T-cell leukemia and melanoma*. *Proc Natl Acad Sci U S A*, 1989. **86**(13): p. 5039-43.
139. Raghavan, S.C., I.R. Kirsch, and M.R. Lieber, *Analysis of the V(D)J recombination efficiency at lymphoid chromosomal translocation breakpoints*. *J Biol Chem*, 2001. **276**(31): p. 29126-33.
140. Aplan, P.D., et al., *Disruption of the human SCL locus by "illegitimate" V-(D)-J recombinase activity*. *Science*, 1990. **250**(4986): p. 1426-9.
141. Van Vlierberghe, P., et al., *The cryptic chromosomal deletion del(11)(p12p13) as a new activation mechanism of LMO2 in pediatric T-cell acute lymphoblastic leukemia*. *Blood*, 2006. **108**(10): p. 3520-9.
142. Kovanen, P.E. and W.J. Leonard, *Cytokines and immunodeficiency diseases: critical roles of the gamma(c)-dependent cytokines interleukins 2, 4, 7, 9, 15, and 21, and their signaling pathways*. *Immunol Rev*, 2004. **202**: p. 67-83.
143. Hacein-Bey-Abina, S., et al., *LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1*. *Science*, 2003. **302**(5644): p. 415-9.

144. Hacein-Bey-Abina, S., et al., *Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1*. J Clin Invest, 2008. **118**(9): p. 3132-42.
145. Howe, S.J., et al., *Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients*. J Clin Invest, 2008. **118**(9): p. 3143-50.
146. Bash, R.O., et al., *Does activation of the TAL1 gene occur in a majority of patients with T-cell acute lymphoblastic leukemia? A pediatric oncology group study*. Blood, 1995. **86**(2): p. 666-76.
147. Ferrando, A.A., et al., *Biallelic transcriptional activation of oncogenic transcription factors in T-cell acute lymphoblastic leukemia*. Blood, 2004. **103**(5): p. 1909-11.
148. Larson, R.C., et al., *Protein dimerization between Lmo2 (Rbtn2) and Tall alters thymocyte development and potentiates T cell tumorigenesis in transgenic mice*. Embo J, 1996. **15**(5): p. 1021-7.
149. Robb, L., et al., *scl, a gene frequently activated in human T cell leukaemia, does not induce lymphomas in transgenic mice*. Oncogene, 1995. **10**(1): p. 205-9.
150. Aplan, P.D., et al., *An scl gene product lacking the transactivation domain induces bony abnormalities and cooperates with LMO1 to generate T-cell malignancies in transgenic mice*. Embo J, 1997. **16**(9): p. 2408-19.
151. Kelliher, M.A., D.C. Seldin, and P. Leder, *Tal-1 induces T cell acute lymphoblastic leukemia accelerated by casein kinase IIalpha*. Embo J, 1996. **15**(19): p. 5160-6.
152. Condorelli, G.L., et al., *T-cell-directed TAL-1 expression induces T-cell malignancies in transgenic mice*. Cancer Res, 1996. **56**(22): p. 5113-9.
153. Fisch, P., et al., *T-cell acute lymphoblastic lymphoma induced in transgenic mice by the RBTN1 and RBTN2 LIM-domain genes*. Oncogene, 1992. **7**(12): p. 2389-97.
154. Larson, R.C., et al., *T cell tumours of disparate phenotype in mice transgenic for Rbtn-2*. Oncogene, 1994. **9**(12): p. 3675-81.
155. Larson, R.C., et al., *The oncogenic LIM protein Rbtn2 causes thymic developmental aberrations that precede malignancy in transgenic mice*. Oncogene, 1995. **11**(5): p. 853-62.
156. Elwood, N.J., et al., *SCL, the gene implicated in human T-cell leukaemia, is oncogenic in a murine T-lymphocyte cell line*. Oncogene, 1993. **8**(11): p. 3093-101.
157. Curtis, D.J., et al., *The CD2-scl transgene alters the phenotype and frequency of T-lymphomas in N-ras transgenic or p53 deficient mice*. Oncogene, 1997. **15**(24): p. 2975-83.
158. Royer-Pokora, B., et al., *The TTG-2/RBTN2 T cell oncogene encodes two alternative transcripts from two promoters: the distal promoter is removed by most 11p13 translocations in acute T cell leukaemia's (T-ALL)*. Oncogene, 1995. **10**(7): p. 1353-60.
159. Crable, S.C. and K.P. Anderson, *A PAR domain transcription factor is involved in the expression from a hematopoietic-specific promoter for the human LMO2 gene*. Blood, 2003. **101**(12): p. 4757-64.

160. Landry, J.R., et al., *Fli1, Elf1, and Ets1 regulate the proximal promoter of the LMO2 gene in endothelial cells*. *Blood*, 2005. **106**(8): p. 2680-7.
161. Landry, J.R., et al., *Expression of the leukemia oncogene Lmo2 is controlled by an array of tissue-specific elements dispersed over 100 kb and bound by Tall/Lmo2, Ets, and Gata factors*. *Blood*, 2009. **113**(23): p. 5783-92.
162. Zhu, H., et al., *Regulation of the lmo2 promoter during hematopoietic and vascular development in zebrafish*. *Dev Biol*, 2005. **281**(2): p. 256-69.
163. Tsunoda, T., et al., *Immune-related zinc finger gene ZFAT is an essential transcriptional regulator for hematopoietic differentiation in blood islands*. *Proc Natl Acad Sci U S A*, 2010. **107**(32): p. 14199-204.
164. Bianchi, E., et al., *c-myb supports erythropoiesis through the transactivation of KLF1 and LMO2 expression*. *Blood*, 2010. **116**(22): p. e99-110.
165. Hammond, S.M., S.C. Crable, and K.P. Anderson, *Negative regulatory elements are present in the human LMO2 oncogene and may contribute to its expression in leukemia*. *Leuk Res*, 2005. **29**(1): p. 89-97.
166. Zhu, T., G. Qin, and B. Royer-Pokora, *A novel post-transcriptional splicing form of the acute T cell leukemia proto-oncogene Lmo2*. *Sci China C Life Sci*, 2001. **44**(6): p. 561-9.
167. Wang, Q., et al., *A novel transcript of the LMO2 gene, LMO2-c, is regulated by GATA-1 and PU.1 and encodes an antagonist of LMO2*. *Leukemia*, 2007. **21**(5): p. 1015-25.
168. Oram, S.H., et al., *A previously unrecognized promoter of LMO2 forms part of a transcriptional regulatory circuit mediating LMO2 expression in a subset of T-acute lymphoblastic leukaemia patients*. *Oncogene*, 2010. **29**(43): p. 5796-808.
169. Yuan, J.Y., et al., *MicroRNA-223 reversibly regulates erythroid and megakaryocytic differentiation of K562 cells*. *J Cell Mol Med*, 2009. **13**(11-12): p. 4551-9.
170. Felli, N., et al., *MicroRNA 223-dependent expression of LMO2 regulates normal erythropoiesis*. *Haematologica*, 2009. **94**(4): p. 479-86.
171. Malumbres, R., et al., *Differentiation stage-specific expression of microRNAs in B lymphocytes and diffuse large B-cell lymphomas*. *Blood*, 2009. **113**(16): p. 3754-64.
172. Zhang, J., et al., *Patterns of microRNA expression characterize stages of human B-cell differentiation*. *Blood*, 2009. **113**(19): p. 4586-94.
173. Biryukova, I., et al., *Drosophila mir-9a regulates wing development via fine-tuning expression of the LIM only factor, dLMO*. *Dev Biol*, 2009. **327**(2): p. 487-96.
174. Bazzoni, F., et al., *Induction and regulatory function of miR-9 in human monocytes and neutrophils exposed to proinflammatory signals*. *Proc Natl Acad Sci U S A*, 2009. **106**(13): p. 5282-7.
175. Monticelli, S., et al., *MicroRNA profiling of the murine hematopoietic system*. *Genome Biol*, 2005. **6**(8): p. R71.
176. Yuva-Aydemir, Y., et al., *MicroRNA-9: Functional evolution of a conserved small regulatory RNA*. *RNA Biol*, 2011. **8**(4).

177. Lecuyer, E., et al., *Protein stability and transcription factor complex assembly determined by the SCL-LMO2 interaction*. J Biol Chem, 2007. **282**(46): p. 33649-58.
178. Ostendorff, H.P., et al., *Ubiquitination-dependent cofactor exchange on LIM homeodomain transcription factors*. Nature, 2002. **416**(6876): p. 99-103.
179. Xu, Z., et al., *Single-stranded DNA-binding proteins regulate the abundance of LIM domain and LIM domain-binding proteins*. Genes Dev, 2007. **21**(8): p. 942-55.
180. Perez-Alvarado, G.C., et al., *Structure of the carboxy-terminal LIM domain from the cysteine rich protein CRP*. Nat Struct Biol, 1994. **1**(6): p. 388-98.
181. Perez-Alvarado, G.C., et al., *Structure of the cysteine-rich intestinal protein, CRIP*. J Mol Biol, 1996. **257**(1): p. 153-74.
182. Schmeichel, K.L. and M.C. Beckerle, *The LIM domain is a modular protein-binding interface*. Cell, 1994. **79**(2): p. 211-9.
183. Feuerstein, R., et al., *The LIM/double zinc-finger motif functions as a protein dimerization domain*. Proc Natl Acad Sci U S A, 1994. **91**(22): p. 10655-9.
184. Arber, S. and P. Caroni, *Specificity of single LIM motifs in targeting and LIM/LIM interactions in situ*. Genes Dev, 1996. **10**(3): p. 289-300.
185. Kadrmas, J.L. and M.C. Beckerle, *The LIM domain: from the cytoskeleton to the nucleus*. Nat Rev Mol Cell Biol, 2004. **5**(11): p. 920-31.
186. Guy, P.M., D.A. Kenny, and G.N. Gill, *The PDZ domain of the LIM protein enigma binds to beta-tropomyosin*. Mol Biol Cell, 1999. **10**(6): p. 1973-84.
187. Maul, R.S., et al., *EPLIN regulates actin dynamics by cross-linking and stabilizing filaments*. J Cell Biol, 2003. **160**(3): p. 399-407.
188. Xia, H., et al., *Actinin-associated LIM protein: identification of a domain interaction between PDZ and spectrin-like repeat motifs*. J Cell Biol, 1997. **139**(2): p. 507-15.
189. Faulkner, G., et al., *ZASP: a new Z-band alternatively spliced PDZ-motif protein*. J Cell Biol, 1999. **146**(2): p. 465-75.
190. Luo, G., et al., *Complete cDNA sequence and tissue localization of N-RAP, a novel nebulin-related protein of striated muscle*. Cell Motil Cytoskeleton, 1997. **38**(1): p. 75-90.
191. Suzuki, T., et al., *MICAL, a novel CasL interacting molecule, associates with vimentin*. J Biol Chem, 2002. **277**(17): p. 14933-41.
192. Roof, D.J., et al., *Molecular characterization of abLIM, a novel actin-binding and double zinc finger protein*. J Cell Biol, 1997. **138**(3): p. 575-88.
193. Grubinger, M. and M. Gimona, *CRP2 is an autonomous actin-binding protein*. FEBS Lett, 2004. **557**(1-3): p. 88-92.
194. Sadler, I., et al., *Zyxin and cCRP: two interactive LIM domain proteins associated with the cytoskeleton*. J Cell Biol, 1992. **119**(6): p. 1573-87.
195. Turner, C.E., J.R. Glenney, Jr., and K. Burridge, *Paxillin: a new vinculin-binding protein present in focal adhesions*. J Cell Biol, 1990. **111**(3): p. 1059-68.
196. Garvalov, B.K., et al., *The conformational state of Tes regulates its zyxin-dependent recruitment to focal adhesions*. J Cell Biol, 2003. **161**(1): p. 33-9.

197. Matsuya, M., et al., *Cell adhesion kinase beta forms a complex with a new member, Hic-5, of proteins localized at focal adhesions*. J Biol Chem, 1998. **273**(2): p. 1003-14.
198. Tu, Y., et al., *Migfilin and Mig-2 link focal adhesions to filamin and the actin cytoskeleton and function in cell shape modulation*. Cell, 2003. **113**(1): p. 37-47.
199. Wixler, V., et al., *The LIM-only protein DRAL/FHL2 binds to the cytoplasmic domain of several alpha and beta integrin chains and is recruited to adhesion complexes*. J Biol Chem, 2000. **275**(43): p. 33669-78.
200. Brown, S., et al., *Characterization of two isoforms of the skeletal muscle LIM protein 1, SLIM1. Localization of SLIM1 at focal adhesions and the isoform slimmer in the nucleus of myoblasts and cytoplasm of myotubes suggests distinct roles in the cytoskeleton and in nuclear-cytoplasmic communication*. J Biol Chem, 1999. **274**(38): p. 27083-91.
201. Coghill, I.D., et al., *FHL3 is an actin-binding protein that regulates alpha-actinin-mediated actin bundling: FHL3 localizes to actin stress fibers and enhances cell spreading and stress fiber disassembly*. J Biol Chem, 2003. **278**(26): p. 24139-52.
202. Tu, Y., et al., *The LIM-only protein PINCH directly interacts with integrin-linked kinase and is recruited to integrin-rich sites in spreading cells*. Mol Cell Biol, 1999. **19**(3): p. 2425-34.
203. Zhang, Y., et al., *Characterization of PINCH-2, a new focal adhesion protein that regulates the PINCH-1-ILK interaction, cell spreading, and migration*. J Biol Chem, 2002. **277**(41): p. 38328-38.
204. Braun, A., et al., *PINCH2 is a new five LIM domain protein, homologous to PINCH and localized to focal adhesions*. Exp Cell Res, 2003. **284**(2): p. 239-50.
205. Ooshio, T., et al., *Involvement of LMO7 in the association of two cell-cell adhesion molecules, nectin and E-cadherin, through afadin and alpha-actinin in epithelial cells*. J Biol Chem, 2004. **279**(30): p. 31365-73.
206. Loughran, G., et al., *Mystique is a new insulin-like growth factor-I-regulated PDZ-LIM domain protein that promotes cell attachment and migration and suppresses Anchorage-independent growth*. Mol Biol Cell, 2005. **16**(4): p. 1811-22.
207. Jani, K. and F. Schock, *Zasp is required for the assembly of functional integrin adhesion sites*. J Cell Biol, 2007. **179**(7): p. 1583-97.
208. Wu, C., *Focal adhesion: a focal point in current cell biology and molecular medicine*. Cell Adh Migr, 2007. **1**(1): p. 13-8.
209. Carragher, N.O. and M.C. Frame, *Focal adhesion and actin dynamics: a place where kinases and proteases meet to promote invasion*. Trends Cell Biol, 2004. **14**(5): p. 241-9.
210. Arber, S., et al., *Regulation of actin dynamics through phosphorylation of cofilin by LIM-kinase*. Nature, 1998. **393**(6687): p. 805-9.
211. Yang, N., et al., *Cofilin phosphorylation by LIM-kinase 1 and its role in Rac-mediated actin reorganization*. Nature, 1998. **393**(6687): p. 809-12.
212. Maekawa, M., et al., *Signaling from Rho to the actin cytoskeleton through protein kinases ROCK and LIM-kinase*. Science, 1999. **285**(5429): p. 895-8.



213. Louis, H.A., et al., *Comparison of three members of the cysteine-rich protein family reveals functional conservation and divergent patterns of gene expression*. J Biol Chem, 1997. **272**(43): p. 27484-91.
214. Arber, S., et al., *MLP-deficient mice exhibit a disruption of cardiac cytoarchitectural organization, dilated cardiomyopathy, and heart failure*. Cell, 1997. **88**(3): p. 393-403.
215. Crone, J., et al., *Zyxin is a critical regulator of the apoptotic HIPK2-p53 signaling axis*. Cancer Res, 2011. **71**(6): p. 2350-9.
216. Hervy, M., et al., *The LIM Protein Zyxin Binds CARP-1 and Promotes Apoptosis*. Genes Cancer, 2010. **1**(5): p. 506-515.
217. James, V., et al., *LIM-domain proteins, LIMD1, Ajuba, and WTIP are required for microRNA-mediated gene silencing*. Proc Natl Acad Sci U S A, 2010. **107**(28): p. 12499-504.
218. Hirota, T., et al., *Aurora-A and an interacting activator, the LIM protein Ajuba, are required for mitotic commitment in human cells*. Cell, 2003. **114**(5): p. 585-98.
219. Ferrand, A., et al., *Ajuba: a new microtubule-associated protein that interacts with BUBR1 and Aurora B at kinetochores in metaphase*. Biol Cell, 2009. **101**(4): p. 221-35.
220. Nix, D.A. and M.C. Beckerle, *Nuclear-cytoplasmic shuttling of the focal contact protein, zyxin: a potential mechanism for communication between sites of cell adhesion and the nucleus*. J Cell Biol, 1997. **138**(5): p. 1139-47.
221. Yang, N. and K. Mizuno, *Nuclear export of LIM-kinase 1, mediated by two leucine-rich nuclear-export signals within the PDZ domain*. Biochem J, 1999. **338 ( Pt 3)**: p. 793-8.
222. Smolich, B., et al., *Cloning and biochemical characterization of LIMK-2, a protein kinase containing two LIM domains*. J Biochem, 1997. **121**(2): p. 382-8.
223. Yang, L., et al., *Interaction of the tau2 transcriptional activation domain of glucocorticoid receptor with a novel steroid receptor coactivator, Hic-5, which localizes to both focal adhesions and the nuclear matrix*. Mol Biol Cell, 2000. **11**(6): p. 2007-18.
224. Kanungo, J., et al., *Ajuba, a cytosolic LIM protein, shuttles into the nucleus and affects embryonal cell proliferation and fate decisions*. Mol Biol Cell, 2000. **11**(10): p. 3299-313.
225. Petit, M.M., et al., *LPP, an actin cytoskeleton protein related to zyxin, harbors a nuclear export signal and transcriptional activation capacity*. Mol Biol Cell, 2000. **11**(1): p. 117-29.
226. Wang, Y. and T.D. Gilmore, *LIM domain protein Trip6 has a conserved nuclear export signal, nuclear targeting sequences, and multiple transactivation domains*. Biochim Biophys Acta, 2001. **1538**(2-3): p. 260-72.
227. Cattaruzza, M., C. Lattrich, and M. Hecker, *Focal adhesion protein zyxin is a mechanosensitive modulator of gene expression in vascular smooth muscle cells*. Hypertension, 2004. **43**(4): p. 726-30.
228. Muller, J.M., et al., *The transcriptional coactivator FHL2 transmits Rho signals from the cell membrane into the nucleus*. Embo J, 2002. **21**(4): p. 736-48.

229. Kasai, M., et al., *The Group 3 LIM domain protein paxillin potentiates androgen receptor transactivation in prostate cancer cell lines*. *Cancer Res*, 2003. **63**(16): p. 4927-35.
230. Campana, W.M., R.R. Myers, and A. Rearden, *Identification of PINCH in Schwann cells and DRG neurons: shuttling and signaling after nerve injury*. *Glia*, 2003. **41**(3): p. 213-23.
231. Srichai, M.B., et al., *A WTI co-regulator controls podocyte phenotype by shuttling between adhesion structures and nucleus*. *J Biol Chem*, 2004. **279**(14): p. 14398-408.
232. Shibanuma, M., et al., *Hic-5 communicates between focal adhesions and the nucleus through oxidant-sensitive nuclear export signal*. *Mol Biol Cell*, 2003. **14**(3): p. 1158-71.
233. Holaska, J.M., S. Rais-Bahrami, and K.L. Wilson, *Lmo7 is an emerin-binding protein that regulates the transcription of emerin and many other muscle-relevant genes*. *Hum Mol Genet*, 2006. **15**(23): p. 3459-72.
234. Osada, H., et al., *Subcellular localization and protein interaction of the human LIMK2 gene expressing alternative transcripts with tissue-specific regulation*. *Biochem Biophys Res Commun*, 1996. **229**(2): p. 582-9.
235. Kong, Y., et al., *Muscle LIM protein promotes myogenesis by enhancing the activity of MyoD*. *Mol Cell Biol*, 1997. **17**(8): p. 4750-60.
236. Wang, H., et al., *Novel function of androgen receptor-associated protein 55/Hic-5 as a negative regulator of Smad3 signaling*. *J Biol Chem*, 2005. **280**(7): p. 5154-62.
237. Guo, B., et al., *The LIM domain protein LPP is a coactivator for the ETS domain transcription factor PEA3*. *Mol Cell Biol*, 2006. **26**(12): p. 4529-38.
238. Shibanuma, M., et al., *A LIM protein, Hic-5, functions as a potential coactivator for Sp1*. *J Cell Biochem*, 2004. **91**(3): p. 633-45.
239. Montoya-Durango, D.E., et al., *Ajuba functions as a histone deacetylase-dependent co-repressor for autoregulation of the growth factor-independent-1 transcription factor*. *J Biol Chem*, 2008. **283**(46): p. 32056-65.
240. Fujimoto, N., et al., *Cloning and characterization of androgen receptor coactivator, ARA55, in human prostate*. *J Biol Chem*, 1999. **274**(12): p. 8316-21.
241. Muller, J.M., et al., *FHL2, a novel tissue-specific coactivator of the androgen receptor*. *Embo J*, 2000. **19**(3): p. 359-69.
242. Ding, L., et al., *FHL1 interacts with oestrogen receptors and regulates breast cancer cell growth*. *J Cell Mol Med*, 2011. **15**(1): p. 72-85.
243. Drori, S., et al., *Hic-5 regulates an epithelial program mediated by PPARgamma*. *Genes Dev*, 2005. **19**(3): p. 362-75.
244. Kassel, O., et al., *A nuclear isoform of the focal adhesion LIM-domain protein Trip6 integrates activating and repressing signals at AP-1- and NF-kappaB-regulated promoters*. *Genes Dev*, 2004. **18**(20): p. 2518-28.
245. Hou, Z., et al., *LIM protein Ajuba functions as a nuclear receptor corepressor and negatively regulates retinoic acid signaling*. *Proc Natl Acad Sci U S A*, 2010. **107**(7): p. 2938-43.

246. Hill, A.A. and P.R. Riley, *Differential regulation of Hand1 homodimer and Hand1-E12 heterodimer activity by the cofactor FHL2*. Mol Cell Biol, 2004. **24**(22): p. 9835-47.
247. Fimia, G.M., D. De Cesare, and P. Sassone-Corsi, *A family of LIM-only transcriptional coactivators: tissue-specific expression and selective activation of CREB and CREM*. Mol Cell Biol, 2000. **20**(22): p. 8613-22.
248. Morlon, A. and P. Sassone-Corsi, *The LIM-only protein FHL2 is a serum-inducible transcriptional coactivator of AP-1*. Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3977-82.
249. Gunther, T., et al., *Fhl2 deficiency results in osteopenia due to decreased activity of osteoblasts*. Embo J, 2005. **24**(17): p. 3049-56.
250. Du, X., et al., *The LIM-only coactivator FHL2 modulates WT1 transcriptional activity during gonadal differentiation*. Biochim Biophys Acta, 2002. **1577**(1): p. 93-101.
251. Martin, B., et al., *The LIM-only protein FHL2 interacts with beta-catenin and promotes differentiation of mouse myoblasts*. J Cell Biol, 2002. **159**(1): p. 113-22.
252. Shi, X., K.M. Bowlin, and D.J. Garry, *Fhl2 interacts with Foxk1 and corepresses Foxo4 activity in myogenic progenitors*. Stem Cells, 2010. **28**(3): p. 462-9.
253. Xiong, Z., et al., *Synergistic repression of estrogen receptor transcriptional activity by FHL2 and Smad4 in breast cancer cells*. IUBMB Life, 2010. **62**(9): p. 669-76.
254. Turner, J., et al., *The LIM protein FHL3 binds basic Kruppel-like factor/Kruppel-like factor 3 and its co-repressor C-terminal-binding protein 2*. J Biol Chem, 2003. **278**(15): p. 12786-95.
255. Takahashi, K., C. Matsumoto, and C. Ra, *FHL3 negatively regulates human high-affinity IgE receptor beta-chain gene expression by acting as a transcriptional co-repressor of MZF-1*. Biochem J, 2005. **386**(Pt 1): p. 191-200.
256. Cottle, D.L., et al., *FHL3 binds MyoD and negatively regulates myotube formation*. J Cell Sci, 2007. **120**(Pt 8): p. 1423-35.
257. Fimia, G.M., D. De Cesare, and P. Sassone-Corsi, *CBP-independent activation of CREM and CREB by the LIM-only protein ACT*. Nature, 1999. **398**(6723): p. 165-9.
258. Lange, S., et al., *Subcellular targeting of metabolic enzymes to titin in heart muscle may be mediated by DRAL/FHL-2*. J Cell Sci, 2002. **115**(Pt 24): p. 4925-36.
259. Canault, M., et al., *FHL2 interacts with both ADAM-17 and the cytoskeleton and regulates ADAM-17 localization and activity*. J Cell Physiol, 2006. **208**(2): p. 363-72.
260. Park, J., et al., *Deficiency in the LIM-only protein FHL2 impairs assembly of extracellular matrix proteins*. FASEB J, 2008. **22**(7): p. 2508-20.
261. Sheppard, S.A., T. Savinova, and D. Loayza, *TRIP6 and LPP, but not Zyxin, are present at a subset of telomeres in human cells*. Cell Cycle, 2011. **10**(11): p. 1726-30.

262. Sheppard, S.A. and D. Loayza, *LIM-domain proteins TRIP6 and LPP associate with shelterin to mediate telomere protection*. Aging (Albany NY), 2010. **2**(7): p. 432-44.
263. Shawlot, W. and R.R. Behringer, *Requirement for Lim1 in head-organizer function*. Nature, 1995. **374**(6521): p. 425-30.
264. Porter, F.D., et al., *Lhx2, a LIM homeobox gene, is required for eye, forebrain, and definitive erythrocyte development*. Development, 1997. **124**(15): p. 2935-44.
265. Sheng, H.Z., et al., *Specification of pituitary cell lineages by the LIM homeobox gene Lhx3*. Science, 1996. **272**(5264): p. 1004-7.
266. Sheng, H.Z., et al., *Multistep control of pituitary organogenesis*. Science, 1997. **278**(5344): p. 1809-12.
267. Agulnick, A.D., et al., *Interactions of the LIM-domain-binding factor Ldb1 with LIM homeodomain proteins*. Nature, 1996. **384**(6606): p. 270-2.
268. Thaler, J.P., et al., *LIM factor Lhx3 contributes to the specification of motor neuron and interneuron identity through cell-type-specific protein-protein interactions*. Cell, 2002. **110**(2): p. 237-49.
269. Jurata, L.W., S.L. Pfaff, and G.N. Gill, *The nuclear LIM domain interactor NLI mediates homo- and heterodimerization of LIM domain transcription factors*. J Biol Chem, 1998. **273**(6): p. 3152-7.
270. Mukhopadhyay, M., et al., *Functional ablation of the mouse Ldb1 gene results in severe patterning defects during gastrulation*. Development, 2003. **130**(3): p. 495-505.
271. Zhu, T.H., et al., *A single ancestral gene of the human LIM domain oncogene family LMO in Drosophila: characterization of the Drosophila Dlmo gene*. Oncogene, 1995. **11**(7): p. 1283-90.
272. Zeng, C., et al., *The Drosophila LIM-only gene, dLMO, is mutated in Beadex alleles and might represent an evolutionarily conserved function in appendage development*. Proc Natl Acad Sci U S A, 1998. **95**(18): p. 10637-42.
273. Milan, M., F.J. Diaz-Benjumea, and S.M. Cohen, *Beadex encodes an LMO protein that regulates Apterous LIM-homeodomain activity in Drosophila wing development: a model for LMO oncogene function*. Genes Dev, 1998. **12**(18): p. 2912-20.
274. Shores, M., et al., *Overexpression Beadex mutations and loss-of-function heldup-a mutations in Drosophila affect the 3' regulatory and coding components, respectively, of the Dlmo gene*. Genetics, 1998. **150**(1): p. 283-99.
275. Milan, M. and S.M. Cohen, *Temporal regulation of apterous activity during development of the Drosophila wing*. Development, 2000. **127**(14): p. 3069-78.
276. Milan, M. and S.M. Cohen, *Regulation of LIM homeodomain activity in vivo: a tetramer of dLDB and apterous confers activity and capacity for regulation by dLMO*. Mol Cell, 1999. **4**(2): p. 267-73.
277. Zenvirt, S., et al., *Drosophila LIM-only is a positive regulator of transcription during thoracic bristle development*. Genetics, 2008. **179**(4): p. 1989-99.

278. McGuire, E.A., et al., *The t(11;14)(p15;q11) in a T-cell acute lymphoblastic leukemia cell line activates multiple transcripts, including Ttg-1, a gene encoding a potential zinc finger protein.* Mol Cell Biol, 1989. **9**(5): p. 2124-32.
279. McGuire, E.A., et al., *Thymic overexpression of Ttg-1 in transgenic mice results in T-cell acute lymphoblastic leukemia/lymphoma.* Mol Cell Biol, 1992. **12**(9): p. 4186-96.
280. Chervinsky, D.S., et al., *Disordered T-cell development and T-cell malignancies in SCL LMO1 double-transgenic mice: parallels with E2A-deficient mice.* Mol Cell Biol, 1999. **19**(7): p. 5025-35.
281. Greenberg, J.M., et al., *Segmental and developmental regulation of a presumptive T-cell oncogene in the central nervous system.* Nature, 1990. **344**(6262): p. 158-60.
282. Wang, K., et al., *Integrative genomics identifies LMO1 as a neuroblastoma oncogene.* Nature, 2011. **469**(7329): p. 216-20.
283. Gratzinger, D., et al., *The transcription factor LMO2 is a robust marker of vascular endothelium and vascular neoplasms and selected other entities.* Am J Clin Pathol, 2009. **131**(2): p. 264-78.
284. Lossos, I.S., et al., *Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes.* N Engl J Med, 2004. **350**(18): p. 1828-37.
285. Nakata, K., et al., *LMO2 is a novel predictive marker for a better prognosis in pancreatic cancer.* Neoplasia, 2009. **11**(7): p. 712-9.
286. Cobanoglu, U., et al., *The expression of LMO2 protein in acute B-cell and myeloid leukemia.* Hematology, 2010. **15**(3): p. 132-4.
287. Ma, S., et al., *The significance of LMO2 expression in the progression of prostate cancer.* J Pathol, 2007. **211**(3): p. 278-85.
288. Malumbres, R., et al., *LMO2 expression reflects the different stages of blast maturation and genetic features in B-cell acute lymphoblastic leukemia and predicts clinical outcome.* Haematologica, 2011. **96**(7): p. 980-6.
289. Aoyama, M., et al., *LMO3 interacts with neuronal transcription factor, HEN2, and acts as an oncogene in neuroblastoma.* Cancer Res, 2005. **65**(11): p. 4587-97.
290. Isogai, E., et al., *Oncogenic LMO3 Collaborates with HEN2 to Enhance Neuroblastoma Cell Growth through Transactivation of Mash1.* PLoS One, 2011. **6**(5): p. e19297.
291. Tse, E., et al., *Null mutation of the Lmo4 gene or a combined null mutation of the Lmo1/Lmo3 genes causes perinatal lethality, and Lmo4 controls neural tube development in mice.* Mol Cell Biol, 2004. **24**(5): p. 2063-73.
292. Kenny, D.A., et al., *Identification and characterization of LMO4, an LMO gene with a novel pattern of expression during embryogenesis.* Proc Natl Acad Sci U S A, 1998. **95**(19): p. 11257-62.
293. Sugihara, T.M., et al., *Mouse deformed epidermal autoregulatory factor 1 recruits a LIM domain factor, LMO-4, and CLIM coregulators.* Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15418-23.
294. Racevskis, J., et al., *Molecular cloning of LMO41, a new human LIM domain gene.* Biochim Biophys Acta, 1999. **1445**(1): p. 148-53.

295. Yu, J., et al., *LIM only 4 is overexpressed in late stage pancreas cancer*. Mol Cancer, 2008. **7**: p. 93.
296. Visvader, J.E., et al., *The LIM domain gene LMO4 inhibits differentiation of mammary epithelial cells in vitro and is overexpressed in breast cancer*. Proc Natl Acad Sci U S A, 2001. **98**(25): p. 14452-7.
297. Sum, E.Y., et al., *Overexpression of LMO4 induces mammary hyperplasia, promotes cell invasion, and is a predictor of poor outcome in breast cancer*. Proc Natl Acad Sci U S A, 2005. **102**(21): p. 7659-64.
298. Montanez-Wiscovich, M.E., et al., *LMO4 is an essential mediator of ErbB2/HER2/Neu-induced breast cancer cell cycle progression*. Oncogene, 2009. **28**(41): p. 3608-18.
299. Montanez-Wiscovich, M.E., et al., *Aberrant expression of LMO4 induces centrosome amplification and mitotic spindle abnormalities in breast cancer cells*. J Pathol, 2010. **222**(3): p. 271-81.
300. Hahm, K., et al., *Defective neural tube closure and anteroposterior patterning in mice lacking the LIM protein LMO4 or its interacting partner Deaf-1*. Mol Cell Biol, 2004. **24**(5): p. 2074-82.
301. Lee, S.K., et al., *The LIM domain-only protein LMO4 is required for neural tube closure*. Mol Cell Neurosci, 2005. **28**(2): p. 205-14.
302. Sum, E.Y., et al., *Loss of the LIM domain protein Lmo4 in the mammary gland during pregnancy impedes lobuloalveolar development*. Oncogene, 2005. **24**(30): p. 4820-8.
303. Huang, Z., et al., *Transcription factor Lmo4 defines the shape of functional areas in developing cortices and regulates sensorimotor control*. Dev Biol, 2009. **327**(1): p. 132-42.
304. Duquette, P.M., et al., *Loss of LMO4 in the retina leads to reduction of GABAergic amacrine cells and functional deficits*. PLoS One, 2010. **5**(10): p. e13232.
305. Lasek, A.W., et al., *Lmo4 in the nucleus accumbens regulates cocaine sensitivity*. Genes Brain Behav, 2010. **9**(7): p. 817-24.
306. Tsai, L.T., et al., *Lmo mutants reveal a novel role for circadian pacemaker neurons in cocaine-induced behaviors*. PLoS Biol, 2004. **2**(12): p. e408.
307. Asprer, J.S., et al., *LMO4 functions as a co-activator of neurogenin 2 in the developing cortex*. Development, 2011. **138**(13): p. 2823-32.
308. Han, C., et al., *Human Bex2 interacts with LMO2 and regulates the transcriptional activity of a novel DNA-binding complex*. Nucleic Acids Res, 2005. **33**(20): p. 6555-65.
309. Manetopoulos, C., et al., *The LIM-only protein LMO4 modulates the transcriptional activity of HEN1*. Biochem Biophys Res Commun, 2003. **307**(4): p. 891-9.
310. Kudryavtseva, E.I., et al., *Identification and characterization of Grainyhead-like epithelial transactivator (GET-1), a novel mammalian Grainyhead-like factor*. Dev Dyn, 2003. **226**(4): p. 604-17.
311. Hislop, N.R., et al., *Grhl3 and Lmo4 play coordinate roles in epidermal migration*. Dev Biol, 2008. **321**(1): p. 263-72.

312. Sum, E.Y., et al., *The LIM domain protein LMO4 interacts with the cofactor CtIP and the tumor suppressor BRCA1 and inhibits BRCA1 activity.* J Biol Chem, 2002. **277**(10): p. 7849-56.
313. Singh, R.R., et al., *Negative regulation of estrogen receptor alpha transactivation functions by LIM domain only 4 protein.* Cancer Res, 2005. **65**(22): p. 10594-601.
314. Setogawa, T., et al., *The tumor suppressor LKB1 induces p21 expression in collaboration with LMO4, GATA-6, and Ldb1.* Biochem Biophys Res Commun, 2006. **343**(4): p. 1186-90.
315. Joshi, K., et al., *LMO4 controls the balance between excitatory and inhibitory spinal V2 interneurons.* Neuron, 2009. **61**(6): p. 839-51.
316. Valge-Archer, V.E., et al., *The LIM protein RBTN2 and the basic helix-loop-helix protein TAL1 are present in a complex in erythroid cells.* Proc Natl Acad Sci U S A, 1994. **91**(18): p. 8617-21.
317. Wadman, I., et al., *Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia.* Embo J, 1994. **13**(20): p. 4831-9.
318. Osada, H., et al., *Association of erythroid transcription factors: complexes involving the LIM protein RBTN2 and the zinc-finger protein GATA1.* Proc Natl Acad Sci U S A, 1995. **92**(21): p. 9585-9.
319. Osada, H., et al., *LIM-only protein Lmo2 forms a protein complex with erythroid transcription factor GATA-1.* Leukemia, 1997. **11 Suppl 3**: p. 307-12.
320. Wadman, I.A., et al., *The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins.* Embo J, 1997. **16**(11): p. 3145-57.
321. Grutz, G.G., et al., *The oncogenic T cell LIM-protein Lmo2 forms part of a DNA-binding complex specifically in immature T cells.* Embo J, 1998. **17**(16): p. 4594-605.
322. Ono, Y., N. Fukuhara, and O. Yoshie, *TAL1 and LIM-only proteins synergistically induce retinaldehyde dehydrogenase 2 expression in T-cell acute lymphoblastic leukemia by acting as cofactors for GATA3.* Mol Cell Biol, 1998. **18**(12): p. 6939-50.
323. Hsu, H.L., et al., *Enhancer-binding activity of the tal-1 oncoprotein in association with the E47/E12 helix-loop-helix proteins.* Mol Cell Biol, 1991. **11**(6): p. 3037-42.
324. Voronova, A.F. and F. Lee, *The E2A and tal-1 helix-loop-helix proteins associate in vivo and are modulated by Id proteins during interleukin 6-induced myeloid differentiation.* Proc Natl Acad Sci U S A, 1994. **91**(13): p. 5952-6.
325. O'Neil, J., et al., *The DNA binding activity of TAL-1 is not required to induce leukemia/lymphoma in mice.* Oncogene, 2001. **20**(29): p. 3897-905.
326. Goldfarb, A.N., K. Lewandowska, and M. Shoham, *Determinants of helix-loop-helix dimerization affinity. Random mutational analysis of SCL/tal.* J Biol Chem, 1996. **271**(5): p. 2683-8.

327. Kassouf, M.T., et al., *Genome-wide identification of TAL1's functional targets: insights into its mechanisms of action in primary erythroid cells*. Genome Res, 2010. **20**(8): p. 1064-83.
328. Vyas, P., et al., *Different sequence requirements for expression in erythroid and megakaryocytic cells within a regulatory element upstream of the GATA-1 gene*. Development, 1999. **126**(12): p. 2799-811.
329. Goardon, N., et al., *ETO2 coordinates cellular proliferation and differentiation during erythropoiesis*. Embo J, 2006. **25**(2): p. 357-66.
330. Pali, C.G., et al., *Differential genomic targeting of the transcription factor TAL1 in alternate haematopoietic lineages*. Embo J, 2011. **30**(3): p. 494-509.
331. Lecuyer, E., et al., *The SCL complex regulates c-kit expression in hematopoietic cells through functional interaction with Sp1*. Blood, 2002. **100**(7): p. 2430-40.
332. Talora, C., et al., *Cross talk among Notch3, pre-TCR, and Tall in T-cell development and leukemogenesis*. Blood, 2006. **107**(8): p. 3313-20.
333. Bain, G., et al., *E2A deficiency leads to abnormalities in alphabeta T-cell development and to rapid development of T-cell lymphomas*. Mol Cell Biol, 1997. **17**(8): p. 4782-91.
334. Barndt, R., M.F. Dai, and Y. Zhuang, *A novel role for HEB downstream or parallel to the pre-TCR signaling pathway during alpha beta thymopoiesis*. J Immunol, 1999. **163**(6): p. 3331-43.
335. O'Neil, J., et al., *TAL1/SCL induces leukemia by inhibiting the transcriptional activity of E47/HEB*. Cancer Cell, 2004. **5**(6): p. 587-96.
336. Tremblay, M., et al., *Regulation of pT alpha gene expression by a dosage of E2A, HEB, and SCL*. J Biol Chem, 2003. **278**(15): p. 12680-7.
337. Bernard, M., et al., *Helix-loop-helix (E2-5, HEB, TAL1 and Id1) protein interaction with the TCRalphadelta enhancers*. Int Immunol, 1998. **10**(10): p. 1539-49.
338. Park, S.T. and X.H. Sun, *The Tall oncoprotein inhibits E47-mediated transcription. Mechanism of inhibition*. J Biol Chem, 1998. **273**(12): p. 7030-7.
339. Wilson, N.K., et al., *Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators*. Cell Stem Cell, 2010. **7**(4): p. 532-44.
340. Kusy, S., et al., *NKX3.1 is a direct TAL1 target gene that mediates proliferation of TAL1-expressing human T cell acute lymphoblastic leukemia*. J Exp Med, 2010. **207**(10): p. 2141-56.
341. Yuan, W., et al., *Downregulation of microRNA-142 by proto-oncogene LMO2 and its co-factors*. Leukemia, 2008. **22**(5): p. 1067-71.
342. Palomero, T., et al., *Transcriptional regulatory networks downstream of TAL1/SCL in T-cell acute lymphoblastic leukemia*. Blood, 2006. **108**(3): p. 986-92.
343. Tripic, T., et al., *SCL and associated proteins distinguish active from repressive GATA transcription factor complexes*. Blood, 2009. **113**(10): p. 2191-201.
344. Soler, E., et al., *The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation*. Genes Dev, 2010. **24**(3): p. 277-89.



345. Wilkinson, D.A., et al., *Elf-2, a rhombotin-2 binding ets transcription factor: discovery and potential role in T cell leukemia*. *Leukemia*, 1997. **11**(1): p. 86-96.
346. Mao, S., G.A. Neale, and R.M. Goorha, *T-cell oncogene rhombotin-2 interacts with retinoblastoma-binding protein 2*. *Oncogene*, 1997. **14**(13): p. 1531-9.
347. Begay-Muller, V., S. Ansieau, and A. Leutz, *The LIM domain protein Lmo2 binds to AF6, a translocation partner of the MLL oncogene*. *FEBS Lett*, 2002. **521**(1-3): p. 36-8.
348. Zhadanov, A.B., et al., *Absence of the tight junctional protein AF-6 disrupts epithelial cell-cell junctions and cell polarity during mouse development*. *Curr Biol*, 1999. **9**(16): p. 880-8.
349. Meyer, C., et al., *New insights to the MLL recombinome of acute leukemias*. *Leukemia*, 2009. **23**(8): p. 1490-9.
350. De Braekeleer, E., et al., *Identification of a MLL-MLLT4 fusion gene resulting from a t(6;11)(q27;q23) presenting as a del(11q) in a child with T-cell acute lymphoblastic leukemia*. *Leuk Lymphoma*, 2010. **51**(8): p. 1570-3.
351. Rabbitts, T.H., *Chromosomal translocations in human cancer*. *Nature*, 1994. **372**(6502): p. 143-9.
352. Look, A.T., *Oncogenic transcription factors in the human acute leukemias*. *Science*, 1997. **278**(5340): p. 1059-64.
353. Groth, A., et al., *Chromatin challenges during DNA replication and repair*. *Cell*, 2007. **128**(4): p. 721-33.
354. Murr, R., *Interplay between different epigenetic modifications and mechanisms*. *Adv Genet*, 2010. **70**: p. 101-41.
355. Hanawalt, P.C., *Paradigms for the three rs: DNA replication, recombination, and repair*. *Mol Cell*, 2007. **28**(5): p. 702-7.
356. Bell, S.P. and B. Stillman, *ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex*. *Nature*, 1992. **357**(6374): p. 128-34.
357. Masai, H., et al., *Eukaryotic chromosome DNA replication: where, when, and how?* *Annu Rev Biochem*, 2010. **79**: p. 89-130.
358. Mechali, M., *Eukaryotic DNA replication origins: many choices for appropriate answers*. *Nat Rev Mol Cell Biol*, 2010. **11**(10): p. 728-38.
359. Cotobal, C., M. Segurado, and F. Antequera, *Structural diversity and dynamics of genomic replication origins in Schizosaccharomyces pombe*. *Embo J*, 2010. **29**(5): p. 934-42.
360. Huberman, J.A. and A.D. Riggs, *Autoradiography of chromosomal DNA fibers from Chinese hamster cells*. *Proc Natl Acad Sci U S A*, 1966. **55**(3): p. 599-606.
361. Cadoret, J.C., et al., *Genome-wide studies highlight indirect links between human replication origins and gene regulation*. *Proc Natl Acad Sci U S A*, 2008. **105**(41): p. 15837-42.
362. Berezney, R., D.D. Dubey, and J.A. Huberman, *Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci*. *Chromosoma*, 2000. **108**(8): p. 471-84.

363. Sequeira-Mendes, J., et al., *Transcription initiation activity sets replication origin efficiency in mammalian cells*. PLoS Genet, 2009. **5**(4): p. e1000446.
364. Sasaki, T. and D.M. Gilbert, *The many faces of the origin recognition complex*. Curr Opin Cell Biol, 2007. **19**(3): p. 337-43.
365. MacAlpine, D.M., H.K. Rodriguez, and S.P. Bell, *Coordination of replication and transcription along a Drosophila chromosome*. Genes Dev, 2004. **18**(24): p. 3094-105.
366. MacAlpine, H.K., et al., *Drosophila ORC localizes to open chromatin and marks sites of cohesin complex loading*. Genome Res, 2010. **20**(2): p. 201-11.
367. Spradling, A. and T. Orr-Weaver, *Regulation of DNA replication during Drosophila development*. Annu Rev Genet, 1987. **21**: p. 373-403.
368. Nordman, J., et al., *Developmental control of the DNA replication and transcription programs*. Genome Res, 2011. **21**(2): p. 175-81.
369. Roy, S., et al., *Identification of functional elements and regulatory circuits by Drosophila modENCODE*. Science, 2010. **330**(6012): p. 1787-97.
370. Abdurashidova, G., et al., *Start sites of bidirectional DNA synthesis at the human lamin B2 origin*. Science, 2000. **287**(5460): p. 2023-6.
371. Cohen, S.M., et al., *Mapping of an origin of DNA replication near the transcriptional promoter of the human HPRT gene*. J Cell Biochem, 2002. **85**(2): p. 346-56.
372. Tao, L., et al., *Major DNA replication initiation sites in the c-myc locus in human cells*. J Cell Biochem, 2000. **78**(3): p. 442-57.
373. Taira, T., S.M. Iguchi-Ariga, and H. Ariga, *A novel DNA replication origin identified in the human heat shock protein 70 gene promoter*. Mol Cell Biol, 1994. **14**(9): p. 6386-97.
374. Kamath, S. and M. Leffak, *Multiple sites of replication initiation in the human beta-globin gene locus*. Nucleic Acids Res, 2001. **29**(3): p. 809-17.
375. Kitsberg, D., et al., *Replication structure of the human beta-globin gene domain*. Nature, 1993. **366**(6455): p. 588-90.
376. Cohen, S.M., et al., *Same origins of DNA replication function on the active and inactive human X chromosomes*. J Cell Biochem, 2003. **88**(5): p. 923-31.
377. Keller, C., et al., *The origin recognition complex marks a replication origin in the human TOP1 gene promoter*. J Biol Chem, 2002. **277**(35): p. 31430-40.
378. Ladenburger, E.M., C. Keller, and R. Knippers, *Identification of a binding region for human origin recognition complex proteins 1 and 2 that coincides with an origin of DNA replication*. Mol Cell Biol, 2002. **22**(4): p. 1036-48.
379. Romero, J. and H. Lee, *Asymmetric bidirectional replication at the human DBF4 origin*. Nat Struct Mol Biol, 2008. **15**(7): p. 722-9.
380. Gray, S.J., et al., *An origin of DNA replication in the promoter region of the human fragile X mental retardation (FMR1) gene*. Mol Cell Biol, 2007. **27**(2): p. 426-37.
381. Kholodii, G., O. Dantsevich, and V. Tarantul, *New origins of DNA replication identified near the human IFNA2 gene*. Cell Cycle, 2008. **7**(24): p. 3949-51.
382. Little, R.D., T.H. Platt, and C.L. Schildkraut, *Initiation and termination of DNA replication in human rRNA genes*. Mol Cell Biol, 1993. **13**(10): p. 6600-13.

383. Norio, P., et al., *Progressive activation of DNA replication initiation in large domains of the immunoglobulin heavy chain locus during B cell development*. Mol Cell, 2005. **20**(4): p. 575-87.
384. Zhou, J., et al., *The origin of a developmentally regulated Igh replicon is located near the border of regulatory domains for Igh replication and expression*. Proc Natl Acad Sci U S A, 2002. **99**(21): p. 13693-8.
385. Gerhardt, J., et al., *Identification of new human origins of DNA replication by an origin-trapping assay*. Mol Cell Biol, 2006. **26**(20): p. 7731-46.
386. Araujo, F.D., et al., *Identification of initiation sites for DNA replication in the human dnmt1 (DNA-methyltransferase) locus*. J Biol Chem, 1999. **274**(14): p. 9335-41.
387. Hu, L., X. Xu, and M.S. Valenzuela, *Initiation sites for human DNA replication at a putative ribulose-5-phosphate 3-epimerase gene*. Biochem Biophys Res Commun, 2004. **320**(3): p. 648-55.
388. Callejo, M., et al., *Identification and functional analysis of a human homologue of the monkey replication origin ors8*. J Cell Biochem, 2006. **99**(6): p. 1606-15.
389. Dijkwel, P.A., S. Wang, and J.L. Hamlin, *Initiation sites are distributed at frequent intervals in the Chinese hamster dihydrofolate reductase origin of replication but are used with very different efficiencies*. Mol Cell Biol, 2002. **22**(9): p. 3053-65.
390. Kobayashi, T., T. Rein, and M.L. DePamphilis, *Identification of primary initiation sites for DNA replication in the hamster dihydrofolate reductase gene initiation zone*. Mol Cell Biol, 1998. **18**(6): p. 3266-77.
391. Todorovic, V., et al., *Human origins of DNA replication selected from a library of nascent DNA*. Mol Cell, 2005. **19**(4): p. 567-75.
392. Cayrou, C., et al., *Genome-scale analysis of metazoan replication origins reveals their organization in specific but flexible sites defined by conserved features*. Genome Res, 2011. **21**(9): p. 1438-49.
393. Birney, E., et al., *Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project*. Nature, 2007. **447**(7146): p. 799-816.
394. Karnani, N., et al., *Genomic study of replication initiation in human chromosomes reveals the influence of transcription regulation and chromatin structure on origin selection*. Mol Biol Cell, 2010. **21**(3): p. 393-404.
395. Pina, B., U. Bruggemeier, and M. Beato, *Nucleosome positioning modulates accessibility of regulatory proteins to the mouse mammary tumor virus promoter*. Cell, 1990. **60**(5): p. 719-31.
396. Yuan, G.C., et al., *Genome-scale identification of nucleosome positions in S. cerevisiae*. Science, 2005. **309**(5734): p. 626-30.
397. Rando, O.J. and K. Ahmad, *Rules and regulation in the primary structure of chromatin*. Curr Opin Cell Biol, 2007. **19**(3): p. 250-6.
398. Nieduszynski, C.A., Y. Knox, and A.D. Donaldson, *Genome-wide identification of replication origins in yeast by comparative genomics*. Genes Dev, 2006. **20**(14): p. 1874-9.
399. Lipford, J.R. and S.P. Bell, *Nucleosomes positioned by ORC facilitate the initiation of DNA replication*. Mol Cell, 2001. **7**(1): p. 21-30.

400. Eaton, M.L., et al., *Conserved nucleosome positioning defines replication origins*. Genes Dev, 2010. **24**(8): p. 748-53.
401. Berbenetz, N.M., C. Nislow, and G.W. Brown, *Diversity of eukaryotic DNA replication origins revealed by genome-wide analysis of chromatin structure*. PLoS Genet, 2010. **6**(9).
402. Deal, R.B., J.G. Henikoff, and S. Henikoff, *Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones*. Science, 2010. **328**(5982): p. 1161-4.
403. Miotto, B. and K. Struhl, *HBO1 histone acetylase activity is essential for DNA replication licensing and inhibited by Geminin*. Mol Cell, 2010. **37**(1): p. 57-66.
404. Tardat, M., et al., *The histone H4 Lys 20 methyltransferase PR-Set7 regulates replication origins in mammalian cells*. Nat Cell Biol, 2010. **12**(11): p. 1086-93.
405. Cadoret, J.C. and M.N. Prioleau, *Genome-wide approaches to determining origin distribution*. Chromosome Res, 2010. **18**(1): p. 79-89.
406. Gomez, M. and N. Brockdorff, *Heterochromatin on the inactive X chromosome delays replication timing without affecting origin usage*. Proc Natl Acad Sci U S A, 2004. **101**(18): p. 6923-8.
407. Bell, S.P., et al., *The multidomain structure of Orc1p reveals similarity to regulators of DNA replication and transcriptional silencing*. Cell, 1995. **83**(4): p. 563-8.
408. Zhang, Z., et al., *Structure and function of the BAH-containing domain of Orc1p in epigenetic silencing*. Embo J, 2002. **21**(17): p. 4600-11.
409. Palacios DeBeer, M.A., U. Muller, and C.A. Fox, *Differential DNA affinity specifies roles for the origin recognition complex in budding yeast heterochromatin*. Genes Dev, 2003. **17**(15): p. 1817-22.
410. Hayashi, M.T., et al., *The heterochromatin protein Swi6/HP1 activates replication origins at the pericentromeric region and silent mating-type locus*. Nat Cell Biol, 2009. **11**(3): p. 357-62.
411. Kohzaki, H. and Y. Murakami, *Transcription factors and DNA replication origin selection*. Bioessays, 2005. **27**(11): p. 1107-16.
412. Di Micco, R., et al., *Oncogene-induced senescence is a DNA damage response triggered by DNA hyper-replication*. Nature, 2006. **444**(7119): p. 638-42.
413. Bartkova, J., et al., *Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints*. Nature, 2006. **444**(7119): p. 633-7.
414. Studzinski, G.P., et al., *Participation of c-myc protein in DNA synthesis of human cells*. Science, 1986. **234**(4775): p. 467-70.
415. Gutierrez, C., et al., *Is c-myc protein directly involved in DNA replication?* Science, 1988. **240**(4856): p. 1202-3.
416. Dominguez-Sola, D., et al., *Non-transcriptional control of DNA replication by c-Myc*. Nature, 2007. **448**(7152): p. 445-51.
417. Herold, S., B. Herkert, and M. Eilers, *Facilitating replication under stress: an oncogenic function of MYC?* Nat Rev Cancer, 2009. **9**(6): p. 441-4.

418. Kumagai, A., et al., *TopBP1 activates the ATR-ATRIP complex*. Cell, 2006. **124**(5): p. 943-55.
419. Robinson, K., et al., *c-Myc accelerates S-Phase and requires WRN to avoid replication stress*. PLoS One, 2009. **4**(6): p. e5951.
420. Oh, I.H. and E.P. Reddy, *The myb gene family in cell growth, differentiation and apoptosis*. Oncogene, 1999. **18**(19): p. 3017-33.
421. Ramsay, R.G. and T.J. Gonda, *MYB function in normal and cancer cells*. Nat Rev Cancer, 2008. **8**(7): p. 523-34.
422. Beall, E.L., et al., *Role for a Drosophila Myb-containing protein complex in site-specific DNA replication*. Nature, 2002. **420**(6917): p. 833-7.
423. Manak, J.R., N. Mitiku, and J.S. Lipsick, *Mutation of the Drosophila homologue of the Myb protooncogene causes genomic instability*. Proc Natl Acad Sci U S A, 2002. **99**(11): p. 7438-43.
424. Markey, M.P., et al., *Unbiased analysis of RB-mediated transcriptional repression identifies novel targets and distinctions from E2F action*. Cancer Res, 2002. **62**(22): p. 6587-97.
425. Angus, S.P., et al., *RB reversibly inhibits DNA replication via two temporally distinct mechanisms*. Mol Cell Biol, 2004. **24**(12): p. 5404-20.
426. Sterner, J.M., et al., *Negative regulation of DNA replication by the retinoblastoma protein is mediated by its association with MCM7*. Mol Cell Biol, 1998. **18**(5): p. 2748-57.
427. Pennaneach, V., et al., *The large subunit of replication factor C promotes cell survival after DNA damage in an LxCxE motif- and Rb-dependent manner*. Mol Cell, 2001. **7**(4): p. 715-27.
428. Mendoza-Maldonado, R., et al., *Interaction of the retinoblastoma protein with Orc1 and its recruitment to human origins of DNA replication*. PLoS One, 2010. **5**(11): p. e13720.
429. Takemura, M., et al., *Phosphorylated retinoblastoma protein stimulates DNA polymerase alpha*. Oncogene, 1997. **15**(20): p. 2483-92.
430. Kennedy, B.K., et al., *Nuclear organization of DNA replication in primary mammalian cells*. Genes Dev, 2000. **14**(22): p. 2855-68.
431. Maser, R.S., et al., *Mre11 complex and DNA replication: linkage to E2F and sites of DNA synthesis*. Mol Cell Biol, 2001. **21**(17): p. 6006-16.
432. Royzman, I., et al., *ORC localization in Drosophila follicle cells and the effects of mutations in dE2F and dDP*. Genes Dev, 1999. **13**(7): p. 827-40.
433. Bosco, G., W. Du, and T.L. Orr-Weaver, *DNA replication control through interaction of E2F-RB and the origin recognition complex*. Nat Cell Biol, 2001. **3**(3): p. 289-95.
434. Avni, D., et al., *Active localization of the retinoblastoma protein in chromatin and its response to S phase DNA damage*. Mol Cell, 2003. **12**(3): p. 735-46.
435. Luo, L., et al., *The cell-cycle regulator geminin inhibits Hox function through direct and polycomb-mediated interactions*. Nature, 2004. **427**(6976): p. 749-53.
436. de Stanchina, E., et al., *Selection of homeotic proteins for binding to a human DNA replication origin*. J Mol Biol, 2000. **299**(3): p. 667-80.
437. Comelli, L., et al., *The homeotic protein HOXC13 is a member of human DNA replication complexes*. Cell Cycle, 2009. **8**(3): p. 454-9.

438. Marchetti, L., et al., *Homeotic proteins participate in the function of human-DNA replication origins*. Nucleic Acids Res, 2010. **38**(22): p. 8105-19.
439. Salsi, V., et al., *HOXD13 binds DNA replication origins to promote origin licensing and is inhibited by geminin*. Mol Cell Biol, 2009. **29**(21): p. 5775-88.
440. Karmakar, S., et al., *A multiprotein complex necessary for both transcription and DNA replication at the beta-globin locus*. Embo J, 2010. **29**(19): p. 3260-71.
441. Schubeler, D., et al., *Genome-wide DNA replication profile for Drosophila melanogaster: a link between transcription and replication timing*. Nat Genet, 2002. **32**(3): p. 438-42.
442. Farkash-Amar, S., et al., *Global organization of replication time zones of the mouse genome*. Genome Res, 2008. **18**(10): p. 1562-70.
443. Hiratani, I., et al., *Global reorganization of replication domains during embryonic stem cell differentiation*. PLoS Biol, 2008. **6**(10): p. e245.
444. Perry, P., et al., *A dynamic switch in the replication timing of key regulator genes in embryonic stem cells upon neural induction*. Cell Cycle, 2004. **3**(12): p. 1645-50.
445. Woodfine, K., et al., *Replication timing of the human genome*. Hum Mol Genet, 2004. **13**(2): p. 191-202.
446. White, E.J., et al., *DNA replication-timing analysis of human chromosome 22 at high resolution and different developmental states*. Proc Natl Acad Sci U S A, 2004. **101**(51): p. 17771-6.
447. Jeon, Y., et al., *Temporal profile of replication of human chromosomes*. Proc Natl Acad Sci U S A, 2005. **102**(18): p. 6419-24.
448. Karnani, N., et al., *Pan-S replication patterns and chromosomal domains defined by genome-tiling arrays of ENCODE genomic areas*. Genome Res, 2007. **17**(6): p. 865-76.
449. Raghuraman, M.K., et al., *Replication dynamics of the yeast genome*. Science, 2001. **294**(5540): p. 115-21.
450. Argiropoulos, B. and R.K. Humphries, *Hox genes in hematopoiesis and leukemogenesis*. Oncogene, 2007. **26**(47): p. 6766-76.
451. Boxer, L.M. and C.V. Dang, *Translocations involving c-myc and c-myc function*. Oncogene, 2001. **20**(40): p. 5595-610.
452. Hoang, T. and T. Hoang, *The T-ALL paradox in cancer*. Nat Med, 2010. **16**(11): p. 1185-6.
453. Negrini, S., V.G. Gorgoulis, and T.D. Halazonetis, *Genomic instability--an evolving hallmark of cancer*. Nat Rev Mol Cell Biol, 2010. **11**(3): p. 220-8.
454. Mailand, N. and J.F. Diffley, *CDKs promote DNA replication origin licensing in human cells by protecting Cdc6 from APC/C-dependent proteolysis*. Cell, 2005. **122**(6): p. 915-26.
455. Nevis, K.R., M. Cordeiro-Stone, and J.G. Cook, *Origin licensing and p53 status regulate Cdk2 activity during G(1)*. Cell Cycle, 2009. **8**(12): p. 1952-63.
456. Shima, N., et al., *A viable allele of Mcm4 causes chromosome instability and mammary adenocarcinomas in mice*. Nat Genet, 2007. **39**(1): p. 93-8.

457. Pruitt, S.C., K.J. Bailey, and A. Freeland, *Reduced Mcm2 expression results in severe stem/progenitor cell deficiency and cancer*. *Stem Cells*, 2007. **25**(12): p. 3121-32.
458. Sheen, J.H. and R.B. Dickson, *Overexpression of c-Myc alters G(1)/S arrest following ionizing radiation*. *Mol Cell Biol*, 2002. **22**(6): p. 1819-33.
459. Land, H., L.F. Parada, and R.A. Weinberg, *Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes*. *Nature*, 1983. **304**(5927): p. 596-602.
460. Nam, C.H. and T.H. Rabbitts, *The role of LMO2 in development and in T cell leukemia after chromosomal translocation or retroviral insertion*. *Mol Ther*, 2006. **13**(1): p. 15-25.
461. Lecuyer, E. and T. Hoang, *SCL: from the origin of hematopoiesis to stem cells and leukemia*. *Exp Hematol*, 2004. **32**(1): p. 11-24.
462. Jurata, L.W., D.A. Kenny, and G.N. Gill, *Nuclear LIM interactor, a rhombotin and LIM homeodomain interacting protein, is expressed early in neuronal development*. *Proc Natl Acad Sci U S A*, 1996. **93**(21): p. 11693-8.
463. Chang, P.Y., et al., *NFKB1 is a direct target of the TAL1 oncoprotein in human T leukemia cells*. *Cancer Res*, 2006. **66**(12): p. 6008-13.
464. Chagraoui, H., et al., *SCL-mediated regulation of the cell-cycle regulator p21 is critical for murine megakaryopoiesis*. *Blood*, 2011. **118**(3): p. 723-35.
465. Stadanlick, J.E., et al., *Developmental Arrest of T Cells in Rpl22-Deficient Mice Is Dependent upon Multiple p53 Effectors*. *J Immunol*, 2011. **187**(2): p. 664-75.
466. Mendez, J. and B. Stillman, *Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis*. *Mol Cell Biol*, 2000. **20**(22): p. 8602-12.
467. Gladden, A.B. and J.A. Diehl, *The cyclin D1-dependent kinase associates with the pre-replication complex and modulates RB.MCM7 binding*. *J Biol Chem*, 2003. **278**(11): p. 9754-60.
468. Kitamura, T., et al., *Identification and analysis of human erythropoietin receptors on a factor-dependent cell line, TF-1*. *Blood*, 1989. **73**(2): p. 375-80.
469. Bell, S.P. and A. Dutta, *DNA replication in eukaryotic cells*. *Annu Rev Biochem*, 2002. **71**: p. 333-74.
470. Waga, S. and B. Stillman, *Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication in vitro*. *Nature*, 1994. **369**(6477): p. 207-12.
471. Mimura, S., et al., *Central role for cdc45 in establishing an initiation complex of DNA replication in Xenopus egg extracts*. *Genes Cells*, 2000. **5**(6): p. 439-52.
472. Wohlschlegel, J.A., et al., *Inhibition of eukaryotic DNA replication by geminin binding to Cdt1*. *Science*, 2000. **290**(5500): p. 2309-12.
473. Geller, R., R. Levenson, and D. Housman, *Significance of the cell cycle in commitment of murine erythroleukemia cells to erythroid differentiation*. *J Cell Physiol*, 1978. **95**(2): p. 213-22.

474. Hawley, R.G., *High-titer retroviral vectors for efficient transduction of functional genes into murine hematopoietic stem cells*. Ann N Y Acad Sci, 1994. **716**: p. 327-30.
475. De Keersmaecker, K., et al., *The TLX1 oncogene drives aneuploidy in T cell transformation*. Nat Med, 2010. **16**(11): p. 1321-7.
476. Snyder, M., X.Y. Huang, and J.J. Zhang, *The minichromosome maintenance proteins 2-7 (MCM2-7) are necessary for RNA polymerase II (Pol II)-mediated transcription*. J Biol Chem, 2009. **284**(20): p. 13466-72.
477. Snyder, M., W. He, and J.J. Zhang, *The DNA replication factor MCM5 is essential for Stat1-mediated transcriptional activation*. Proc Natl Acad Sci U S A, 2005. **102**(41): p. 14539-44.
478. Kroschl, G., et al., *Transcription factor SCL is required for c-kit expression and c-Kit function in hemopoietic cells*. J Exp Med, 1998. **188**(3): p. 439-50.
479. Zhao, X.F. and P.D. Aplan, *SCL binds the human homologue of DRG in vivo*. Biochim Biophys Acta, 1998. **1448**(1): p. 109-14.
480. Huang, S., et al., *p300 functions as a transcriptional coactivator for the TAL1/SCL oncoprotein*. Oncogene, 1999. **18**(35): p. 4958-67.
481. Huang, S. and S.J. Brandt, *mSin3A regulates murine erythroleukemia cell differentiation through association with the TAL1 (or SCL) transcription factor*. Mol Cell Biol, 2000. **20**(6): p. 2248-59.
482. Mahajan, M.A., S.T. Park, and X.H. Sun, *Association of a novel GTP binding protein, DRG, with TAL oncogenic proteins*. Oncogene, 1996. **12**(11): p. 2343-50.
483. Rual, J.F., et al., *Towards a proteome-scale map of the human protein-protein interaction network*. Nature, 2005. **437**(7062): p. 1173-8.
484. Johannessen, M., et al., *The multifunctional roles of the four-and-a-half-LIM only protein FHL2*. Cell Mol Life Sci, 2006. **63**(3): p. 268-84.
485. Wang, J.C. and J.E. Dick, *Cancer stem cells: lessons from leukemia*. Trends Cell Biol, 2005. **15**(9): p. 494-501.
486. Lane, S.W. and D.G. Gilliland, *Leukemia stem cells*. Semin Cancer Biol, 2010. **20**(2): p. 71-6.
487. Grondin, B., et al., *c-Jun homodimers can function as a context-specific coactivator*. Mol Cell Biol, 2007. **27**(8): p. 2919-33.
488. Trost, M., et al., *The phagosomal proteome in interferon-gamma-activated macrophages*. Immunity, 2009. **30**(1): p. 143-54.
489. Sauvageau, G., et al., *Overexpression of HOXB4 in hematopoietic cells causes the selective expansion of more primitive populations in vitro and in vivo*. Genes Dev, 1995. **9**(14): p. 1753-65.
490. Antonchuk, J., G. Sauvageau, and R.K. Humphries, *HOXB4-induced expansion of adult hematopoietic stem cells ex vivo*. Cell, 2002. **109**(1): p. 39-45.
491. Oshima, M., et al., *Genome-wide analysis of target genes regulated by HoxB4 in hematopoietic stem and progenitor cells developing from embryonic stem cells*. Blood, 2011. **117**(15): p. e142-50.
492. Dubois, N.C., et al., *Placental rescue reveals a sole requirement for c-Myc in embryonic erythroblast survival and hematopoietic stem cell function*. Development, 2008. **135**(14): p. 2455-65.



493. Baena, E., et al., *c-Myc is essential for hematopoietic stem cell differentiation and regulates Lin(-)Sca-1(+)c-Kit(-) cell generation through p21*. Exp Hematol, 2007. **35**(9): p. 1333-43.
494. Wilson, A., et al., *c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation*. Genes Dev, 2004. **18**(22): p. 2747-63.
495. Calvi, L.M., et al., *Osteoblastic cells regulate the haematopoietic stem cell niche*. Nature, 2003. **425**(6960): p. 841-6.
496. Duncan, A.W., et al., *Integration of Notch and Wnt signaling in hematopoietic stem cell maintenance*. Nat Immunol, 2005. **6**(3): p. 314-22.
497. Moore, M.A., et al., *NUP98 dysregulation in myeloid leukemogenesis*. Ann N Y Acad Sci, 2007. **1106**: p. 114-42.
498. Wei, G., et al., *HIPK2 represses beta-catenin-mediated transcription, epidermal stem cell expansion, and skin tumorigenesis*. Proc Natl Acad Sci U S A, 2007. **104**(32): p. 13040-5.
499. Hofmann, T.G., et al., *Regulation of p53 activity by its interaction with homeodomain-interacting protein kinase-2*. Nat Cell Biol, 2002. **4**(1): p. 1-10.
500. D'Orazi, G., et al., *Homeodomain-interacting protein kinase-2 phosphorylates p53 at Ser 46 and mediates apoptosis*. Nat Cell Biol, 2002. **4**(1): p. 11-9.
501. Gresko, E., et al., *PML tumor suppressor is regulated by HIPK2-mediated phosphorylation in response to DNA damage*. Oncogene, 2009. **28**(5): p. 698-708.
502. Aikawa, Y., et al., *Roles of HIPK1 and HIPK2 in AML1- and p300-dependent transcription, hematopoiesis and blood vessel formation*. Embo J, 2006. **25**(17): p. 3955-65.
503. Li, X.L., et al., *Mutations of the HIPK2 gene in acute myeloid leukemia and myelodysplastic syndrome impair AML1- and p53-mediated transcription*. Oncogene, 2007. **26**(51): p. 7231-9.
504. Hui, L., et al., *The oncoprotein LMO3 interacts with calcium- and integrin-binding protein CIB*. Brain Res, 2009. **1265**: p. 24-9.
505. Schaffar, G., et al., *LIM-only protein 4 interacts directly with the repulsive guidance molecule A receptor Neogenin*. J Neurochem, 2008. **107**(2): p. 418-31.
506. Li, H.Y., et al., *Protein-protein interaction of FHL3 with FHL2 and visualization of their interaction by green fluorescent proteins (GFP) two-fusion fluorescence resonance energy transfer (FRET)*. J Cell Biochem, 2001. **80**(3): p. 293-303.
507. Labalette, C., et al., *Interaction and functional cooperation between the LIM protein FHL2, CBP/p300, and beta-catenin*. Mol Cell Biol, 2004. **24**(24): p. 10689-702.
508. McLoughlin, P., et al., *The LIM-only protein DRAL/FHL2 interacts with and is a corepressor for the promyelocytic leukemia zinc finger protein*. J Biol Chem, 2002. **277**(40): p. 37045-53.
509. Purcell, N.H., et al., *Extracellular signal-regulated kinase 2 interacts with and is negatively regulated by the LIM-only protein FHL2 in cardiomyocytes*. Mol Cell Biol, 2004. **24**(3): p. 1081-95.
510. Lin, J., et al., *The four and a half LIM domain protein 2 interacts with and regulates the HERG channel*. FEBS J, 2008. **275**(18): p. 4531-9.

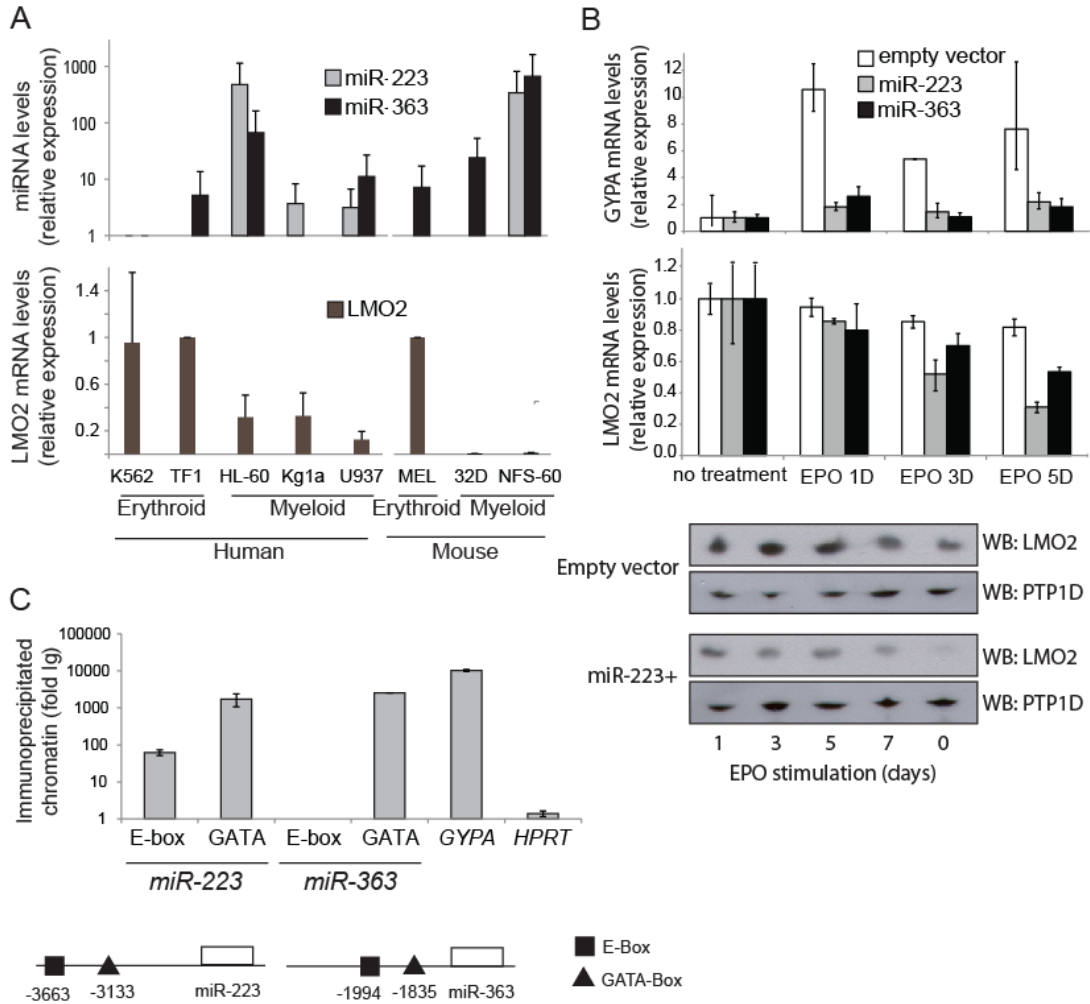
511. Johnson, J.D., et al., *Transcriptional synergy between LIM-homeodomain proteins and basic helix-loop-helix proteins: the LIM2 domain determines specificity*. Mol Cell Biol, 1997. **17**(7): p. 3488-96.
512. Chang, D.F., et al., *Cysteine-rich LIM-only proteins CRP1 and CRP2 are potent smooth muscle differentiation cofactors*. Dev Cell, 2003. **4**(1): p. 107-18.
513. Lee, S.W., E.J. Kim, and S.J. Um, *FHL2 mediates p53-induced transcriptional activation through a direct association with HIPK2*. Biochem Biophys Res Commun, 2006. **339**(4): p. 1056-62.
514. Cao, H., A.R. Sanguinetti, and C.C. Mastick, *Oxidative stress activates both Src-kinases and their negative regulator Csk and induces phosphorylation of two targeting proteins for Csk: caveolin-1 and paxillin*. Exp Cell Res, 2004. **294**(1): p. 159-71.
515. Chan, K.K., et al., *Protein-protein interaction of FHL2, a LIM domain protein preferentially expressed in human heart, with hCDC47*. J Cell Biochem, 2000. **76**(3): p. 499-508.
516. Lee, Y.N., et al., *The function of lysyl-tRNA synthetase and Ap4A as signaling regulators of MITF activity in FcepsilonRI-activated mast cells*. Immunity, 2004. **20**(2): p. 145-51.
517. Lee, Y.N. and E. Razin, *Nonconventional involvement of LysRS in the molecular mechanism of USF2 transcriptional activity in FcepsilonRI-activated mast cells*. Mol Cell Biol, 2005. **25**(20): p. 8904-12.
518. Moore, M.J. and N.J. Proudfoot, *Pre-mRNA processing reaches back to transcription and ahead to translation*. Cell, 2009. **136**(4): p. 688-700.
519. Squires, C.L. and D. Zaporozhets, *Proteins shared by the transcription and translation machines*. Annu Rev Microbiol, 2000. **54**: p. 775-98.
520. Wang, K.C., et al., *A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression*. Nature, 2011. **472**(7341): p. 120-4.
521. Huarte, M., et al., *A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response*. Cell, 2010. **142**(3): p. 409-19.
522. Tsai, M.C., et al., *Long noncoding RNA as modular scaffold of histone modification complexes*. Science, 2010. **329**(5992): p. 689-93.
523. Khalil, A.M., et al., *Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression*. Proc Natl Acad Sci U S A, 2009. **106**(28): p. 11667-72.
524. Miyamoto, K., et al., *Foxo3a is essential for maintenance of the hematopoietic stem cell pool*. Cell Stem Cell, 2007. **1**(1): p. 101-12.
525. Iizuka, M. and B. Stillman, *Histone acetyltransferase HBO1 interacts with the ORC1 subunit of the human initiator protein*. J Biol Chem, 1999. **274**(33): p. 23027-34.
526. Burke, T.W., et al., *Replication factors MCM2 and ORC1 interact with the histone acetyltransferase HBO1*. J Biol Chem, 2001. **276**(18): p. 15397-408.
527. Iizuka, M., et al., *Regulation of replication licensing by acetyltransferase Hbo1*. Mol Cell Biol, 2006. **26**(3): p. 1098-108.
528. Mishima, Y., et al., *The Hbo1-Brd1/Brpf2 complex is responsible for global acetylation of H3K14 and required for fetal liver erythropoiesis*. Blood, 2011.

529. Jorgensen, S., et al., *The histone methyltransferase SET8 is required for S-phase progression*. J Cell Biol, 2007. **179**(7): p. 1337-45.
530. Huen, M.S., et al., *Direct interaction between SET8 and proliferating cell nuclear antigen couples H4-K20 methylation with DNA replication*. J Biol Chem, 2008. **283**(17): p. 11073-7.
531. Tardat, M., et al., *PR-Set7-dependent lysine methylation ensures genome replication and stability through S phase*. J Cell Biol, 2007. **179**(7): p. 1413-26.
532. Collins, N., et al., *An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin*. Nat Genet, 2002. **32**(4): p. 627-32.
533. Esteve, P.O., et al., *Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication*. Genes Dev, 2006. **20**(22): p. 3089-103.
534. Uemura, T., et al., *Temporal and spatial localization of novel nuclear protein NP95 in mitotic and meiotic cells*. Cell Struct Funct, 2000. **25**(3): p. 149-59.
535. Miura, M., et al., *Dynamic changes in subnuclear NP95 location during the cell cycle and its spatial relationship with DNA replication foci*. Exp Cell Res, 2001. **263**(2): p. 202-8.
536. Bonapace, I.M., et al., *Np95 is regulated by E1A during mitotic reactivation of terminally differentiated cells and is essential for S phase entry*. J Cell Biol, 2002. **157**(6): p. 909-14.
537. Papait, R., et al., *Np95 is implicated in pericentromeric heterochromatin replication and in major satellite silencing*. Mol Biol Cell, 2007. **18**(3): p. 1098-106.
538. Sharif, J., et al., *The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA*. Nature, 2007. **450**(7171): p. 908-12.
539. Fujimori, A., et al., *Cloning and mapping of Np95 gene which encodes a novel nuclear protein associated with cell proliferation*. Mamm Genome, 1998. **9**(12): p. 1032-5.
540. Liu, L., et al., *Identification of a novel protein, PDIP38, that interacts with the p50 subunit of DNA polymerase delta and proliferating cell nuclear antigen*. J Biol Chem, 2003. **278**(12): p. 10041-7.
541. Bartkova, J., et al., *DNA damage response as a candidate anti-cancer barrier in early human tumorigenesis*. Nature, 2005. **434**(7035): p. 864-70.
542. Gorgoulis, V.G., et al., *Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions*. Nature, 2005. **434**(7035): p. 907-13.
543. Karakaidos, P., et al., *Overexpression of the replication licensing regulators hCdt1 and hCdc6 characterizes a subset of non-small-cell lung carcinomas: synergistic effect with mutant p53 on tumor growth and chromosomal instability--evidence of E2F-1 transcriptional control over hCdt1*. Am J Pathol, 2004. **165**(4): p. 1351-65.
544. Vaziri, C., et al., *A p53-dependent checkpoint pathway prevents rereplication*. Mol Cell, 2003. **11**(4): p. 997-1008.

545. Williams, G.H., et al., *Improved cervical smear assessment using antibodies against proteins that regulate DNA replication*. Proc Natl Acad Sci U S A, 1998. **95**(25): p. 14932-7.
546. Gonzalez, M.A., et al., *Control of DNA replication and its potential clinical exploitation*. Nat Rev Cancer, 2005. **5**(2): p. 135-41.
547. Lei, M., Y. Kawasaki, and B.K. Tye, *Physical interactions among Mcm proteins and effects of Mcm dosage on DNA replication in Saccharomyces cerevisiae*. Mol Cell Biol, 1996. **16**(9): p. 5081-90.
548. Liang, D.T., J.A. Hodson, and S.L. Forsburg, *Reduced dosage of a single fission yeast MCM protein causes genetic instability and S phase delay*. J Cell Sci, 1999. **112 ( Pt 4)**: p. 559-67.
549. Lengronne, A. and E. Schwob, *The yeast CDK inhibitor Sic1 prevents genomic instability by promoting replication origin licensing in late G(1)*. Mol Cell, 2002. **9**(5): p. 1067-78.
550. Tanaka, S. and J.F. Diffley, *Deregulated G1-cyclin expression induces genomic instability by preventing efficient pre-RC formation*. Genes Dev, 2002. **16**(20): p. 2639-49.
551. Ekholm-Reed, S., et al., *Deregulation of cyclin E in human cells interferes with prereplication complex assembly*. J Cell Biol, 2004. **165**(6): p. 789-800.
552. Spruck, C.H., K.A. Won, and S.I. Reed, *Deregulated cyclin E induces chromosome instability*. Nature, 1999. **401**(6750): p. 297-300.

# Annexe 1

## Boucle d'auto-régulation entre LMO2 et deux microARNs



**Figure A.1 Boucle d'auto-régulation entre LMO2 et deux microARNs**

A) Les niveaux transcriptionnels de LMO2 (bas) sont inversement corrélés aux niveaux d'expression de miR-223 et de miR-363 (haut) dans différentes lignées hématopoïétiques. L'ARN total, incluant les petits ARNs, a été extrait et soumis au RT-PCR. Les niveaux d'ARN ont été normalisés par rapport au gène contrôle *Hprt* (pour les niveaux de LMO2) ou *U6* (pour les niveaux de microARNs). B) La surexpression de miR-223 ou de miR-363 compromet la différenciation érythrocytaire des cellules TF-1 (haut) en faisant diminuer les niveaux de LMO2 (centre et bas). La différenciation érythrocytaire est induite par traitement à l'érythropoïétine (EPO) pour les temps indiqués en jours (D). L'expression de la glycophorine A (*GYP A*) par RT-PCR (haut) est employée comme marqueur de différenciation. L'expression de LMO2 a été mesurée par RT-PCR (centre) et par Western blot (bas). Les Western blot sont représentatifs de deux expériences indépendantes. C) Haut : LMO2 s'associe aux promoteurs de *miR-223* et de *miR-363*. Immunoprécipitation de la chromatine (ChIP) de cellules TF-1 avec un anticorps contre LMO2. Le gène *GYP A* est une cible transcriptionnelle connue de LMO2 employée comme contrôle positif, alors que le gène *HPRT* est un contrôle négatif. Bas : Schéma représentant les promoteurs de *miR-223* et *miR-363*. Les sites consensus de liaison de LMO2 sont indiqués. Les données de RT-PCR et de ChIP représentent la moyenne de deux expériences indépendantes effectuées en duplicata. Cette figure est tirée d'un manuscrit en préparation et a été adaptée avec la permission de tous les co-auteurs (Véronique Lisi, Marie-Claude Sincennes, Trang Hoang, François Major – Institut de Recherches en Immunologie et Cancer, Montréal).

## Annexe 2

### **Protein Stability and Transcription Factor Complex Assembly Determined by the SCL-LMO2 Interaction**

Eric Lécuyer<sup>1</sup>, Simon Larivière<sup>1</sup>, Marie-Claude Sincennes<sup>1</sup>, André Haman<sup>1</sup>, Rachid Lahli<sup>1</sup>, Margarita Todorova<sup>1</sup>, Mathieu Tremblay<sup>1</sup>, Brian C. Wilkes<sup>2</sup>, Trang Hoang<sup>1,3</sup>

<sup>1</sup> Institute of Research in Immunology and Cancer (IRIC) – University of Montreal, Montréal, Québec, Canada H3C 3J7

<sup>2</sup> Institut de Recherche Clinique de Montréal, Montréal, Québec, Canada, H2W 1R7

<sup>3</sup> Pharmacology, Biochemistry and Molecular Biology Departments, University of Montréal, Québec, Canada H3C 3J7

This research was originally published in The Journal of Biological Chemistry.

Eric Lécuyer, Simon Larivière, Marie-Claude Sincennes, André Haman, Rachid Lahli, Margarita Todorova, Mathieu Tremblay, Brian C. Wilkes, Trang Hoang. Protein Stability and Transcription Factor Complex Assembly Determined by the SCL-LMO2 Interaction. *The Journal of Biological Chemistry*. 2007; 282:33649-33658.

© the American Society for Biochemistry and Molecular Biology.

Supplemental Material can be found at:  
<http://www.jbc.org/content/suppl/2007/09/18/M703939200.DC1.html>

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 46, PP. 33649–33658, NOVEMBER 16, 2007  
 © 2007 BY THE AMERICAN SOCIETY FOR BIOCHEMISTRY AND MOLECULAR BIOLOGY, INC. PRINTED IN THE U.S.A.

## Protein Stability and Transcription Factor Complex Assembly Determined by the SCL-LMO2 Interaction<sup>\*[5]</sup>

Received for publication, May 14, 2007, and in revised form, August 20, 2007. Published, JBC Papers in Press, September 17, 2007, DOI 10.1074/jbc.M703939200

Eric Lécuyer<sup>†§1,2</sup>, Simon Larivière<sup>†§</sup>, Marie-Claude Sincennes<sup>†§1</sup>, André Haman<sup>†</sup>, Rachid Lahlii<sup>†</sup>, Margarita Todorova<sup>†§</sup>, Mathieu Tremblay<sup>†§1</sup>, Brian C. Wilkes<sup>†</sup>, and Trang Hoang<sup>†§||\*\*3,4</sup>

From the <sup>†</sup>Institut de Recherche en Immunologie et Cancérologie and the Departments of <sup>§</sup>Molecular Biology, <sup>||</sup>Pharmacology, and <sup>\*\*</sup>Biochemistry, Université de Montréal, Montréal, Québec H3C 2J7, Canada and the <sup>1</sup>Institut de Recherche Clinique de Montréal, Montréal, Québec H2W 1R7, Canada

Gene expression programs are established by networks of interacting transcription factors. The basic helix-loop-helix factor SCL and the LIM-only protein LMO2 are components of transcription factor complexes that are essential for hematopoiesis. Here we show that LMO2 and SCL are predominant interaction partners in hematopoietic cells and that this interaction occurs through a conserved interface residing in the loop and helix 2 of SCL. This interaction nucleates the assembly of SCL complexes on DNA and is required for target gene induction and for the stimulation of erythroid and megakaryocytic differentiation. We also demonstrate that SCL determines LMO2 protein levels in hematopoietic cells and reveal that interaction with SCL prevents LMO2 degradation by the proteasome. We propose that the SCL-LMO2 interaction couples protein stabilization with higher order protein complex assembly, thus providing a powerful means of modulating the stoichiometry and spatiotemporal activity of SCL complexes. This interaction likely provides a rate-limiting step in the transcriptional control of hematopoiesis and leukemia, and similar mechanisms may operate to control the assembly of diverse protein modules.

Tissue-specific programs of gene expression are largely controlled by networks of interacting transcription factors that provide flexibility and specificity in gene regulation (1, 2). This theme finds particular relevance in the context of the hematopoietic system, where the generation of diverse cellular lineages is controlled by combinations of hematopoietic-specific and broadly expressed transcription factors (3–5). Regulation by networking is exemplified by the SCL/TAL-1 (hereafter referred to as SCL) transcription factor, which is essential for

the establishment of the hematopoietic system and plays important functions at several branch points in the hematopoietic hierarchy (6).

SCL is a member of a subfamily of tissue-specific bHLH transcription factors that includes two other hematopoietic factors, TAL2 and LYL-1 (7), and two neurogenic factors, NSCL1/nHLH-1 and NSCL2 (8). Transcription regulation by SCL requires its integration within multifactorial complexes (SCL complexes) containing the ubiquitously expressed bHLH factors encoded by the *E2A* gene (*E47* and *E12*), LMO<sup>5</sup> proteins, the LIM domain-binding protein Ldb1, and hematopoietic GATA family members (9–11). These complexes govern the switch between a proliferative state in erythroid progenitors and a commitment to terminal differentiation (12), and different variants of these complexes activate transcription of target genes such as *c-kit*, glycophorin A (*GPA*), and protein 4.2 in progenitor and erythroid cells (13–15). These observations are consistent with the view that subtle variations in the composition of transcription factor complexes modulate choices in cell fate. Interestingly, SCL complexes exhibit an all-or-nothing behavior in transcription activation requiring the simultaneous presence of each partner for robust target gene activation (13, 15). These complexes have, therefore, been proposed to form multi-input motifs within the hematopoietic regulatory hierarchy (5), a network feature that ensures specificity and flexibility in gene regulation. Despite these findings, the mechanisms governing the assembly of SCL-containing complexes on target gene regulatory elements remain ill-defined.

It has recently been shown that the SCL-LMO2 interaction is essential for hematopoietic cell fate specification (16–18), consistent with the phenotypic similarity of *SCL*<sup>-/-</sup> and *LMO2*<sup>-/-</sup> embryos, which exhibit early lethality due to a complete absence of blood cells (19–24). In addition to their important function during normal hematopoiesis, the *SCL* and *LMO2* genes are the most frequent targets of chromosomal rearrangements in pediatric T-cell acute lymphoblastic leukemia (7, 25).

\* This work was funded by grants from the Canadian Institutes of Health Research and the National Cancer Institute of Canada. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[5] The on-line version of this article (available at <http://www.jbc.org>) contains supplemental Figs. S1–S3 and Table S1.

<sup>1</sup> Recipients of Canadian Institutes of Health Research studentships.

<sup>2</sup> Present address: Terrence Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, Toronto, Ontario M5S 3E1 Canada.

<sup>3</sup> Recipient of the Canada Research Chair in Cell Differentiation and the Genetics of Acute Leukemias.

<sup>4</sup> To whom correspondence should be addressed: Institut de Recherche en Immunologie et Cancérologie, Université de Montréal, P. O. Box 6128, Downtown Branch, Montréal, Québec H3C 2J7, Canada. Tel.: 514-343-6970; Fax: 514-987-5757; E-mail: Trang.hoang@umontreal.ca.

<sup>5</sup> The abbreviations used are: BFU-E, burst-forming unit erythroid; bHLH, basic helix-loop-helix; CFU-GM/Meg, colony-forming unit granulocyte-macrophage/megakaryocyte; EMSA, electrophoretic mobility shift assay; *Épor*, erythropoietin receptor gene; F-LMO2, FLAG-LMO2; GFP, green fluorescence protein; GPA, glycophorin A; IB, immunoblotting; IP, immunoprecipitation; IRES, internal ribosome entry site; Ldb1, LIM domain-binding protein-1; LIM, Lin-Is1-Mec; LMO, LIM-only; MSCV, murine stem cell virus; NSCL1, neuronal SCL-1; SCL/TAL1, stem cell leukemia/T-cell acute leukemia-1; IF, immunofluorescence; PBS, phosphate-buffered saline; HA, hemagglutinin; RT, reverse transcription; GST, glutathione S-transferase.



### Molecular Functions of the SCL-LMO2 Interaction

Furthermore, retroviral integration within the *LMO2* gene, leading to a severe lymphoproliferative disorder, has been noted after gene therapy in one clinical trial (26). Finally, *SCL* genetically interacts with *LMO1/2* to induce aggressive T-cell tumors in transgenic mice (7), underscoring the importance of the SCL-LMO2 interaction *in vivo*.

Despite this extensive body of evidence conveying the biological importance of the SCL-LMO2 interaction, it is not clear why this interaction is important at the molecular level. In addition, because LIM domain proteins, including LMO2, are known to be targeted for proteasomal degradation (27, 28), this raises the possibility that LMO2 degradation may represent a rate-limiting step in the assembly of SCL complexes. In the present study we show that SCL prevents the degradation of LMO2 via direct protein-protein interaction, allowing for the nucleation of multifactorial complexes with proper SCL/LMO2 stoichiometry.

### EXPERIMENTAL PROCEDURES

**Plasmids and Mutagenesis**—A detailed description of plasmid construction methodology is provided as supplemental material.

**Cell Culture, Infections, Transfections, Immunofluorescence (IF), and Protein Extraction**—NIH 3T3, 293, *ts20*, and TF-1 cells were cultured as described previously (15, 29, 30). Primary fetal liver cells were isolated from E12.5 *Epor*<sup>-/-</sup> and *SCL*<sup>lacZ/WT</sup> embryos, sorted for TER119 expression, and cultured as described before (15). Fetal liver cell infections were performed by overnight co-culture on GP+E packaging cells stably expressing the viruses encoding SCL and SCL-M13. After infection, the cells were cultured in the presence of interleukin-3 (50 ng/ml) and steel factor (5 ng/ml) for an additional 24 h before performing further analyses. Infections of TF-1 cells with control retroviruses (MSCV-neo) or viruses encoding SCL and SCL-M13 were performed as detailed previously (15). Transient transactivation assays were performed as previously described (13, 15). Expression vector doses used in specific experiments are indicated in figure legends. In all samples the amount of total DNA was kept constant at 4.5  $\mu$ g with pGem4, and 100 ng of the cytomegalovirus- $\beta$ -galactosidase vector was included to normalize luciferase values.

For IF analysis, TF-1 or transfected 293 cells were harvested, washed in PBS, and fixed in 500  $\mu$ l of Bouin's fixative for 15 min at room temperature. Cells were then washed several times in PBS by successive centrifugations and spun onto glass slides. In a humidified chamber at room temperature, the samples were blocked with PBS plus 15% normal goat serum for 15 min, washed in PBS, and incubated for 2 h with primary mouse anti-SCL and goat anti-LMO2 or anti-Ldb1 antibodies (see description below). After washing the samples 5 $\times$  for 10 min in PBS, they were incubated for 2 h with secondary donkey anti-mouse and anti-goat antibodies coupled to fluorescein isothiocyanate and Cy3, respectively. All antibodies were used at a 1/100 dilution in PBS. Cells were then counter-stained with Hoechst for 1 min, washed, and resuspended in Vectashield (Vector Laboratories Inc, Burlington, ON, Canada) before analysis on an LSM 510 laser scanning confocal system (Carl Zeiss Inc., Maple Grove, MN).

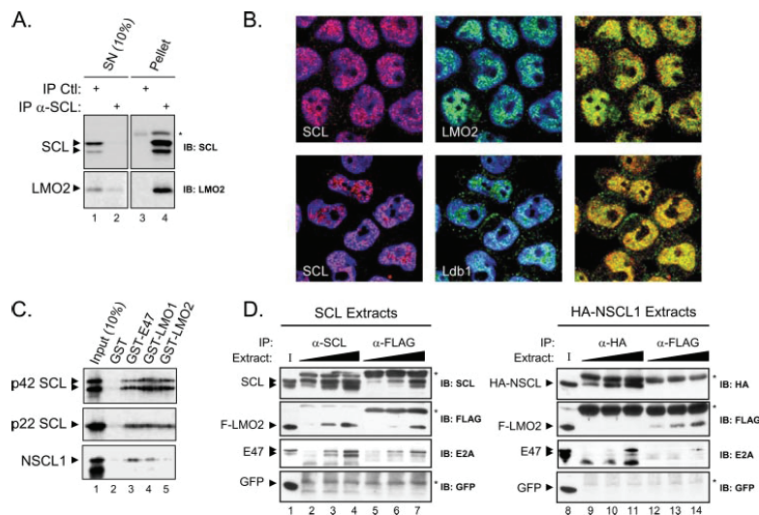
TF-1 and 293 cell nuclear extracts were prepared as documented previously (15). Protein extracts from primary fetal liver from wild type and *SCL*<sup>lacZ/WT</sup> embryos were prepared as detailed before (31). For 293 cell extracts used in immunoprecipitations (IPs) and electrophoretic mobility shift assays (EMSA), 4.2 $\times$ 10<sup>6</sup> cells were seeded in 100-mm plates and transfected with combinations of expression vectors (5  $\mu$ g of each) as described in the figure legends. The total amount of DNA was kept at 25–30  $\mu$ g using pGem4. For studies of LMO2 protein expression/stability in 293 cells, 7 $\times$ 10<sup>5</sup> cells/well were seeded in 6-well plates and transfected with 1  $\mu$ g of each expression vector indicated in figure legends. The total amount of DNA was kept at 5  $\mu$ g with pGem4. Proteasomal inhibitor (MG132, lactacystin) or cycloheximide treatments were carried out as described in figure legends from 4 h to overnight. After treatment, the cells were washed twice with cold PBS and instantaneously frozen on liquid nitrogen. Total cell protein extracts were then prepared by lysing the cells for 20 min at 4  $^{\circ}$ C with 400  $\mu$ l of radioimmunoprecipitation assay buffer (10 mM Tris-HCl (pH 8.0), 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% SDS, 0.1% deoxycholate) supplemented with protease inhibitors.

**EMSA, Pulldown Assays, IBs, and IPs**—For EMSA, each binding reaction was performed using 10  $\mu$ g of TF-1 or 293 cell nuclear extract in 20 mM HEPES (pH 7.5), 50 mM KCl, 1 mM dithiothreitol, 1 mM EDTA, 5% glycerol, 10  $\mu$ g bovine serum albumin, 0.5  $\mu$ g of dI-dC, and the GPA-84 probe (10,000 cpm) in a final volume of 20  $\mu$ l. After 15 min at room temperature, protein-DNA complexes were resolved at 150 V on a 4% PAGE gel in 0.5 $\times$  Tris-buffered EDTA at 4  $^{\circ}$ C for 4 h. Pulldown experiments were performed as previously described (13, 15). For IPs, 293 (10–60  $\mu$ g) or TF-1 (500  $\mu$ g) cell nuclear extracts were incubated overnight at 4  $^{\circ}$ C with 3  $\mu$ g of antibody in 1 ml of IP buffer (20 mM Tris-HCl (pH 8.0), 137 mM NaCl, 1% Nonidet P-40, 10% glycerol, 1 mM EDTA). Protein complexes were precipitated by adding appropriately conjugated Pansorbin cells (Calbiochem) for 30–120 min at 4  $^{\circ}$ C, washed 3 times with 1 ml of IP buffer, and subjected to SDS-PAGE. After transfer on PVDF membranes, proteins were visualized by immunoblotting (IB) using ECL plus (GE Healthcare). The following antibodies were used for IP, IB, and IF analysis. The mouse anti-E2A (YAE), rat anti-GATA-1 (N6), rabbit anti-GFP (FL), goat anti-LMO2 (N-16), and goat anti-Ldb1/CLIM-2 (N-18) were all from Santa Cruz Biotechnology Inc (Santa-Cruz, CA). The BTL-73 and 2TL-136 mouse anti-SCL antisera were provided by Dr. D. Mathieu (Institut de Génétique Moléculaire, Montpellier, France). The mouse anti-HA, anti-FLAG, and anti-P21 antibodies were obtained from Covance (Richmond, VA), Stratagene (La Jolla, CA), and BD Biosciences (Mississauga, ON, Canada), respectively.

**RT-PCR Analysis**—Total RNA was extracted from TF-1 or 293 cells as previously described (15). To eliminate contaminant DNA molecules, 500 ng of each sample was then subjected to digestion with EcoRI and DNase I in REact3 buffer (Invitrogen). After the nucleases were heat-inactivated for 10 min at 65  $^{\circ}$ C, the samples were subjected to reverse transcription using the Superscript first strand cDNA synthesis system (Invitrogen). The quantification of GPA,  $\beta$ -major, LMO2, and S14 and

Supplemental Material can be found at:  
<http://www.jbc.org/content/suppl/2007/09/18/M703939200.DC1.html>

### Molecular Functions of the SCL-LMO2 Interaction



**FIGURE 1. SCL predominantly and specifically interacts with LMO2.** *A*, in hematopoietic progenitor cells most of LMO2 is associated with SCL. TF-1 cell nuclear extracts (500  $\mu$ g) were subjected to IP using an anti-SCL antibody or control immunoglobulins. Samples of the precipitated (*Pellet*) or unbound supernatant (*SN*) fractions were analyzed by IB as indicated. The asterisk indicates Ig heavy chains. *B*, SCL co-localizes with LMO2 and Ldb1 in the nuclei of hematopoietic cells. TF-1 cells were double-stained with antibodies directed against SCL and LMO2 (*upper panels*) or with SCL and Ldb1 (*lower panels*). Red, SCL; green, LMO2 or Ldb1; blue, Hoechst-stained nuclei. C, SCL but not NSCL1, interacts with LMO2 *in vitro*. Pull-down experiments were performed with immobilized GST, GST-LMO1, GST-LMO2, GST-E47, and  $^{35}$ S-labeled SCL and NSCL1. Input samples (*lane 1*) represent 10% of the amount used for the binding reactions (*lanes 2–5*). *D*, SCL but not NSCL1 interacts with LMO2 in transfected cells. Extracts of 293 cells (10, 30, and 60  $\mu$ g) transfected with expression vectors (5  $\mu$ g each) for E47, F-LMO2, Ldb1, GFP, and SCL (*lanes 1–7*) or HA-NSCL1 (*lanes 8–14*) were immunoprecipitated with the indicated antibodies and revealed by IB with antibodies listed to the right of each panel. *Lane 1* represents input (*I*) samples (10  $\mu$ g), and arrowheads point to specific bands. Asterisks indicate Ig light or heavy chains.

S16 mRNAs was performed by real time PCR on a MX3000 apparatus (Stratagene, La Jolla, CA) using QuantiTect SYBR green PCR kit (Qiagen, Mississauga, ON, Canada) according to the manufacturer's instructions and under the following conditions: 95  $^{\circ}$ C for 15 min followed by 50 cycles of 94  $^{\circ}$ C for 30 s, 57  $^{\circ}$ C for 30 s, and 72  $^{\circ}$ C for 30 s. Oligonucleotides used for RT-PCR analysis were described previously (15).

**Structural Modeling**—The structure of the E47 homodimer-DNA complex (32) was obtained from the Brookhaven protein data base and was subjected to 1000 steps of conjugate gradient minimization. One of the E47 moieties was transformed into human SCL using homology modeling techniques. After the structure of the DNA bound SCL/E47 heterodimer was obtained, the side chains of the mutated amino acids were packed using SYBYL (Tripos, St. Louis, MO). In this procedure the backbone dihedral angles are held fixed, whereas the side chains of the individual amino acids are rotated one at a time until a sterically acceptable conformation was obtained. This structure was again minimized, at which time no major conformational changes were observed.

## RESULTS

**SCL and LMO2 Are Predominant and Specific Interaction Partners**—The SCL and LMO2 interaction is essential for hematopoiesis. To determine what proportion of LMO2 is

associated with SCL in hematopoietic cells, we performed quantitative IPs of nuclear extracts of CD34+ TF-1 cells with an antibody against SCL or control immunoglobulins. As shown in Fig. 1A, all of the SCL protein was precipitated with an anti-SCL antibody (*lanes 2 and 4*), whereas the protein remained in the supernatant with a control anti-serum (*lanes 1 and 3*). Under these conditions, most of LMO2 co-precipitated with SCL, and only 20% of LMO2 was found in the supernatant (*lane 2*). Furthermore, by performing successive SCL and Ldb1 IPs, we found that the majority LMO2 that did not come down with the SCL antibody was co-precipitated with anti-Ldb1 antisera (data not shown). In contrast, all of LMO2 was recovered in the supernatant of control IPs (*lane 1*). Therefore, we conclude that the majority (~90%) of LMO2 is associated with SCL and Ldb1 in TF-1 cell extracts. In agreement with these findings, all three of these factors were found to exhibit overlapping nucleoplasmic localization patterns, as determined by IF analysis in TF-1 cells (Fig. 1B). Together, our observations concur with the view that SCL, LMO2, and

Ldb1 are predominant interaction partners in hematopoietic cells.

SCL has recently been shown to exhibit exquisite specificity in comparison to other bHLH factors in inducing hematopoietic cell fate commitment in embryonic stem cells (17). SCL shares 64% identity within the bHLH domain with the closely related neurogenic factor NSCL1, and both proteins heterodimerize with the ubiquitous bHLH factor E47. Furthermore, LMO2, which is 59% identical to LMO1 within the LIM domains, interacts with SCL but not E47. To assess the specificity of the SCL-LMO2 interaction, we compared the binding properties of SCL and NSCL1 in pull-down assays. As expected, both SCL and NSCL1 interact with GST-E47 (Fig. 1C). Interaction with LMO proteins, however, revealed a striking difference. Both of the main isoforms of SCL (p42 and p22) associate equally well with LMO1 and LMO2. In contrast, NSCL1 was only weakly retained on GST-LMO1 columns and was not retained on GST-LMO2 columns. These results indicate that interaction with LMO2 is specific to SCL.

To ascertain whether this interaction specificity is also observed in transfected cells, we next performed IPs using nuclear extracts from 293 cells expressing SCL or HA-NSCL1, in combination with FLAG-LMO2 (F-LMO2), E47, and Ldb1. Increasing amounts of nuclear extracts were subjected to IP using anti-SCL, anti-FLAG, or anti-HA antibodies. In SCL-containing extracts, IP with both the anti-SCL and anti-FLAG

Supplemental Material can be found at:  
<http://www.jbc.org/content/suppl/2007/09/18/M70393200.DC1.html>

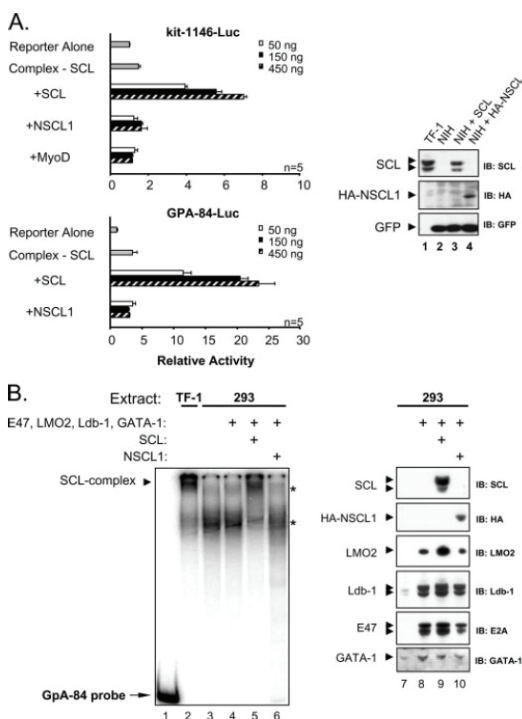
## Molecular Functions of the SCL-LMO2 Interaction

antibodies efficiently precipitated SCL, F-LMO2, and E47 (Fig. 1D, lanes 1–7), showing that these factors associate in transfected cells. With HA-NSCL1 extracts, however, the anti-HA antibody brought down HA-NSCL1 and E47 but was unable to co-precipitate F-LMO2 (lanes 9–11). In agreement, the anti-FLAG antibody precipitated F-LMO2 but not HA-NSCL1 or E47 (lanes 12–14). Finally, co-expressed GFP protein, included as a negative control, was not precipitated with any of the antibodies. Because LMO2 and E47 only co-precipitate in the presence of SCL, we conclude that SCL acts as a bridging factor between its two partners. More importantly, whereas both SCL and NSCL1 are able to associate with E47, interaction with LMO2 is restricted to SCL.

**Specificity of SCL in Hematopoietic Target Gene Activation and Assembly of Transcription Factor Complexes**—To determine whether SCL also demonstrates specificity at the molecular level during hematopoietic target gene activation, we next performed transactivation assays of the *c-kit* and GPA promoters, two well characterized targets of SCL-containing complexes (13, 15). This approach previously revealed an essential requirement for each partner of the SCL complex in transcriptional regulation (13, 15). As suspected, although SCL-containing complexes efficiently activate the *c-kit* and GPA promoters, neither the neurogenic bHLH factor NSCL1 nor the myogenic factor MyoD is able to substitute for SCL function (Fig. 2A, left panels) despite being efficiently expressed in transfected cells (Fig. 2A, right panel). In contrast to SCL, other partners of the complex are functionally redundant with members of their respective families (supplemental Fig. S1).

To assess whether SCL is specifically required for the assembly of SCL complexes on DNA, we next performed EMSA using a probe derived from the GPA promoter (GPA-84), to which the SCL complex was previously shown to bind with high affinity (15). As shown in Fig. 2B, SCL and its partners, expressed in 293 cells, form a low mobility complex (arrowhead) on the GPA-84 probe (lane 5), which is comparable with that seen with endogenous proteins from hematopoietic TF-1 cells (lane 2) and is distinct from nonspecific complexes (asterisks) observed with untransfected 293 cell extracts (lane 3). This low mobility complex is supershifted by antibodies against SCL and its partners (data not shown; Ref. 13) and is not seen with 293 extracts lacking SCL but expressing the other partners of the complex (lane 4). Interestingly, NSCL1 was unable to substitute for SCL in nucleating the assembly of this complex on DNA (lane 6), in agreement with transactivation experiments. The expression of SCL and its partners as well as NSCL1 in 293 cell extracts was confirmed by IB (lanes 7–11). We conclude that SCL confers specificity in hematopoietic target gene activation and in the assembly SCL complexes on DNA.

**SCL Interacts with LMO2 via Multiple Residues in the Loop and Helix 2**—Schlaeger *et al.* (17) previously identified residues of the HLH domain that are critical for interaction with LMO2 and for the rescue of primitive erythropoiesis in *SCL*<sup>-/-</sup> ES cells. These include residues His-217 of the loop region and Phe-238 residing in Helix 2, which was also recently shown to be required for the hematopoietic inducing potential of SCL in zebrafish (18). It is not known, however, whether the same residues are required for the other molecular and biological func-



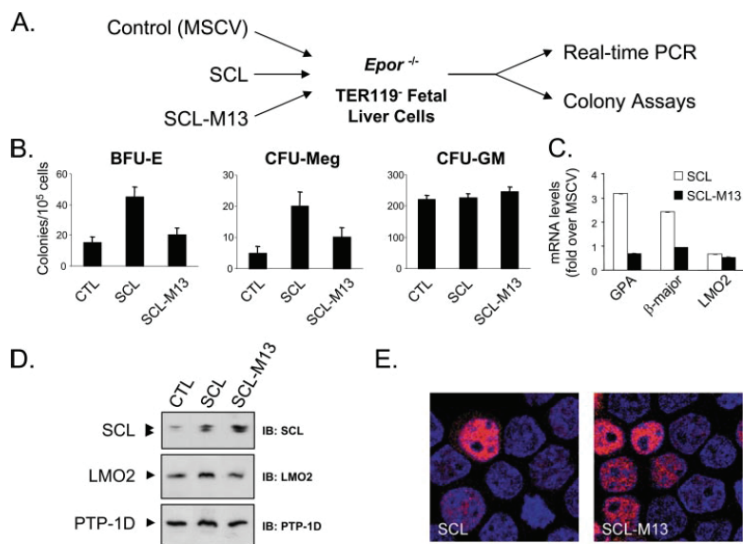
**FIGURE 2. Non-redundant function of SCL in hematopoietic gene activation and complex formation on DNA.** A, SCL is non-redundant with other bHLH factors in *c-kit* and GPA promoter activation. The kit-1146-Luc or GPA-84-Luc reporter constructs (1500 ng) were transfected into NIH 3T3 cells with expression vectors for E47 (150 ng), LMO2 (750 ng), Ldb1 (750 ng), and GATA-1 or -2 (150 ng) and the indicated amounts of SCL, NSCL1, and MyoD vectors (50–450 ng). Transactivation assays with kit-1146-Luc were conducted using GATA-1-containing complexes (upper graph), whereas GPA-84-Luc was assayed with complexes containing GATA-2 (lower graph). Results are expressed as fold activation relative to the reporter vector transfected alone, represent the average  $\pm$  S.D. of triplicate determinations, and are representative of (n) independent experiments. Right panel, nuclear extracts of TF-1 cells (10  $\mu$ g) or NIH 3T3 cells (10  $\mu$ g) expressing SCL or HA-NSCL1 were analyzed by IB with the indicated antibodies. Expression of GFP in transfected cells was monitored as a control for equal loading. Arrowheads indicate specific bands. B, SCL is specifically required to assemble a high molecular weight complex on the GPA promoter sequences. EMSAs were performed using the GPA-84 probe (arrow), which covers positions -84 to -30 of the human GPA promoter (15), and nuclear extracts (10  $\mu$ g) of TF-1 cells (lane 2) or 293 cells transfected with the indicated expression vectors (lanes 3–6). Note the low mobility complex containing SCL and its partners (arrowhead) (15), which is not formed in the presence of NSCL1. Asterisks indicate nonspecific complexes. Proteins expressed in the 293 cell extracts (10  $\mu$ g) used for EMSA were revealed by IB with the indicated antibodies (lanes 7–10).

tions of SCL, *i.e.* the induction of definitive erythropoiesis and the inhibition of lymphoid lineages. Therefore, to identify the determinants of SCL molecular specificity, based on comparisons with NSCL1, we first generated a series of SCL-NSCL1 domain swap mutants and tested their functional activity in transactivation assays. As expected, this approach revealed that specificity in transcriptional activation and interaction with LMO2 is imparted by the loop and helix 2 region of SCL (supplemental Fig. S2).



Supplemental Material can be found at:  
<http://www.jbc.org/content/suppl/2007/09/18/M703939200.DC1.html>

### Molecular Functions of the SCL-LMO2 Interaction



**FIGURE 4. The SCL-LMO2 interaction is required for the stimulation of erythroid differentiation and endogenous GPA gene induction by SCL.** *A*, day 12.5 *Epor*<sup>-/-</sup> mice TER119<sup>+</sup> fetal liver cells, which are blocked at the pro-erythroblast stage, were infected with control retroviruses (CTL) or viruses encoding SCL or SCL-M13. Cells were assessed for the frequency of the different types of progenitors in methyl cellulose assays. In parallel, gene expression was assessed by real time RT-PCR analysis. *B*, enforced expression of SCL stimulates the formation of erythroid (BFU-E) and megakaryocytic (CFU-Meg) progenitors, whereas no effect was observed on the number of myeloid (CFU-GM) progenitors. In contrast, transduction of the SCL-M13 mutant has no effect on the number of BFU-E and CFU-Meg progenitors. *C*, the SCL-LMO2 interaction is required for the induction of the endogenous GPA gene in primary fetal liver cells. Endogenous GPA mRNA levels were analyzed by real time RT-PCR. SCL induces a 4-fold increase in GPA mRNA levels, whereas no effect was observed with SCL-M13. S16 was used as a normalization control. *D*, SCL and SCL-M13 were efficiently expressed in infected hematopoietic cells. TF-1 cells were infected in parallel with the same viral stocks as primary *Epor*<sup>-/-</sup> cells, and nuclear extracts (10  $\mu$ g) were analyzed by IB with the indicated antibodies. The limited number of infected primary *Epor*<sup>-/-</sup> fetal liver cells precludes analysis by IB. *E*, SCL and SCL-M13 exhibit similar nucleoplasmic localization in transfected 293 cells. Red, SCL or SCL-M13; blue, Hoechst-stained nuclei.

variations further support the view that other residues can compensate for the absence of Phe-238. Importantly, mutation of three residues that are conserved in *D. melanogaster* SCL and divergent in NSCL1 (H217L, N226I, and K234C, mutant M12) leads to a significant loss of activity (Fig. 3, *A* and *B*, left graph) and a 10-fold reduction in LMO2 interaction (Fig. 3C, lanes 8–14). These substitutions change the charge and hydrophilicity at the corresponding positions, consistent with their importance at the functional interface with LMO2. Finally, complete loss of SCL activity and interaction with LMO2 requires additional mutations in divergent helix 2 residues (mutant M13, Fig. 3, *A* and *B* and lanes 15–17 in *C*). Therefore, although our results are in general agreement with Schlaeger *et al.* (17) regarding the importance of His-217 and Phe-238, they extend previous findings by revealing that the SCL-LMO2 interaction is specified by an extended interface that includes two other residues, Asn-226 and Lys-234.

As mentioned above, several proteins define the SCL subfamily of bHLH transcription factors, including three hematopoietic factors (SCL, TAL2, and LYL1), all of which are implicated in chromosomal rearrangements in acute leukemias, and two neuronal factors (NSCL1 and NSCL2). Significantly, all of the residues that we find to be important for LMO2 interaction

are conserved in SCL, TAL2, and LYL-1 but diverge in NSCL1/2 (Fig. 3D). To visualize the structural arrangement of these residues, a three-dimensional model of the SCL/E47 heterodimer bound to DNA was generated using the crystallographic coordinates of the published E47/E47 homodimer (32). As represented in Fig. 3E, His-217 and Phe-238 are in close proximity in the model despite being far apart in the primary sequence. Together with Lys-234, these residues lie on a single surface oriented away from the interaction interface of SCL with E47 or of SCL with DNA. Although Asn-226 lies farther away at the beginning of helix 2, it may also be involved in the docking of LMO2 molecules onto SCL/E47 heterodimers. In summary, the positioning of these residues is consistent with their function as an interface for interaction with LMO2 that is distinct from that involved in E47 heterodimerization. In agreement with this view, the M13 mutant, which is unable to associate with LMO2 and to nucleate higher order complexes on GPA promoter sequences (supplemental Fig. S3), efficiently co-immunoprecipitates with E47 (data not shown) and is as active as wild type SCL in inhibiting

E-protein-dependent activation of the PT $\alpha$  enhancer (Fig. 3B, right graph) (33). This type of configuration explains how the HLH domain can exert a nucleation function by simultaneously accommodating multiple protein interactions within larger protein complexes.

Taken together these results indicate that the SCL-LMO2 interaction is specified by four residues, His-217, Asn-226, Lys-234, and Phe-238 that appear to lie within an exposed surface of the SCL molecule. This interaction is required for assembling SCL complexes on erythroid target gene regulatory elements and transcriptional activation, whereas it is dispensable for the inhibition of lymphoid E-protein target genes.

*The SCL-LMO2 Interaction Is Required for Endogenous Erythroid Gene Activation and Erythroid Differentiation*—Definitive erythropoiesis takes place in the fetal liver at day E12.5 in the mouse. To address the functional importance of SCL-LMO2 interaction in erythroid differentiation (34, 35), SCL or SCL-M13 were expressed in primary Epo receptor (EpoR)-deficient fetal liver cells by retroviral transduction (Fig. 4A). In the absence of EpoR signaling, erythroid progenitors are generated, but these cells are blocked in differentiation and do not express globin genes (36). Enforced expression of SCL leads to a 3-fold enhancement in the number of erythroid (BFU-E) and

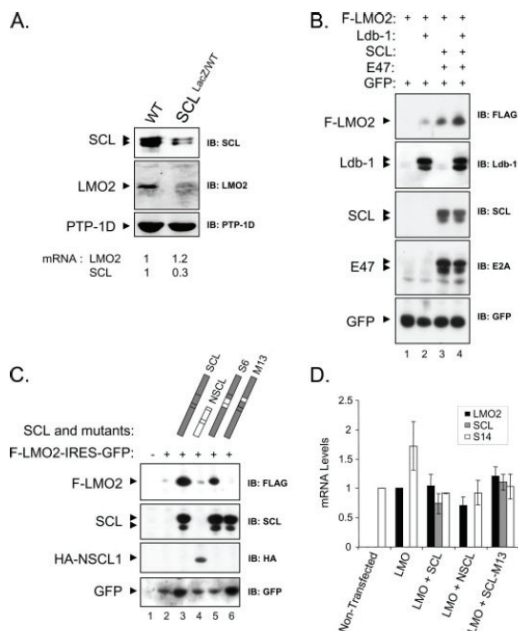
Supplemental Material can be found at:  
<http://www.jbc.org/content/suppl/2007/09/18/M703939200.DC1.html>

### Molecular Functions of the SCL-LMO2 Interaction

megakaryocyte (CFU-Meg) progenitors assayed in methylcellulose cultures (Fig. 4B) without affecting the number of granulocyte-macrophage progenitors (CFU-GM). Furthermore, SCL transduction induces *GPA* and  $\beta$ -globin expression by 2–3-fold in infected cells (Fig. 4C). In contrast, SCL-M13 transduction had no effect on *GPA* expression levels or on progenitor cell profiles despite being efficiently expressed in cells infected with the same virus stocks (Fig. 4D). Furthermore, IF analysis of 293 cells expressing SCL or SCL-M13 shows that the mutant protein exhibits normal nucleoplasmic localization (Fig. 4E), indicating that its lack of activity is not due to inappropriate subcellular distribution. Finally, SCL-M13 was also defective in inducing erythroid differentiation and *GPA* gene activation in transduced TF-1 cells (data not shown). We conclude that the stimulatory effect of SCL on erythroid differentiation at the onset of definitive erythropoiesis and its capacity to activate erythroid target genes in chromatin requires direct interaction with LMO2.

**Interaction with SCL Enhances LMO2 Protein Levels**—While conducting the above experiments we observed an intriguing correlation between SCL and LMO2 protein levels. First, we observed a modest but reproducible (~2-fold) increase in LMO2 protein expression in SCL-transduced TF-1 cells, as revealed after normalization to protein-tyrosine phosphatase-1P phosphatase levels used as a loading control, whereas LMO2 levels are unaffected in cells overexpressing SCL-M13 (Fig. 4D). Secondly, a similar correlation was seen in primary hematopoietic cells in which SCL protein levels are genetically reduced by *lacZ* insertion into one allele in the *SCL* locus (*SCL<sup>lacZ/WT</sup>*) (37). As shown in Fig. 5A, a striking decrease in LMO2 protein levels is observed in primary *SCL<sup>lacZ/WT</sup>* fetal liver cells, whereas LMO2 mRNA levels are unaffected. Significantly, this decrease correlates with reduced output in erythroid progenitors (data not shown). Third, we observed that LMO2 expression is barely detected when transfected alone in various heterologous cell lines (Fig. 5B, lane 1; data not shown), whereas it is readily detectable upon co-expression of either Ldb1 or SCL/E47 (lanes 2 and 3) and is further augmented in the presence of all three partners (lane 4). In contrast, GFP expression from an independent vector is constant in all samples. A similar correlation in SCL and LMO2 expression was recently observed (17). Together, these observations suggest that LMO2 protein expression is enhanced in the presence of its interaction partners.

We next asked whether an intact SCL-LMO2 interaction interface is required for the enhancement of LMO2 expression. As shown in Fig. 5C, LMO2 protein is strongly expressed when either SCL or the active SCL-S6 chimera is present (lanes 3 and 5). In contrast, LMO2 is barely detectable when co-expressed with NSCL1 (lane 4) or SCL-M13 (lane 6). In this case the levels of the control GFP protein, translated from the same mRNA molecules as LMO2 via an internal ribosome entry site (IRES), are comparable in all samples. These differences in protein levels are not due to variations in *LMO2* mRNA levels, as these were found to be comparable in transfected cell samples by quantitative real time RT-PCR analysis (Fig. 5D). We conclude that the effect of SCL on LMO2 protein expression is post-translational and requires the direct interaction of both factors.

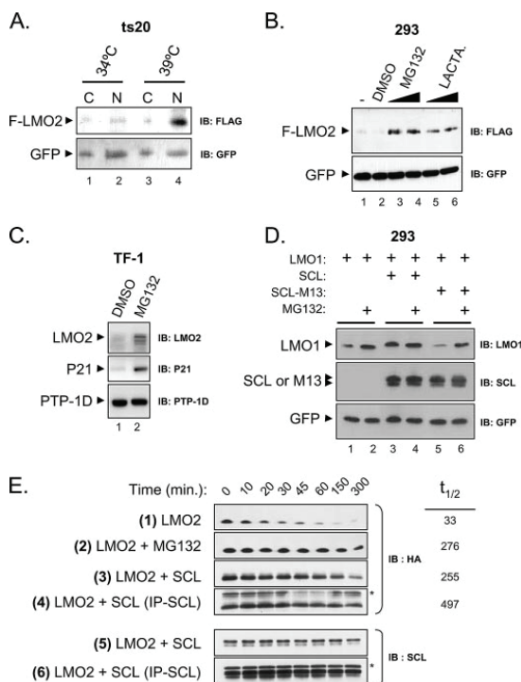


**FIGURE 5. LMO2 protein levels are modulated by SCL.** A, reduced SCL expression perturbs LMO2 levels *in vivo*. Protein extracts from unsorted E12.5 fetal liver cells were isolated from wild type (WT) and *SCL<sup>lacZ/WT</sup>* mice and subjected to IB with the indicated antibodies. The fetal liver contained 70–80% TER119+ cells irrespective of the genotype. Note the concomitant decrease in SCL and LMO2 proteins in *SCL<sup>lacZ/WT</sup>* cells. Normalized SCL and LMO2 mRNA levels, measured by real time RT-PCR analysis, are indicated below the panels. Note that LMO2 mRNA levels are unaffected in *SCL<sup>lacZ/WT</sup>* cells. PTP-1D, protein-tyrosine phosphatase-1D. B, expression of the LMO2 protein requires the presence of its interacting partners. 293 cells were co-transfected with the expression vector combinations (5  $\mu$ g each) indicated at the top of the panels together with MSCV-GFP as a control. After 36 h nuclear extracts were prepared and analyzed by IB with the indicated antibodies. C, enhanced LMO2 expression requires direct interaction with SCL. 293 cells were transfected with the MSCV-F-LMO2-IRES-GFP vector, which produces a bicistronic mRNA encoding F-LMO2 and GFP as well as vectors for the SCL, NSCL1, or different SCL mutants (1  $\mu$ g of each). Note that SCL and the active SCL-S6 chimera enhance LMO2 protein levels, whereas NSCL1 and SCL-M13 do not. D, LMO2, SCL, and S14 mRNA levels in non-transfected 293 cells or cells transfected with the indicated combinations of expression vectors, as assessed by real time RT-PCR analysis. Shown are the average  $\pm$  S.D. of two independent experiments performed in duplicate.  $\Delta$ CT and relative quantifications ( $2^{-\Delta(\Delta C T_{exp} - \Delta C T_{control})}$ ) were evaluated. Amplification was in the linear range (cycle threshold of 18 for LMO2, 14 for SCL, and 15 for S14).

**Interaction with SCL Prevents LMO2 Proteasomal Degradation**—Recent studies have shown that Ldb1 and LMO2 proteins are targeted for proteasomal degradation (27, 28, 38). To address the possibility that LMO2 may be degraded via the ubiquitin conjugation pathway, we made use of the *ts20* cell line that expresses a temperature-sensitive ubiquitin-activating E1 enzyme (30, 39). The MSCV-F-LMO2-IRES-GFP vector was transfected into *ts20* cells, and both cytoplasmic and nuclear extracts were prepared from cells grown either at the permissive (34 °C) or restrictive temperatures (39 °C). As shown in Fig. 6A, LMO2 levels accumulate substantially in the nuclei of cells grown at 39 °C, thus providing genetic evidence that LMO2 is degraded by the ubiquitin conjugation system. Consistent with these findings, LMO2 expression in transfected 293 cells is

Supplemental Material can be found at:  
<http://www.jbc.org/content/suppl/2007/09/18/M703939200.DC1.html>

### Molecular Functions of the SCL-LMO2 Interaction



**FIGURE 6. LMO2 degradation by the proteasome is prevented through interaction with SCL.** *A*, LMO2 degradation requires a functional ubiquitin conjugation pathway. The MSCV-F-LMO2-IRES-GFP vector was transfected into ts20 cells, which harbor a temperature-sensitive allele of the ubiquitin-activating enzyme E1. Duplicate samples were kept either at the permissive (34 °C) or restrictive (39 °C) temperatures for 12 h, then cytoplasmic (C) and nuclear (N) extracts were prepared and subjected to IB. *B* and *C*, LMO2 is targeted for proteasomal degradation in transfected cells and in hematopoietic cells. *B*, the MSCV-F-LMO2-IRES-GFP vector was transfected into 293 cells. After 36 h the cells were treated with MG132 (12.5 and 25  $\mu$ M), and lactacystin (10 and 20  $\mu$ M) and total cell lysates were analyzed by IB. A control treatment was performed with Me<sub>2</sub>SO (DMSO). *C*, treatment of SCL-expressing TF-1 cells with MG132 (25  $\mu$ M) for 4 h stabilizes LMO2 and P21 expression. TF-1 cell nuclear extracts were subjected to IB with the indicated antibodies. *D*, LMO1 proteasomal degradation is also suppressed by co-expression of SCL. 293 cells expressing LMO1 with or without SCL were treated with MG132 where indicated. Nuclear extracts were subjected to IB analysis with the indicated antibodies. The MSCV-GFP vector was co-transfected as a control for normalization. *E*, the half-life of LMO2 is strongly increased in the presence of SCL. Transfections were carried out in 293 cells with an expression vector encoding HA-LMO2 alone (*panels 1* and *2*) or in combination with SCL (*panels 3–6*). Cells were then treated with cycloheximide (100  $\mu$ g/ml) for the indicated times (0–300 min), and nuclear extracts were prepared. Samples were analyzed by IB (*panels 1, 2, 3, and 5*), and a portion of the extracts was subjected to IP with an anti-SCL antibody before IB. Note the striking increase in the half-life of LMO2 after MG132 treatment (25  $\mu$ M) or when SCL is present.

strikingly increased after treatment with either of two proteasome inhibitors, MG132, or lactacystin (Fig. 6*B*, lanes 3–6) but not in control untreated or Me<sub>2</sub>SO-treated cells (lanes 1 and 2). Again, GFP expression from bicistronic mRNA molecules is comparable in all samples. Furthermore, MG132 treatment of hematopoietic TF-1 cells increases endogenous LMO2 expression (Fig. 6*C*) as well as p21 levels, a protein known to be regulated by proteasomal degradation (30).

As seen with LMO2, expression of the LMO1 protein is also increased by treatment with MG132 or by co-expression

of SCL (Fig. 6*D*). However, MG132 does not further increase LMO1 protein levels when SCL is present. As expected, the SCL-M13 mutant is inactive in this co-transfection assay. Taken together, our observations are consistent with the view that interaction with SCL protects LMO proteins from proteasomal degradation.

Finally, to directly assess whether the stability of LMO2 is influenced by SCL, we measured the half-life of LMO2 in the absence or presence of SCL through the use of cycloheximide, an inhibitor of *de novo* protein synthesis. When expressed alone in 293 cells, LMO2 is rapidly degraded and exhibits a half-life of 33 min (Fig. 6*E*, *panel 1*). As expected, LMO2 is stabilized after MG132 treatment, with its half-life increasing to 276 min (*panel 2*). Interestingly, co-expression of SCL leads to a similar increase in LMO2 half-life (255 min), and when the SCL-associated fraction of LMO2 protein is isolated via SCL IP, the half-life value increases to 497 min (*panel 4*). SCL expression is also stable in these experiments (*panels 5* and *6*). Finally, because most of the LMO2 pool is associated with SCL in hematopoietic cells (Fig. 1*A*), our observations are consistent with the view that excess LMO2 is degraded by the proteasome and that this degradation is prevented by interaction with SCL.

### DISCUSSION

In the present study we define molecular determinants that control the regulatory activity and assembly of SCL complexes. We find that the SCL-LMO2 interaction provides biological specificity for the nucleation and function of these complexes during hematopoietic gene activation. Furthermore, we demonstrate that this interaction prevents LMO2 protein degradation, thus revealing an additional post-translational mechanism regulating SCL complexes during hematopoiesis.

**The SCL-LMO2 Interaction Interface and the Control of Hematopoietic Gene Expression**—The biological functions of the SCL and LMO2 genes are intimately intertwined during normal hematopoiesis (16, 19–24, 40) and leukemogenesis (7). Both factors also play important functions at different branch points of the hematopoietic cell hierarchy. Indeed, elegant gene complementation experiments in SCL<sup>-/-</sup> ES cells and morpholino-treated zebrafish embryos have shown the SCL-LMO2 interaction is required at the onset of hematopoiesis to specify mesodermal precursors to a hematopoietic fate (16–18). Therefore, SCL and LMO2 appear to act as commitment factors in hemangioblast development, thus placing them at the top of the regulatory hierarchy governing hematopoietic development (18). Conditional gene knock-out studies have also established that SCL plays an essential function in adult erythropoiesis and megakaryopoiesis (41, 42). Importantly, here we show that the SCL-LMO2 interaction is required for the stimulatory effect of SCL on erythroid and megakaryocytic progenitors. Because we find that SCL expands these lineages without affecting myeloid progenitor numbers, this suggests that SCL functions as an expansion factor within the erythroid and megakaryocytic pathways rather than a lineage commitment factor.

Our results suggest that the SCL-LMO2 interaction represents a primary nucleation event in the assembly of SCL complexes on DNA, a property that relies on the modularity of these

Supplemental Material can be found at:  
<http://www.jbc.org/content/suppl/2007/09/18/M703939200.DC1.html>

### Molecular Functions of the SCL-LMO2 Interaction

factors and their ability to simultaneously accommodate multiple protein interactions. Although Schlaeger *et al* (17) previously highlighted the importance of the His-217 and Phe-238 residues for the interaction with LMO2, our analysis further suggests that the SCL-LMO2 interaction interface involves additional residues located throughout the loop and helix 2 of SCL. Importantly, these residues lie in close proximity to each other in a three-dimensional model of the SCL/E2A heterodimer. We propose that heterodimerization with E2A, occurring via broadly conserved residues in the HLH domain of SCL, exposes the loop and part of Helix2 for interaction with LMO2. Because LMO proteins are likely constitutively associated with Ldb1, which has the capacity to self-dimerize and bridge the recruitment of additional LMO molecules, this enables subsequent assembly of multimeric complexes, as elegantly demonstrated in previous studies (10, 11, 43). The LIM1 and LIM2 domains of LMO2 have been shown to mediate specific protein interactions with Ldb1 and GATA factors, respectively (44, 45). This modularity, thus, enables the tethering of SCL complexes to DNA via different recruitment platforms, such as SCL-E2A heterodimers, GATA factors, or additional DNA-binding proteins, such as Sp1 (10, 11, 13, 15). Additional interactions of SCL complexes with co-factors, such as the ETO2 co-repressor, can further modulate the recruitment of chromatin remodeling and histone modification factors to target gene regulatory elements and, thus, control the transition from transcriptionally active to repressive states or vice versa (12, 46–49). These transitions in chromatin states likely provide epigenetic switches that regulate cell fate commitment and differentiation (12).

The all-or-nothing activity of SCL-containing complexes (13, 15) is reminiscent of enhanceosome complexes that control the expression of the interferon- $\beta$  and T-cell receptor  $\alpha$  genes, in which cooperative assembly of higher order complexes containing tissue-specific and signal-induced transcription factors on enhancer sequences is observed (2, 50–52). The function of these complexes also relies on architectural proteins, such as high mobility group I(Y) and LEF-1, and on interactions with chromatin remodeling complexes. The current study provides another example of how critical protein interactions can nucleate multifactorial complexes, thus providing a binary switch in transcription activation.

**Coupling Protein Stabilization and Protein Complex Assembly**—Proteins with short half-lives often represent rate-limiting components of biological processes. Over the past few years, ubiquitination and proteasomal degradation of transcription factors has emerged as a major mechanism in the control of gene transcription (53, 54). In some instances degradation serves to limit the activity of specific transcription factors, whereas high transcriptional potency is often directly correlated with high rates of transcription factor turnover (53–56). The underlying instability of LMO proteins likely prevents the assembly of higher order complexes in the absence of appropriate interacting partners, thus providing an efficient mechanism to control hematopoietic transcription factor codes.

This study demonstrates that LMO2 is subject to proteasomal degradation and that interaction with SCL stabilizes LMO2. It has recently been shown that regulated degradation

of Ldb1 controls the stoichiometry of Ldb1-containing complexes during development (27, 28, 38). Several studies in *Drosophila* have also revealed that LIM-homeodomain and LMO proteins can be stabilized by association with Chip, the *Drosophila* orthologue of Ldb1 (43, 57, 58). Our finding that efficient LMO2 expression requires the presence of either Ldb1 or SCL, which belong to distinct protein families, suggests that protein stabilization by interacting co-factors represents a major mode of regulation of LIM domain proteins. This view is consistent with a recent study by Xu *et al* (28) who showed that Ldb1 and LMO2 can be stabilized through interaction with yet another category of partners, the single-stranded DNA-binding proteins. Similar mechanisms coupling protein stabilization to protein complex assembly likely operate to control a variety of functional modules within the cell.

Finally, different combinations of bHLH-LIM factor interactions may be utilized as a general strategy to control differentiation in many tissues. For instance, NSCL1 might interact with distinct LMO family members during neuronal development (59), whereas LMO proteins of the cysteine-rich protein family have been shown to enhance the myogenic activity of MyoD and to function as adaptors in multifactorial complexes that potently activate smooth muscle cell gene expression (60, 61). Therefore, we speculate that interactions between bHLH and LIM proteins have evolved to create codes that specify cell fates by enabling the formation of tissue-specific transcription factor complexes.

**Acknowledgments**—We thank members of the Meloche group for helpful suggestions and for the *ts20* cell line.

### REFERENCES

1. Tjian, R., and Maniatis, T. (1994) *Cell* 77, 5–8
2. Grosschedl, R. (1995) *Curr. Opin. Cell Biol.* 7, 362–370
3. Shivdasani, R. A., and Orkin, S. H. (1996) *Blood* 87, 4025–4039
4. Sieweke, M. H., and Graf, T. (1998) *Curr. Opin. Genet. Dev.* 8, 545–551
5. Swiers, G., Patient, R., and Loose, M. (2006) *Dev. Biol.* 294, 525–540
6. Lecuyer, E., and Hoang, T. (2004) *Exp. Hematol.* 32, 11–24
7. Begley, C. G., and Green, A. R. (1999) *Blood* 93, 2760–2770
8. Massari, M. E., and Murre, C. (2000) *Mol. Cell. Biol.* 20, 429–440
9. Visvader, J. E., Mao, X., Fujiwara, Y., Hahm, K., and Orkin, S. H. (1997) *Proc. Natl. Acad. Sci. U. S. A.* 94, 13707–13712
10. Wadman, I. A., Osada, H., Grutz, G. G., Agulnick, A. D., Westphal, H., Forster, A., and Rabbitts, T. H. (1997) *EMBO J.* 16, 3145–3157
11. Grutz, G. G., Bucher, K., Lavenir, I., Larson, T., Larson, R., and Rabbitts, T. H. (1998) *EMBO J.* 17, 4594–4605
12. Goardon, N., Lambert, J. A., Rodriguez, P., Nissaire, P., Herblot, S., Thibault, P., Dumenil, D., Strouboulis, J., Romeo, P. H., and Hoang, T. (2006) *EMBO J.* 25, 357–366
13. Lecuyer, E., Herblot, S., Saint-Denis, M., Martin, R., Begley, C. G., Porcher, C., Orkin, S. H., and Hoang, T. (2002) *Blood* 100, 2430–2440
14. Xu, Z., Huang, S., Chang, L. S., Agulnick, A. D., and Brandt, S. J. (2003) *Mol. Cell. Biol.* 23, 7585–7599
15. Lahlil, R., Lecuyer, E., Herblot, S., and Hoang, T. (2004) *Mol. Cell. Biol.* 24, 1439–1452
16. Mead, P. E., Deconinck, A. E., Huber, T. L., Orkin, S. H., and Zon, L. I. (2001) *Development* 128, 2301–2308
17. Schlaeger, T. M., Schuh, A., Flitter, S., Fisher, A., Mikkola, H., Orkin, S. H., Vyas, P., and Porcher, C. (2004) *Mol. Cell. Biol.* 24, 7491–7502
18. Patterson, L. J., Gering, M., Eckfeldt, C. E., Green, A. R., Verfaillie, C. M., Ekker, S. C., and Patient, R. (2007) *Blood* 109, 2389–2398
19. Warren, A. J., Colledge, W. H., Carlton, M. B., Evans, M. J., Smith, A. J., and



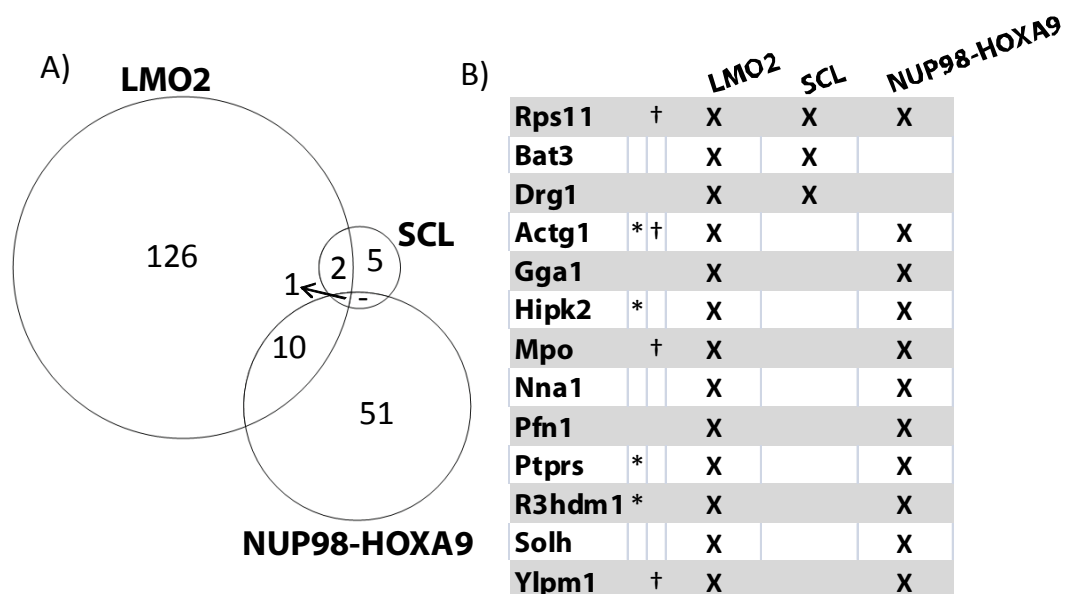
Supplemental Material can be found at:  
<http://www.jbc.org/content/suppl/2007/09/18/M703939200.DC1.html>

### Molecular Functions of the SCL-LMO2 Interaction

- Rabbitts, T. H. (1994) *Cell* **78**, 45–57
20. Robb, L., Lyons, I., Li, R., Hartley, L., Kontgen, F., Harvey, R. P., Metcalf, D., and Begley, C. G. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 7075–7079
  21. Shivdasani, R. A., Mayer, E. L., and Orkin, S. H. (1995) *Nature* **373**, 432–434
  22. Porcher, C., Swat, W., Rockwell, K., Fujiwara, Y., Alt, F. W., and Orkin, S. H. (1996) *Cell* **86**, 47–57
  23. Robb, L., Elwood, N. J., Elefanty, A. G., Kontgen, F., Li, R., Barnett, L. D., and Begley, C. G. (1996) *EMBO J.* **15**, 4123–4129
  24. Yamada, Y., Warren, A. J., Dobson, C., Forster, A., Pannell, R., and Rabbitts, T. H. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 3890–3895
  25. Rabbitts, T. H. (1998) *Genes Dev.* **12**, 2651–2657
  26. Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulfraat, N., Leboulch, P., Lim, A., Osborne, C. S., Pawliuk, R., Morillon, E., Sorensen, R., Forster, A., Fraser, P., Cohen, J. L., de Saint Basile, G., Alexander, I., Wintergerst, U., Frebourg, T., Aurias, A., Stoppa-Lyonnet, D., Romana, S., Radford-Weiss, L., Gross, F., Valensi, F., Delabesse, E., Macintyre, E., Sigaux, F., Soulier, J., Leiva, L. E., Wissler, M., Prinz, C., Rabbitts, T. H., Le Deist, F., Fischer, A., and Cavazzana-Calvo, M. (2003) *Science* **302**, 415–419
  27. Ostendorff, H. P., Peirano, R. I., Peters, M. A., Schluter, A., Bossenz, M., Scheffner, M., and Bach, I. (2002) *Nature* **416**, 99–103
  28. Xu, Z., Meng, X., Cai, Y., Liang, H., Nagarajan, L., and Brandt, S. J. (2007) *Genes Dev.* **21**, 942–955
  29. Krosch, G., He, G., Lefrancois, M., Charron, F., Romeo, P. H., Jolicoeur, P., Kirsch, I. R., Nemer, M., and Hoang, T. (1998) *J. Exp. Med.* **188**, 439–450
  30. Coulombe, P., Rodier, G., Bonneil, E., Thibault, P., and Meloche, S. (2004) *Mol. Cell Biol.* **24**, 6140–6150
  31. Calkhoven, C. F., Muller, C., Martin, R., Krosch, G., Pietsch, H., Hoang, T., and Leutz, A. (2003) *Genes Dev.* **17**, 959–964
  32. Ellenberger, T., Fass, D., Arnaud, M., and Harrison, S. C. (1994) *Genes Dev.* **8**, 970–980
  33. Tremblay, M., Herblot, S., Lecuyer, E., and Hoang, T. (2003) *J. Biol. Chem.* **278**, 12680–12687
  34. Aplan, P. D., Nakahara, K., Orkin, S. H., and Kirsch, I. R. (1992) *EMBO J.* **11**, 4073–4081
  35. Hoang, T., Paradis, E., Brady, G., Billia, F., Nakahara, K., Iscove, N. N., and Kirsch, I. R. (1996) *Blood* **87**, 102–111
  36. Wu, H., Liu, X., Jaenisch, R., and Lodish, H. F. (1995) *Cell* **83**, 59–67
  37. Elefanty, A. G., Begley, C. G., Metcalf, D., Barnett, L., Kontgen, F., and Robb, L. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 11897–11902
  38. Hiratani, I., Yamamoto, N., Mochizuki, T., Ohmori, S. Y., and Taira, M. (2003) *Development* **130**, 4161–4175
  39. Chowdary, D. R., Dermody, J. J., Jha, K. K., and Ozer, H. L. (1994) *Mol. Cell Biol.* **14**, 1997–2003
  40. Silver, L., and Palis, J. (1997) *Blood* **89**, 1154–1164
  41. Hall, M. A., Curtis, D. J., Metcalf, D., Elefanty, A. G., Sourris, K., Robb, L., Gothert, J. R., Jane, S. M., and Begley, C. G. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 992–997
  42. Mikkola, H. K., Klintman, J., Yang, H., Hock, H., Schlaeger, T. M., Fujiwara, Y., and Orkin, S. H. (2003) *Nature* **421**, 547–551
  43. Milan, M., and Cohen, S. M. (1999) *Mol. Cell* **4**, 267–273
  44. Deane, J. E., Mackay, J. P., Kwan, A. H., Sum, E. Y., Visvader, J. E., and Matthews, J. M. (2003) *EMBO J.* **22**, 2224–2233
  45. Terano, T., Zhong, Y., Toyokuni, S., Hiai, H., and Yamada, Y. (2005) *Exp. Hematol.* **33**, 641–651
  46. Meier, N., Krpic, S., Rodriguez, P., Strouboulis, J., Monti, M., Krijgsveld, J., Gering, M., Patient, R., Hostert, A., and Grosveld, F. (2006) *Development* **133**, 4913–4923
  47. Schuh, A. H., Tipping, A. J., Clark, A. J., Hamlett, I., Guyot, B., Iborra, F. J., Rodriguez, P., Strouboulis, J., Enver, T., Vyas, P., and Porcher, C. (2005) *Mol. Cell Biol.* **25**, 10235–10250
  48. Xu, Z., Meng, X., Cai, Y., Koury, M. J., and Brandt, S. J. (2006) *Biochem. J.* **399**, 297–304
  49. Huang, S., Qiu, Y., Shi, Y., Xu, Z., and Brandt, S. J. (2000) *EMBO J.* **19**, 6792–6803
  50. Giese, K., Kingsley, C., Kirshner, J. R., and Grosschedl, R. (1995) *Genes Dev.* **9**, 995–1008
  51. Kim, T. K., and Maniatis, T. (1997) *Mol. Cell* **1**, 119–129
  52. Merika, M., and Thanos, D. (2001) *Curr. Opin. Genet. Dev.* **11**, 205–208
  53. Thomas, D., and Tyers, M. (2000) *Curr. Biol.* **10**, 341–343
  54. Lipford, J. R., and Deshaies, R. J. (2003) *Nat. Cell Biol.* **5**, 845–850
  55. Molinari, E., Gilman, M., and Natesan, S. (1999) *EMBO J.* **18**, 6439–6447
  56. Salghetti, S. E., Muratani, M., Wijnen, H., Futcher, B., and Tansey, W. P. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97**, 3118–3123
  57. Milan, M., and Cohen, S. M. (2000) *Development* **127**, 3069–3078
  58. Weihe, U., Milan, M., and Cohen, S. M. (2001) *Development* **128**, 4615–4622
  59. Bao, J., Talmage, D. A., Role, L. W., and Gautier, J. (2000) *Development* **127**, 425–435
  60. Kong, Y., Flick, M. J., Kudla, A. J., and Konieczny, S. F. (1997) *Mol. Cell Biol.* **17**, 4750–4760
  61. Chang, D. F., Belaguli, N. S., Iyer, D., Roberts, W. B., Wu, S. P., Dong, X. R., Marx, J. G., Moore, M. S., Beckerle, M. C., Majesky, M. W., and Schwartz, R. J. (2003) *Dev. Cell* **4**, 107–118

## Annexe 3

### LMO2 et SCL partagent des partenaires protéiques avec la protéine de fusion oncogénique NUP98-HOXA9



**Figure A.2 Partenaires protéiques communs entre LMO2, SCL et NUP98-HOXA9, identifiés par double hybride.**

A) Diagramme de Venn représentant le nombre de partenaires protéiques différents pour LMO2, SCL et NUP98-HOXA9, révélés par la méthode de double hybride. B) Liste des partenaires protéiques en commun entre LMO2, SCL et NUP98-HOXA9. \*Interaction confirmée *in vitro* (Figure 4.2). †Protéine identifiée sur les gènes cibles de LMO2 (SCL et/ou GYPA) par spectrométrie de masse (Figure 4.4).

