



UNIVERSITÉ DE STRASBOURG

ED414 – école doctorale des sciences de la vie et
de la santé



École Doctorale
des Sciences de la Vie
et de la Santé
STRASBOURG

THÈSE présentée par :

Ava GALLAND

INSERM U1113 – IRFAC-SMART – Strasbourg

Fonction et régulation du gène homéotique Cdx2 dans les leucémies aiguës.

Soutenue le : **17 décembre 2021**

Pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline / Spécialité : **Aspects Moléculaires et Cellulaires de la Biologie**

THÈSE DIRIGÉE PAR :

Mme Claire DOMON-DELL Directeur de thèse CR, IRFAC Strasbourg, Université de Strasbourg

MEMBRES DU JURY :

Mme	Emmanuelle CLAPPIER	PUPH, Hôpital Saint Louis, Paris. (Rapporteuse)
Mme	Monique DONTENWILL	DR, Faculté de Pharmacie, Illkirch-Graffenstaden
M	Philippe KASTNER	MCUPH, IGBMC, Illkirch-Graffenstaden
M	Bernard NAMOUR	PUPH, Faculté de médecine, Vandœuvre-lès-Nancy
M	Éric SOLARY	PUPH, Institut Gustave Roussy, Villejuif. (Rapporteur)

A mes parents,

REMERCIEMENTS

Je tiens tout d'abord à remercier Emmanuelle Clappier, Monique Dontenwill, Philippe Kastner, Bernard Namour et Eric Solary, les membres jury, d'avoir accepté d'évaluer mon travail de thèse. J'espère que la lecture de mon manuscrit vous sera agréable.

Je tiens tout particulièrement à remercier Jean-Noël Freund, directeur de l'unité de recherche, à qui j'avais, en 2016, tout innocemment adressé ma candidature de stage de M1. C'est en effet chronologiquement avec toi que tout a commencé. J'avais particulièrement accroché au cours que tu as dispensé ce 5 Octobre 2016, sur un certain gène homéotique... Car oui, quoi que tu en dises, tu es un excellent pédagogue, la preuve : il s'agissait là du point de départ de mon aventure. Et quelle aventure ! Je suis encore là, 4 ans plus tard, irrémédiablement contaminée par ce virulent virus qu'est Cdx2. Je suis heureuse d'avoir pu côtoyer quelqu'un qui incarne à mes yeux la passion du métier. Apercevoir dans l'entrebâillement de la porte, encore tard le soir, ta silhouette penchée sur le microscope, a été une réelle source d'inspiration et de persévérance.

Revenons donc en cette fin d'année 2016, où ma candidature avait fini par ricocher dans une boîte mail tout à fait inattendue. Alors que je postulais pour de l'intestin, pour de l'épithélium, que dis-je, pour du carcinome ! Je me suis retrouvée face à une problématique toute autre, mais pas pour autant dénuée de charme. Cdx2 oui, mais dans les leucémies. Le sujet m'a rapidement plu, au même titre que sa détentric, qui a de suite su me mettre à l'aise et en confiance par son accueil chaleureux. En adoptant ce sujet de stage de M1, puis de M2, et allez hop' pourquoi pas de thèse aussi tant qu'on y est, j'ai eu l'incroyable chance d'être acceptée telle que je suis, avec mes forces, mais aussi mes faiblesses et mes bizarreries. La confiance en moi et l'assurance me faisant cruellement défaut, je ne te remercierai jamais assez, Claire, d'avoir su me prendre par la main et d'avoir fait preuve de patience et de pédagogie. Humainement, je n'aurai osé espérer autant de la part d'un encadrant. Lors de mes périodes les plus sombres, c'est ta bienveillance qui a continué de me donner envie de faire bien, de faire mieux, de faire plus. Je ressors de cette expérience grandie sur bien des aspects, et c'est pour beaucoup grâce à toi.

Ces presque 5 années que j'ai passées dans l'unité de recherche ont été riches en rencontres, je pense avoir eu la chance d'apprécier les qualités d'un grand nombre d'entre vous. Merci à toi Isabelle D, voisine de bureau, complice de papotages en tout genre quand la journée commence à se faire longue, mais aussi et surtout, source de conseils avisés. Partager le même bureau m'a permis de bénéficier de ton recul

sur les problématiques qui s'opposaient à moi : cette source d'encadrement supplémentaire était réellement agréable. Merci à toi Isabelle G, de t'être systématiquement rendue disponible quand je débarquais littéralement en trombe dans ton bureau, encore et toujours avec des questions techniques. Mais également pour ton efficacité et ta fiabilité en ce qui concerne l'approvisionnement en kits toujours plus variés et farfelus dont je pouvais avoir besoin, tu as toujours veillé à ce que les manip soient faites dans de bonnes conditions, avec le matériel nécessaire, je t'en suis reconnaissante. Merci à toi Elisabeth, pilier inébranlable du laboratoire, pour le soutien constant que tu m'as apporté au cours de mes manip. Merci de m'avoir fait faire mes premiers pas en culture cellulaire, et de toi aussi t'être assurée que l'environnement de travail soit optimal. J'espère avoir appris de ta rigueur. Merci également à tous ceux m'ayant apporté leur expertise dans leur domaine de prédilection. Je pense notamment à Manuela et Laurent : merci à vous deux pour votre implication dans mon projet de thèse, j'ai vraiment apprécié de pouvoir bénéficier de votre regard critique sur mon travail. Je tiens de la même manière à remercier Marie-Pierre Chenard, Laurent Mauvieux et Laurent Miguet de l'hôpital de Hautepierre, Thomas Sexton, Nezih Karasu, Thomas Kleiber, Claudine Ebel et Bernard Jost de l'IGBMC, qui tout en étant extérieurs au laboratoire m'ont donné de leur temps et partagé leur savoir-faire, apportant ainsi leur pierre à l'édifice.

Ahhh, j'en viens maintenant aux doctorantes. La tête toujours dans les manip à courir entre 2 étages, ou alors déjà volatilisée chez moi, je n'ai pas été la plus présente. Et pourtant, je suis si heureuse d'avoir pu côtoyer vos personnalités uniques et riches en couleurs. Grâce à vous, j'ai réalisé que... bin, que l'Homme est un animal sociable en fait ! Vous m'avez, sans doute sans vous en rendre compte, fait réaliser que l'entraide peut faire la différence. Qu'il est possible de demander de l'aide, de compter sur quelqu'un, et cela sans culpabiliser. J'ai beaucoup appris à votre contact, merci d'avoir été mes compagnons de route pour ce trajet de vie. J'en profite ici pour adresser un petit mot à Alicia et François, deux pépites que j'ai eu la chance « d'encadrer », pendant ma dernière année : je ne me fais absolument aucun souci pour la suite de votre parcours, vous êtes brillants ! Il ne tient qu'à vous d'atteindre vos objectifs ! Je crois en vous ! Merci d'avoir ajouté un peu de légèreté dans mon quotidien, pendant une période pas si facile pour moi...

Susa -quoi ? C'est qui ? .. Comment ? Non vraiment, je ne vois pas, jamais entendu parler... Non, Susanna ce n'est pas une simple rencontre, c'est une révélation ! Qui aurait cru que je ressortirai (vivante) de cette aventure, forte en prime d'une indéfectible amitié ? Quel luxe d'avoir auprès de soi un mentor, un partenaire de discussions existentielles, mais surtout quelqu'un d'aussi prévenant et fiable. Qui, je vous le demande, débarquerait chez vous en pleine nuit avec des antibiotiques à 4h du matin ?? Je te suis éternellement reconnaissante pour le soutien dont tu as fait preuve ces 3 dernières années. Jusqu'au bout tu as été là, à me porter à bout de bras, pour que j'atteigne la ligne d'arrivée. Allé, à nos futures années de voyage et de streaming !

Les amis, , je n'en compte qu'une poignée d'irréductibles, mais ô combien précieux et chérissables. Je ne suis pas la plus démonstrative, en bon ours introverti difficile à sortir de sa grotte, mais qu'est-ce que je pense à vous. Merci Raphaëlle et Agathe, avec vous, je peux être un ours tout en sillonnant la France en quête de concerts, festivals (bon d'ailleurs ça reprends quand tout ça ?), de randonnées, et autres destinations réjouissantes. Ces escapades ont activement contribué à me faire garder le cap, et à me vider la tête. Les années qui viennent nous réservent le meilleur, j'ai hâte de continuer les périples à vos côtés, à toute blinde, musique à fond, à coup de « TA TA TA » et « BANG BANG BANG » (comprendra qui pourra). Merci à toi aussi mon cher Michel, tu es mon plus ancien ami ici à Strasbourg, félicitations héhé. Les moments passés en ta compagnie sont toujours synonymes de bonne humeur, de conversations intéressantes ... et de bons repas !! Ces rendez-vous entre grands gastronomes ont activement contribué à me remonter le moral, j'espère que quoi qu'il arrive, on gardera ces bonnes habitudes ! En parlant de vieux amis, quelle n'a pas été ma surprise que nous nous retrouvions finalement à Strasbourg, Anne ! Ça en fait des années qu'on se connaît, et je sais pertinemment que même sans se voir ou se parler, ce lien reste, et nos retrouvailles sont les mêmes : simples, authentiques et bon enfant. Et cela va pour vous aussi, Jenna et Patricia : je vous souhaite le meilleur à toutes ! Vous êtes de belles personnes, j'espère vous revoir davantage les prochaines années. Je clôture ce paragraphe avec une mention spéciale -et un ressort comique- pour K. N., K. S., M. Y., J. H., P. J., K. T. et J. J., qui se reconnaîtraient s'ils passaient par-là. Merci à vous d'avoir apporté un brin de fantaisie et de soleil dans les journées passées en votre compagnie.

A ton tour d'avoir ton petit paragraphe mon Stan, merci de m'avoir épaulée sur la fin de ce chapitre de ma vie. Quoi que tu en penses, tu m'inspires et me donne envie d'aller de l'avant. Merci de croire en moi, de me rassurer, et de m'encourager à suivre mes aspirations. J'espère pouvoir te rendre tout ça au centuple, merci d'être à mes côtés. Le meilleur est à venir. Wonpiw.

Je remercie enfin du fond du cœur ma petite famille, soucieuse de ma réussite mais surtout de mon bonheur. Mamie, marraine, parrain, Alexandre, Victor et Karel : nos entrevues elles aussi sont rares, mais toujours chaleureuses, et emplies de bienveillance envers les uns et les autres. Papa, maman, je suis allée jusqu'au bout du chapitre pour vous, parce que vous avez toujours fait en sorte de mettre toutes les chances de mon côté. L'expression « loin des yeux, près du cœur » n'a jamais autant pris son sens. Je m'estime incroyablement chanceuse de vous avoir. Vous m'avez toujours encouragée, soutenue, et surtout comprise. C'est justement parce que vous ne me demandez rien que j'ai envie de vous rendre fiers. Merci.

SOMMAIRE

TABLE DES ILLUSTRATIONS	- 13 -
ABBREVIATIONS	- 15 -
INTRODUCTION	- 20 -
Chapitre 1.....	- 21 -
L'hématopoïèse	- 21 -
I. Le système hématopoïétique	- 23 -
A. Origine embryonnaire	- 23 -
B. L'Hématopoïèse chez l'adulte	- 25 -
1. Les cellules souches hématopoïétiques	- 25 -
2. Modèles de différenciation hématopoïétique	- 27 -
3. Régulateurs de l'hématopoïèse	- 28 -
II. Microenvironnement et homéostasie.....	- 28 -
A. Régulation des CSH	- 30 -
1. Le concept de niche hématopoïétique.....	- 30 -
2. Organisation spatiale	- 30 -
3. Contribution des populations hématopoïétiques	- 32 -
B. Perturbation de la niche.....	- 34 -
1. Implication dans les hémopathies	- 34 -
2. Hématopoïèse extramédullaire	- 35 -
Chapitre 2.....	- 36 -
La leucémie aiguë myéloïde	- 36 -
I. Historique et Classification des LAM	- 38 -
A. Découverte des leucémies	- 38 -
B. Classification des leucémies	- 38 -
II. Mécanismes physiopathologiques des LAM.....	- 40 -
A. Paysage mutationnel des LAM	- 41 -

B. Modèles de pathogenèse	- 43 -
1. Notion de cellules souches leucémiques	- 43 -
2. Notion d'hétérogénéité et évolution clonale	- 43 -
III. La leucémie aiguë myéloïde, à l'échelle du patient	- 45 -
A. Epidémiologie des LAM.....	- 45 -
1. Statistiques	- 45 -
2. Facteurs de risques	- 45 -
a. Prédispositions héréditaires.....	- 45 -
b. Facteurs iatrogènes	- 45 -
c. Agents leucémogènes.....	- 46 -
d. Facteurs individuels.....	- 46 -
B. Diagnostic et classification des LAM.....	- 48 -
1. Symptomatologie.....	- 48 -
2. Examens de dépistage et diagnostic	- 48 -
3. Etablissement du pronostic.....	- 50 -
C. Traitements de la LAM.....	- 50 -
1. Thérapie conventionnelle.....	- 50 -
a. Chimiothérapie d'induction	- 50 -
b. Thérapie de consolidation	- 51 -
2. Ciblage de protéines mutées.....	- 51 -
3. Ciblage de la voie apoptotique	- 52 -
4. Thérapies immunitaires	- 52 -
a. Anticorps monoclonaux.....	- 52 -
b. Réactivation du système immunitaire.....	- 52 -
5. Radiothérapie externe	- 54 -

Chapitre 3..... - 56 -

Notions d'épigénétique

I. Régulation épigénétique de la transcription.....	- 58 -
A. Mécanismes épigénétiques.....	- 58 -
1. Méthylation de l'ADN.....	- 58 -
2. Modifications post-traductionnelles des histones	- 62 -

a. Acétylation des histones.....	- 62 -
b. Méthylation des histones.....	- 62 -
3. Interactions entre ces mécanismes.....	- 63 -
a. Co-occurrences et exclusions.....	- 63 -
b. Notion de bivalence	- 64 -
B. Dynamisme de la chromatine.....	- 64 -
1. Remodelage des nucléosomes	- 64 -
2. Notion d'enhancer	- 66 -
3. Organisation tridimensionnelle et topologique	- 66 -
a. Que sont les domaines d'associations topologiques (TADs) ?.....	- 66 -
b. Structure des TADs.....	- 68 -
c. Fonctions des TADs.....	- 68 -
II. Épигénétique et leucémie aiguë myéloïde.....	- 69 -
A. Hématopoïèse physiologique	- 69 -
B. Mécanismes perturbés.....	- 70 -
1. Méthylation de l'ADN.....	- 70 -
2. Modifications d'histones.....	- 71 -
3. Organisation de la chromatine	- 73 -
C. Thérapies épigénétiques pour LAM.....	- 74 -
1. Inhibiteurs de méthyltransférase de l'ADN.....	- 74 -
2. Inhibiteurs d'acétylases.....	- 74 -

Chapitre 4..... - 76 -

Le gène homéotique *CDX2*..... - 76 -

I. Cdx2 : gène homéotique à homéoboîte.....	- 78 -
A. Origine, structure et fonction	- 78 -
B. Caudal Type 2 Homeobox gene 2.....	- 78 -
II. Fonctions physiologiques de Cdx2... ..	- 82 -
A. ... au cours de l'embryogenèse	- 82 -
1. Établissement du trophoctoderme.....	- 82 -
2. Élongation postérieure de l'embryon	- 82 -

3. Hématopoïèse primitive	- 84 -
4. Spécification du tissu intestinal	- 86 -
B. ... tout au long de la vie adulte	- 86 -
1. Fonctions transcriptionnelles dans l'intestin.....	- 87 -
2. Fonctions non-transcriptionnelles dans l'intestin.....	- 89 -
III. Fonctions de Cdx2 dans les cancers	- 90 -
A. Perte d'expression dans les CCR.....	- 90 -
B. Expression ectopique dans l'épithélium gastrique.....	- 91 -
C. Expression ectopique dans le compartiment hématopoïétique.....	- 93 -
D. Modélisation de l'expression dérégulée de Cdx2.....	- 93 -
1. Perte d'expression dans l'épithélium intestinal.....	- 93 -
2. Gain d'expression dans l'épithélium gastrique.....	- 94 -
3. Gain d'expression dans le compartiment hématopoïétique	- 95 -
IV. Mécanismes de régulation de l'expression de Cdx2	- 96 -
A. Régulation post-transcriptionnelle et post-traductionnelle.....	- 96 -
B. Régulation transcriptionnelle chez l'embryon précoce.....	- 96 -
C. Régulation transcriptionnelle dans l'intestin et les CCR.....	- 97 -
D. Régulation transcriptionnelle dans les carcinomes gastriques	- 98 -
E. Régulation transcriptionnelle dans les LAM	- 99 -
OBJECTIFS	- 100 -
RÉSULTATS.....	- 102 -
Partie 1	- 103 -
A. Un premier modèle d'induction <i>ex vivo</i>	- 104 -
B. Modèle d'induction <i>in vivo</i>	- 104 -
CDX2 expression in the hematopoietic lineage promotes leukemogenesis via TGFB inhibition.....	- 107 -
1. Introduction	107
2. Materials and methods	108

2.2. Plasmids, siRNA, and luciferase assays	108
2.3. Cell line establishment, cell culture,transfection, and TGF- β treatment.....	108
2.4. Serial replating assays.....	108
2.5. Histology, cytology, and flow cytometry	108
2.6. RNA preparation, RT-qPCR, RNAsequencing, and data analyses	108
2.7. Western blots	109
2.8. Chromatin immunoprecipitation.....	109
2.9. Statistics	109
3. Results and Discussion.....	109
3.1. Oncogenic activity of CDX2 in thehematopoietic lineage <i>in vivo</i>	109
3.2. Myeloid vs lymphoid imbalance driven byectopic CDX2.....	111
4. Conclusion	115
Partie 2	- 119 -
A. Validation de l'expression de <i>CDX2</i>	- 121 -
B. Régions de chromatine active	- 122 -
1. Identification des régions de chromatine actives par ChIP-seq	- 122 -
2. Identification de la chromatine active par ATAC-seq	- 125 -
C. Bivalence du promoteur proximal du gène <i>CDX2</i>	- 128 -
D. Régions en interaction avec le promoteur proximal de <i>CDX2</i>	- 129 -
DISCUSSION	- 133 -
Partie 1	- 134 -
I. Description de la pathologie se développant dans les souris MxCdx2.....	- 135 -
II. Cdx2 transforme les cellules hématopoïétiques	- 137 -
III. Cdx2 interfère avec l'hématopoïèse	- 138 -
A. Cdx2 favorise la différenciation des progéniteurs myéloïdes	- 138 -
1. Pas de modification de l'autorenouvellement et de la différenciation	- 138 -
2. Expansion préférentielle de cellules myéloïdes immatures	- 139 -
B. Cdx2 inhibe l'engagement vers le lignage lymphoïde	- 140 -
C. Qu'en est-il des gènes <i>HOX</i> ?.....	- 141 -
	- 10 -

IV. Perturbation de la voie de signalisation du TGF-β	- 142 -
A. La surexpression du gène Bambi réduit l'action du TGF- β	- 143 -
1. Bambi est une cible directe de Cdx2	- 143 -
2. BAMBI contribue au blocage de la différenciation monocyttaire	- 144 -
B. Impact de la voie sur les populations hématopoïétiques	- 145 -
1. Autorenouvellement, quiescence, prolifération et engagement des CSH	- 145 -
2. Différenciation inégale du lignage myéloïde	- 145 -
V. Cdx2 oncogénique dans le lignage lymphoïde ?	- 146 -
Partie 2	- 148 -
I. Validation de l'expression de Cdx2	- 149 -
A. Lignées intestinales	- 149 -
B. Lignées leucémiques	- 150 -
C. Lignées gastriques	- 150 -
II. Régions de chromatine accessible	- 151 -
A. ChIP-seq et modifications post-traductionnelles d'histones	- 151 -
B. ATAC-seq et espacement des nucléosomes	- 152 -
III. Bivalence du promoteur proximal de Cdx2	- 152 -
IV. Régions en interaction avec le promoteur proximal de Cdx2	- 153 -
CONCLUSION	- 154 -
MATÉRIELS & MÉTHODES	- 157 -
I. Matériels	- 158 -
A. Lignées cellulaires	- 158 -
B. Anticorps	- 158 -
C. Sondes nucléotidiques	- 158 -
II. Méthodes	- 160 -
A. Extraction des ARN totaux	- 160 -
B. Transcription reverse (RT)	- 160 -
C. PCR quantitative en temps réel (qPCR)	- 160 -

D. Western Blot	- 161 -
E. Immunoprécipitation de chromatine (ChIP)	- 161 -
F. Immunoprécipitation de chromatine + séquençage (ChIP-seq).....	- 162 -
G. Assay for Transposase-Accessible Chromatin (ATAC)	- 163 -
H. Rapid Amplification of cDNA Ends (RACE).....	- 164 -
I. Circularised Chromosome Conformation Capture (4C)	- 165 -
BIBLIOGRAPHIE	- 167 -

TABLE DES ILLUSTRATIONS

INTRODUCTION

Chapitre 1. L'hématopoïèse

Fig 1. Mise en place du système hématopoïétique chez l'embryon murin.....	22 -
Fig 2. Mise en place du système hématopoïétique définitif chez l'embryon murin	24 -
Fig 3. Modèles de différenciation hématopoïétique	26 -
Fig 4. Composants cellulaires et moléculaires de la niche hématopoïétique	29 -
Fig 5. Organisation schématisée de la cavité médullaire	31 -
Fig 6. Altérations de la niche hématopoïétique dans les leucémies	33 -

Chapitre 2. La leucémie aiguë myéloïde

Fig 7. Des prémices de l'hématologie.....	37 -
Fig 8. ... A l'avènement de l'analyse cytogénétique.....	39 -
Fig 9. Chevauchement fonctionnel des mutations drivers dans les LAM.....	42 -
Fig 10. Modèles d'évolution clonale de la LAM.	42 -
Fig 11. Epidémiologie	44 -
Fig 12. Anomalies moléculaires récurrentes chez les patients LAM	47 -
Fig 13. Aperçu des approches immunothérapeutiques.	53 -

Chapitre 3. Notions d'épigénétique

Fig 14. Organisation de la chromatine.	55 -
Fig 15. Writers, Readers et Erasers.	57 -
Fig 16. La méthylation de l'ADN.....	59 -
Fig 17. Modifications post-traductionnelles des histones.....	61 -
Fig 18. Concept d'enhancer et de super-enhancer.....	65 -
Fig 19. Organisation tridimensionnelle de la chromatine.....	67 -
Fig 20. Perturbation des TADs.....	72 -

Chapitre 4. Le gène homéotique Cdx2

Fig 21. Les gènes HOX et ParaHox	77 -
Fig 22. Le facteur de transcription homéotique CDX2	79 -

Fig 23. Expression de <i>Cdx2</i> dans les cellules externes de la morula.....	81 -
Fig 24. Expression de <i>Cdx2</i> dans la partie caudale de l'embryon	81 -
Fig 25. Expression de <i>Cdx2</i> pendant l'émergence hématopoïétique.....	83 -
Fig 26. Expression de <i>Cdx2</i> dans l'endoderme postérieur	85 -
Fig 27. Conséquences de la perte de fonct. mosaïque de <i>Cdx2</i> dans l'épithélium intestinal .	88 -
Fig 28. Expression aberrante du gène <i>Cdx2</i> dans la moelle osseuse	92 -
Fig 29. Modèle d'expression inductible de <i>Cdx2</i> dans le compartiment hématopoïétique. -	105 -

RÉSULTATS

Partie 1.

Fig. 1. Leukemia development by ectopic expression of <i>CDX2</i> in the hematopoietic lineage.	110 -
Fig. 2. <i>Ex vivo</i> replating activity of <i>MxCDX2</i> bone marrow cells.....	110 -
Fig. 3. Cellular consequences of the ectopic expression of <i>CDX2</i>	111 -
Fig. 4. Molecular consequences of the ectopic expression of <i>CDX2</i>	113 -
Fig. 5. Involvement of <i>BAMBI</i> in the oncogenic effect of <i>CDX2</i>	113 -

Partie 2.

Fig. 6. Expression du gène <i>CDX2</i>	120 -
Fig. 7. ChIP-seq : principe de la préparation de la chromatine.	122 -
Fig. 8. ChIP-seq : PCR de contrôle des ChIP.....	123 -
Fig. 9. Profils du ChIP-seq.....	124 -
Fig. 10. Initiation de l'ATAC-seq.....	125 -
Fig. 11. Profils de l'ATAC-seq.	126 -
Fig. 12. Superposition des profils de ChIP-seq et ATAC-seq	127 -
Fig. 13. Étude du promoteur proximal du gène <i>CDX2</i>	128 -
Fig. 14. Initiation de l'étude par 4C-seq (partie 1).	130 -
Fig. 15. Choix des amorces pour la PCR.....	131 -
Fig. 16. Contrôles de digestion et ligation.....	132 -

TABLEAUX

Tableau 1. Stratification du risque encouru par les patients – Classification Leukemia Net 2017 ...	49 -
Tableau 2. Sondes utilisées pour la PCR quantitative (Applied Biosystems).....	158 -
Tableau 3. Anticorps et conditions utilisées pour les immunoprécipitations de chromatine.	158 -
Tableau 4. Amorces PCR des bibliothèques d'ATAC-seq	159 -
Tableau 5. Amorces PCR des bibliothèques de 4C-seq	159 -

ABBREVIATIONS

A

Ac : Groupe acétyle
ACE : Angiotensin I Converting Enzyme
ACVR1 : Activin A Receptor type 1
ADN : Acide désoxyribonucléique
AGM : Aorte-Gonades-Mésonephros
AML : Acute Myeloid Leukemia
AMPK : Adenosine Monophosphate-activated Protein Kinase
APC : Adenomatous Polyposis Coli
APL : Acute Promyelocytic Leukemia
ARN : Acide Ribonucléique
ARNm : ARN messenger
ASBT : Apical Sodium dependent Bile acid Transporter (SLC10A2)
ASXL1 : ASXL transcriptional regulator 1
ATP : Adenosine triphosphate

B

BARX1 : BARX Homeobox 1
BCL2 : B-Cell Lymphoma 2
BET : Bromodomain Extra-Terminal
BiKE : Bi-specific Killer cell Engagers
BMP : Bone Morphogenetic Protein
BRAF : B-Raf proto-oncogene, serine/threonine kinase

C

Cad : Caudal

CAR : Chimeric Antigen Receptor
CCR : Cancer Colorectal
CD : Cellule Dendritique
CD/31 : Cluster of Differentiation 31
CDK2 : Cyclin Dependent Kinase 2
Cdx/2 : Caudal type homeobox 2
CEBPA : CCAAT/Enhancer-Binding Protein Alpha
CGH : Comparative Genomic Hybridization
CIRC : Centre International de Recherche sur le Cancer
CLEC12A : C-type Lectin domain family 12, member A
CLP : Common Lymphoid Progenitor
CMP : Common Myeloid Progenitor
CpG : Dinucléotides 5'-C-phosphate-G-3'
CSH : Cellule Souche Hématopoïétique
CSH-CT : CSH dites de « court-terme »
CSH-LT : CSH dites de « long-terme »
CSL : Cellule Souche Leucémique
CTCF : CCCTC-binding Factor
CTLA-4 : Cytotoxic T-lymphocyte Antigen 4 (ou CD152)
CXCL/12 : Chemokine Ligand / 12

D

DAMP : Damage-Associated Molecular Pattern
DART : Dual-affinity Re-targeting Antibody
DCA : Dichloroacetic acid
DII : Delta-like protein

DNMT/1 : DNA-Methyltransferase / 1

DNMTi : DNMT inhibitor

E

E : Embryonic (day)

EHT : Endothelial-Hematopoietic Transition

ELN : « European Leukemia Net »
classification

EMP : Erythro-Myeloid Progenitor

EMT : Epithelial-Mesenchymal Transition

EPO : Erythropoietin

ERG : ETS-Related Gene

ERK1/2 : Extracellular signal-Regulated
Kinases 1/2

ETO : Translocation Eight Twenty-One

eRNA : ARN issu d'un enhancer

EryP : Erythroide Progenitor

ES : Embryonic Stem (Cell)

ETV6 : Variant 6 de ETS, alias de TET

ETS : Erythroblast transformation specific

EZH2 : Enhancer of Zeste 2 polycomb
repressive complex 2 subunit

F

FAB : « French-American-British »
Classification

FISH : Fluorescence in situ hybridization

FLI/1 : Friend Leukemia Integration / 1
transcription factor

FLK1 : vascular endothelial growth factor
receptor 2 (VEGFR)

FLT3 : Fms-Like Tyrosine kinase 3

FLT3-ITD : FLT3-Internal Tandem
Duplication

FLT3-TKD : FLT3 mutation in the Tyrosine
Kinase Domain

FLT3L : FLT3 Ligand

FOXD1 : Forkhead box D1

FT : Facteur de Transcription

FXR : Farnesoid X Receptor

G

G-CSF : Granulocyte Colony-Stimulating
Factor

GATA : Séquence ADN "GATA"

GATA/1 : GATA-binding factor / 1

GFI-1B : Growth Factor Independence-1B

GHRL : Ghrelin

GLUT2 : Glucose Transporter 2

GM-CSF : Granulocyte-Macrophage
Colony-Stimulating Factor

GMP : Granulocyte-Macrophage Progenitor

GSX1 : Genomic screened homeobox 1

GTP : Guanosine Triphosphate

H

H : Histone

HAT : Histone Acetyltransferase

HDAC : Histones-Deacetylase

HDACi : HDAC inhibitor

HMA : Hypomethylating Agent

HMT : Histone Methyltransferase

HNF4-α : Hepatocyte nuclear factor-4 alpha

HOX/A9 : Homeobox gene / A9

HTLV1 : Human T-Lymphotropic Virus type
1

I

IDH/1 : Isocitrate Dehydrogenase 1

IFN-γ : Interferon gamma

IL/2 : Interleukine / 2

iNOS : Inducible Nitric Oxide Synthase

IRS2 : Insulin Receptor Substrate 2

ISX : Intestine Specific homeobox

K

K : Lysine

Kb : Kilobase (1000 paires de bases)

KDM/3A : Lysine (K)-specific Demethylase /
3A

KIT : Mast/Stem Cell growth Factor
Receptor (SCFR ou CD117)

KLF1 : Kruppel Like Factor 1

KRAS : Kirsten rat sarcoma viral Proto-
Oncogene, GTPase

KLF4 : Kruppel-like factor 4

L

LAL : Leucémie Aiguë Lymphoblastique

LAM : Leucémie Aiguë Myéloïde

LepR : Récepteur à la Leptine

LLC : Leucémie Lymphocytaire Chronique

LMO2 : LIM domain Only 2

LMP : Lymphomyeloid Progenitor
Lymphomyéloïde

LYL1 : Lymphoblastic Leukemia derived
sequence 1

M

M-CSF : Macrophage Colony-Stimulating
Factor

MAPK : Mitogen-Activated Protein Kinase

MBP : Methyl-Binding Protein

MCL1 : Induced Myeloid Leukemia Cell

differentiation protein 1

MDS : Myelodysplastic Syndrome

Me/1/2/3 : Mono-(1) / di-(2) / tri-(3) /
methylation

MEIS : Myeloid Ecotropic viral Integration
Site 1

MEIS1 : Meis homeobox 1

MEP : Megakaryocyte-Erythroid Progenitor

MEX3A : Mex /3 RNA binding family
member a

MGG : May-Grünwald Giemsa

miRNA : Micro RNA

MLL : Mixed Lineage Leukemia

MPO : Myeloperoxidase

MPP : Multipotent Progenitor

MRD : Minimal Residual Disease

MUC/1 : Mucin / 1

MYC : MYC proto-oncogene

N

Nanog : Nom dérivé de la mythologie
celtique

NFκB : Nuclear Factor kappa B

NHEJ : Non-Homologous End Joining

NK : Natural Killer

NOD : Non-Obese Diabetic

NPM1 : Nucleophosmin 1

NRAS : Neuroblastoma Ras viral oncogene
homolog

NUP98 : Nucleoporin 98 kD

NuRD : Nucleosome Remodeling and
histone Deacetylase complex

O

Oct4 : Octamer-binding transcription factor

OMS : Organisation Mondiale de la Santé

P

PAX/5 : Paired box / 5

Pb : Paire de bases

PCR : Polymerase Chain Reaction

PD1 : Programmed cell Death protein 1

PDX1 : Pancreatic and Duodenal
homeobox 1

PHD : Plant Homeodomain

PML-RAR α : Promyelocytic Leukaemia -
Retinoic Acid Receptor alpha

PR-1 : Pathogenesis-Related group / 1

PRAME : Preferentially expressed Antigen
in Melanoma

PRC2 : Polycomb Repressive Complex 2

Q

qPCR : Quantitative Polymerase Chain
Reaction

R

R : Arginine

R882 : Résidu arginine 882

Rb1 : Retinoblastoma protein

Rbpj : Recombination Signal Binding
Protein For Immunoglobulin Kappa J
Region

RegIV : Regenerating Family Member IV

RHAMM : Receptor for Hyaluronic acid-
Mediated Motility

RT-qPCR : Reverse Transcriptase -
quantitative PCR

RUNX1 : Runt-related transcription factor 1

S

SAM : S-Adenosyl Methionine

SATB1 : AT-rich Sequence-Binding protein
1

SCF : Skp, Cullin, F-box containing complex

SCL : Stem Cell Leukemia gene

SEER : The « Surveillance, Epidemiology,
and End Results Program”

SHH : Sonic Hedgehog

SMAD : Small Mothers Against
Decapentaplegic

SMC/1 : Structural Maintenance of
Chromosomes protein / 1

SMD : Stanford Microarray Database

SNF : Sucrose Non-Fermentable

SOX/2 : SRY-box transcription factor / 2

STAG/1 : Cohesin subunit SA-1

SWI : Switch

T

TAA : Tumor Associated Antigen

TAD : Topologically Associated Domains

TAL1 : T-cell Acute Leukemia protein 1

TCF4 : Transcription Factor 4 (ITF2)

TEA : DNA-binding domain named after
TEF-1 & abaA

TEAD/4 : TEA Domain transcription factor /
4

TET : Ten-Eleven Translocation

TGF- β : Transforming Growth Factor beta

TIM3 : T cell Immunoglobulin and Mucin
domain-containing protein 3

TNF : Tumor Necrosis Factor

TRAF/3 : TNF Receptor-Associated Factor /
3

TriKE : Tri-specific Killer cell Engagers

TSS : Transcription Start Site

U

UTR : Untranslated Region

V

VCAM-1 : Vascular Cell Adhesion protein 1
(CD106)

VE-cadhérine : Vascular Endothelial
cadhérine (CD144)

W

WES : Whole Exome Sequencing

Wnt : Wingless-type MMPTV integration site

WT1 : Wilms Tumor 1

INTRODUCTION

Chapitre 1

L'hématopoïèse

Le système hématopoïétique se compose de populations cellulaires spécialisées, dotées de fonctions uniques comme le transport de l'oxygène ou les défenses immunitaires. Quotidiennement, environ 350 milliards de cellules sanguines spécialisées sont renouvelées chez l'adulte à partir des cellules souches hématopoïétiques (CSH) qui logent dans la moelle osseuse. L'hématopoïèse est un processus physiologique complexe permettant le renouvellement des cellules du sang, pourvues d'une courte durée de vie pour la plupart. Ce remplacement continu et régulé s'effectue sous l'impulsion d'un système complexe d'interactions avec les différents éléments du microenvironnement. Je ferai le point ici sur les connaissances que l'on a sur l'hématopoïèse chez la souris, pour laquelle davantage d'études ont pu être menées grâce à la disponibilité de nombreux modèles murins. L'hématopoïèse chez l'Homme, bien moins connue, semble conservée mais présente cependant certaines spécificités.

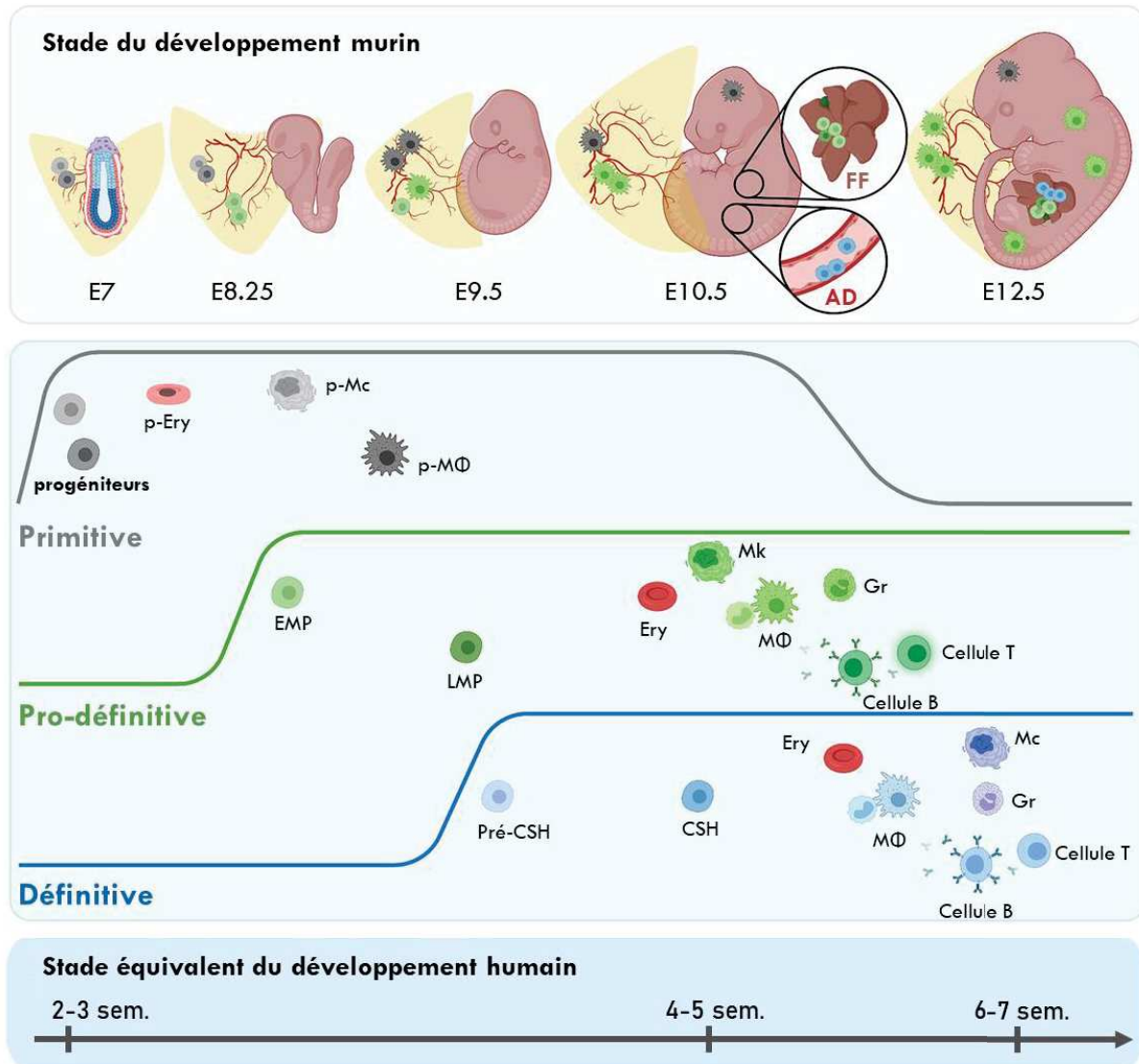


Figure 1. Mise en place du système hématopoïétique chez l'embryon murin

Le développement hématopoïétique s'initie via 3 vagues spatiotemporelles se chevauchant, et produisant des progéniteurs hématopoïétiques distincts. La vague primitive (en gris) génère des érythrocytes (p-Ery), mégacaryocytes (p-Mc) et macrophages (p-MΦ) au niveau du sac vitellin. La vague pro-définitive (en vert) génère des progéniteurs érythromyéloïdes (EMP) et lymphomyéloïdes (LMP) au niveau du sac vitellin. La vague définitive (en bleu) génère des pré-cellules souches hématopoïétiques (pré-CSH) au niveau de l'aorte gonade mésonéphros (AGM). Une fois matures, ces CSH sont capables d'autorenouvellement. Les progéniteurs des vagues pro-définitives et définitives colonisent tous deux le foie fœtal (FF) à E10,5, où ils produisent des cellules matures : érythrocytes (Ery), mégacaryocytes (Mc), granulocytes (Gr), des cellules T et B, et des macrophages (MΦ). E/7 : 7^e jour de développement embryonnaire chez la souris ; sem : semaine de développement embryonnaire chez l'Homme. Figure adaptée de Canu et. al, 2021.

I. Le système hématopoïétique

A. Origine embryonnaire

Le système hématopoïétique est l'un des premiers tissus complexes à se former dans l'embryon. Son apparition prend le relais sur la diffusion passive de l'oxygène et des nutriments, qui devient insuffisante au développement de l'organisme. Il se met en place en 3 vagues chez les mammifères, au niveau de sites anatomiques distincts, avant de se stabiliser dans la moelle osseuse et le thymus (**Figure 1**).

Chez la souris, les premières cellules hématopoïétiques apparaissent dès le 7^e jour de développement embryonnaire (E7), peu après la gastrulation, au sein du sac vitellin, dans des îlots mésodermiques de la splanchnopleure extra-embryonnaire, appelés îlots sanguins. Le sac vitellin est une membrane extra-embryonnaire dérivée de l'endoderme et recouverte de mésoderme extra-embryonnaire. Peu après la gastrulation (vers le 18^e jour), des cellules du mésoderme extra-embryonnaire du sac vitellin s'agrègent en îlots sanguins et vasculoformateurs¹. Ces agrégats mésodermiques sont à l'origine à la fois du système vasculaire et de l'hématopoïèse : un précurseur mésodermique commun, appelé hémangioblaste, semblerait être à l'origine des 2 lignages. Lors de cette hématopoïèse dite « primitive », des cellules hématopoïétiques transitoires sont produites pour remplir les besoins immédiats de l'embryon : des progéniteurs érythroïdes (EryP, larges et nucléés, d'où le terme « primitif ») sont notamment requis pour assurer l'oxygénation de l'embryon, et des macrophages (à ploïdie faible, générant moins de plaquettes) pour permettre le remodelage des tissus et leur maintenance²⁻⁴.

L'hématopoïèse pro-définitive commence à E8,25 à nouveau dans le sac vitellin. Elle génère des progéniteurs érythromyéloïdes (EMPs) et lymphomyéloïdes (LMPs), qui colonisent le foie fœtal, et se différencient en cellules matures⁵. Ces progéniteurs sont issus d'une transition endothélio-hématopoïétique (EHT) : elles dérivent de cellules endothéliales constituant la vascularisation du sac vitellin, dotées d'un potentiel hémogénique^{6,7}.

La 3^e et dernière vague, correspondant à l'hématopoïèse définitive, est marquée par la génération de pré-CSH, entre E9,5 et E11,5, dans la région de l'Aorte-Gonade-Mésonephros (AGM) principalement^{8,9}. L'AGM est un tissu embryonnaire originaire de la splanchnopleure

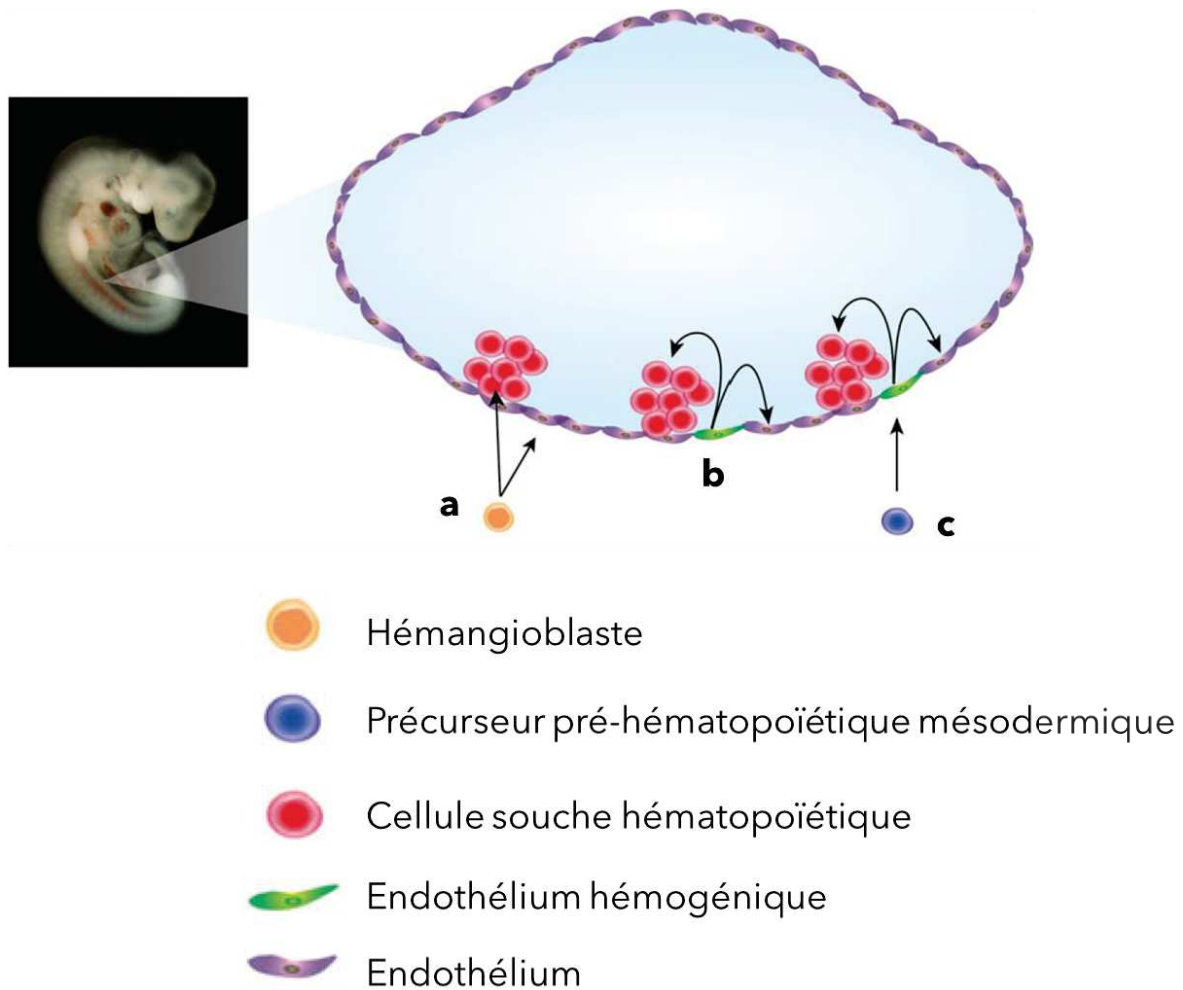


Figure 2. Mise en place du système hématopoïétique définitif chez l'embryon murin

L'origine des cellules souches hématopoïétiques (CSH) intra-aortiques est débattue, 3 modèles sont proposés. **(a)** Un hémangioblaste donne naissance séparément aux CSH et cellules endothéliales. **(b)** Une cellule endothéliale à potentiel hémogénique donne naissance aux CSH et cellules endothéliales. Des preuves de ces deux concepts existent, d'ailleurs ils ne s'excluent pas mutuellement. **(c)** Un mécanisme intermédiaire propose qu'un hémangioblaste (ou précurseur mésodermique pré-hématopoïétique) donne naissance à un endothélium hémogénique intermédiaire, produisant ensuite les CSH et cellules endothéliales. Figure adaptée de Julien et. al, 2016

para-aortique, dérivé de la couche germinale mésodermique, qui contient l'aorte dorsale (artère majeure de l'embryon), la crête génitale et le mésonéphros. Les CSH émergent d'un endothélium dit « hémogénique » au niveau de la paroi ventrale de l'aorte dorsale, via une EHT. L'origine de cet endothélium est encore discutée¹⁰. Le modèle d'émergence privilégié est à nouveau celui de l'endothélium hémogénique, donnant naissance aux CSH via une EHT^{7,11} (**Figure 2**).

Les premières cellules hématopoïétiques d'origine vitelline sont détectées dans le foie fœtal à E10,5. Vers E12,5 celui-ci est colonisé par les CSH qui y subissent une expansion et une différenciation rapide¹². Chez la souris spécifiquement, une transition vers la rate fœtale s'initie vers E15,5, où les CSH poursuivent leur expansion, et dont une fraction subsistera après la naissance et à l'âge adulte¹³. A E17,5 des CSH colonisent enfin la moelle osseuse, où elles assureront la plus grosse partie de l'hématopoïèse tout au long de la vie, tout en étant majoritairement quiescentes¹⁴. Là, elles pourront se différencier en près de 10 types cellulaires distincts via un processus multi étapes et la mise en œuvre d'un programme génique complexe^{4,15}. L'ébauche thymique est colonisée vers E11,5 par des cellules T précurseurs¹⁶. Une fois fonctionnel, le thymus est le site de maturation des lymphocytes T¹⁷.

B. L'Hématopoïèse chez l'adulte

1. Les cellules souches hématopoïétiques

Présentes en très petite quantité (1/10000 des cellules hématopoïétiques de la moelle osseuse), ces cellules sont indifférenciées (dépourvues de marqueurs d'engagement) et présentent à leur surface des marqueurs d'immatunité tels que CD34, SCA-1, CD38, CD117 (c-Kit) permettant de les discriminer des autres populations cellulaires.

Ainsi situées au sommet de la pyramide de différenciation, les CSH sont multipotentes, c'est-à-dire capables de récapituler à elles seules l'entière du système sanguin, par exemple après transplantation chez un sujet irradié. Une fois engagées, les cellules souches vont progressivement perdre cette capacité d'autorenouvellement et réduire leur panel de différenciation. Aussi, l'autorenouvellement des CSH, capacité à donner naissance à deux cellules filles (division symétrique ou asymétrique au besoin), permet le maintien de leur pool¹⁸.

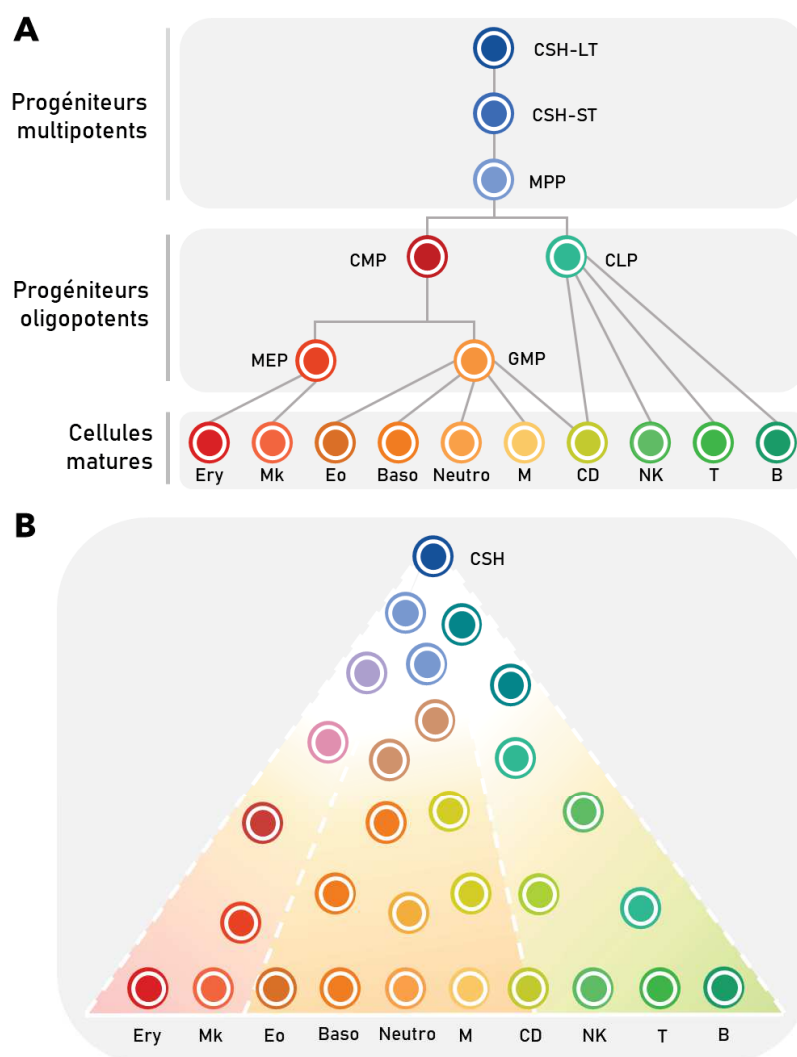


Figure 3. Modèles de différenciation hématopoïétique

(A) Le modèle de différenciation classique des cellules souches hématopoïétiques (CSH) consiste en une hiérarchie branchée, où la différenciation se fait étape par étape selon des décisions binaires. Cette représentation découle d'expériences mesurant le potentiel de transplantation et de formation de colonies. (B) Les nouveaux modèles graphiques représentent la différenciation des CSH comme un processus plus continu : les cellules unipotentes émergeraient directement de cellules souches et progénitrices, sans passer par différentes étapes de potentialité de plus en plus restreintes. CSH-LT : CSH de long terme ; CSH-ST : CSH de court terme ; MPP : Progéniteur multipotent ; CMP : Progéniteur myéloïde commun ; CLP : Progéniteur lymphoïde commun ; MEP : Progéniteur mégacaryocyte-érythroïde ; Ery : Erythrocytes ; Mk : Macrophages ; GMP : Progéniteur des granulocytes et monocytes ; Eo : Eosinophile ; Baso : Basophile ; Neutro : Neutrophile ; M : Monocyte ; CD : Cellule dendritique ; NK : Cellule Natural-killer ; T : Lymphocytes T ; B : Lymphocyte B. Figures inspirées de Antoniani et. al, 2017 et Cheng et. al, 2020.

2. Modèles de différenciation hématopoïétique

Le modèle de différenciation hiérarchique des CSH, proposé pendant plus de 20 ans, suggère un processus séquentiel passant par des stades de progéniteurs multi-, oligo-, et unipotents distincts avant d'atteindre la maturité (**Figure 3A**). Les CSH y sont considérées comme une population homogène, sur la base des marqueurs de surface qu'elles partagent. Une distinction est néanmoins faite dans ce modèle en fonction de leur capacité à reconstituer le système hématopoïétique sur la durée (se traduit par leur potentiel de prise de greffe). Les CSH à la capacité d'autorenouvellement illimitée sont dites de « long-terme » (CSH-LT), elles sont maintenues quiescentes dans la moelle osseuse et se distinguent par l'absence de marqueurs CD34 à leur surface. De ces dernières découlent les CSH dites de « court-terme » (CSH-CT) à la capacité d'autorenouvellement limitée dans le temps. Ce modèle se présente comme suit : les CSH se différencient en progéniteurs multipotents (MPPs), dépourvus de capacité d'auto-renouvellement mais conservant tout leur potentiel de différenciation^{19,20}. La première bifurcation consiste en la différenciation de MPPs en 2 types de progéniteurs oligopotents : les progéniteurs lymphoïdes communs (CLPs) et les progéniteurs myéloïdes communs (CMPs). Ces derniers donnent naissance aux cellules effectrices engagées dans les deux lignages respectifs. Les CMPs évoluent en progéniteurs mégacaryocytes-érythroïdes (MEP) et progéniteurs granulocytes-macrophages (GMP). Les CLPs évoluent en progéniteurs T, B et NK²¹. Les cellules dendritiques (CD) proviennent aussi bien des CLP et CMP²².

Ce modèle de différenciation séquentielle tend à être remplacé par un modèle de différenciation plus « continu », moins cloisonné des CSH, où l'acquisition d'une identité serait graduelle (**Figure 3B**). Les approches *-omics* et *single-cell* des dernières années ont en effet révélé une hétérogénéité au sein des cellules souches et progénitrices²³⁻²⁵. Les CSH semblent se distinguer entre elles par leur capacité de différenciation : certains clones seraient prédestinés à s'engager vers le lignage myéloïde ou lymphoïde. Les marqueurs de surface CD41 et CD86 sont par exemple retrouvés sur les CSH orientées respectivement vers les lignages myéloïde ou lymphoïde^{26,27}. Il en va de même pour les MPPs, subdivisables en 4 populations à la différenciation préférentielle. Ce destin cellulaire, préétabli, serait scellé dès le stade de CSH²⁸.

3. Régulateurs de l'hématopoïèse

Quiescence, auto-renouvellement, expansion et différenciation des CSH vers l'un ou l'autre des lignages sont finement régulés par l'interaction combinée de facteurs de transcription. Dix régulateurs clés de l'hématopoïèse ont été identifiés : SCL/TAL1, LYL1, LMO2, GATA2, RUNX1, MEIS1, PU.1, ERG, FLI-1 et GFI-1B²⁹.

L'organisation de la chromatine, et les modifications d'histones et de l'ARN sont aussi cruciales, puisque ces facteurs déterminent l'activation ou la répression des gènes. L'autorenouvellement, la maintenance et la différenciation des CSH sont ainsi largement influencés par l'épigénome, comme nous le verrons plus tard³⁰.

Les sous-populations de CSH expriment sélectivement les récepteurs de cytokines affiliées à des lignages. On suppose maintenant que des cytokines telles M-CSF, G-CSF, GM-CSF, Epo et Flt3L ne sont pas seulement de simples facteurs de prolifération et de survie des progéniteurs engagés, mais qu'elles contribuent à orienter le devenir cellulaire des CSH³¹. Par exemple, une sous-population de CSH requérant de la thrombopoïétine pour sa maintenance, et donc qui exprime le récepteur associé, donnerait une descendance mégacaryocytaire, à l'origine des plaquettes sanguines, ou en tout cas biaisée vers le myéloïde, si transplantée³². Aussi, le M-CSF ordonnerait un devenir myéloïde via l'induction de l'expression du facteur de transcription PU.1³³. Il a d'ailleurs été observé que des cellules avec un lignage prédéterminé étaient capables d'évoluer en cellules matures différentes du « choix » initial, au contact de cytokines^{34,35}. Cette flexibilité permettrait à l'hématopoïèse de s'adapter à la demande en certains types cellulaires, et renforce l'idée que le développement des CSH est davantage dynamique que linéaire.

II. Microenvironnement et homéostasie

Le microenvironnement médullaire, dans lequel résident les cellules hématopoïétiques, est la source première de facteurs extrinsèques influençant leur différenciation. Les CSH y sont exposées à des cytokines, des facteurs de croissance, des variations de pH et d'oxygène, et des contacts cellulaires.

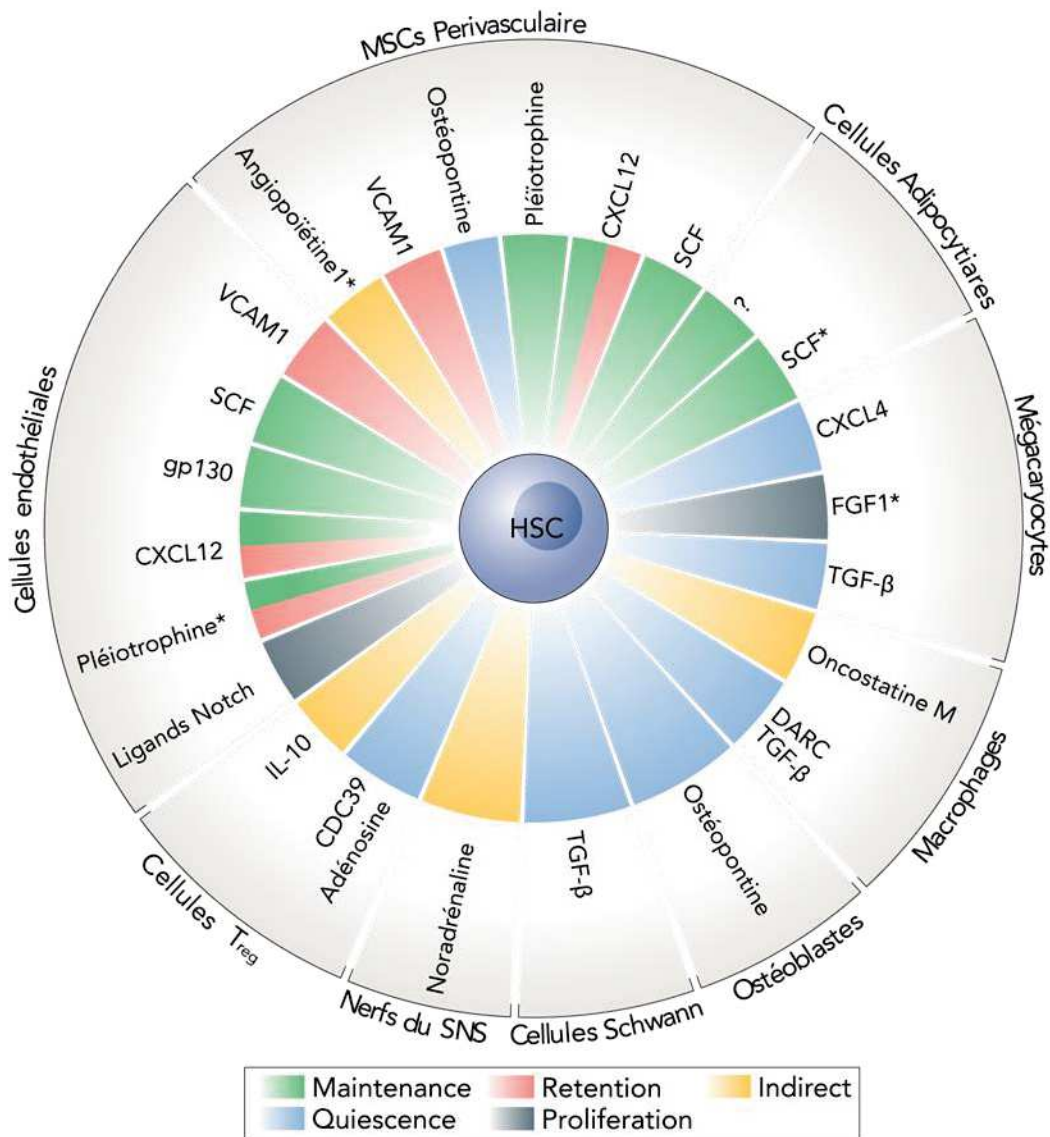


Figure 4. Composants cellulaires et moléculaires de la niche hématopoïétique

Ce schéma illustre comment les cellules de la niche hématopoïétique contribuent à la régulation directe ou indirecte des cellules souches hématopoïétiques (CSH), en sécrétant un panel de molécules. Les couleurs indiquent quel aspect de l'homéostasie elles impactent. Les astérisques indiquent les molécules impliquées dans la régénération de la moelle osseuse après ablation. SCF : Stem Cell Factor; CXCL12 : CXC-Chemokine Ligand 12; CXCL4 : CXC Chemokine Ligand 4; FGF1 : Fibroblast Growth Factor 1; TGF- β : Transforming Growth Factor- β ; DARC : Duffy Antigen Receptor for Chemokines; GP130 : Glycoprotein 130; VCAM1 : Vascular Cell Adhesion Molecule 1; IL-10 : Interleukin-10. Figure adaptée de Pinho et. al, 2019.

A. Régulation des CSH

1. Le concept de niche hématopoïétique

La niche de cellules souches se réfère à l'environnement proche permettant leur maintien au stade indifférencié et leur autorenouvellement, et aiguillant leur devenir. Historiquement, le concept de niche hématopoïétique a été proposé en 1978 par R. Schofield. Elle était décrite comme un réseau multicellulaire fournissant aux CSH les molécules et interactions physiques nécessaires à leur auto-renouvellement, différenciation, expansion et rétention. Il était supposé à cette époque que seules les CSH y étaient maintenues, et que les cellules filles en sortaient pour se différencier. Ce microenvironnement spécialisé se localise dans la cavité osseuse des os longs et axiaux, l'hématopoïèse dans les os longs s'amointrissant au cours de l'adolescence³⁶.

Les structures principales de la niche hématopoïétique sont l'os, le réseau vasculaire et les cellules stromales réticulaires. La vascularisation consiste en des artérioles qui pénètrent dans l'os et s'y transforment en vaisseaux transitionnels, puis en réseau de sinusoides à l'endothélium fenestré, distribués de manière homogène dans la cavité. On trouve également d'autres cellules non hématopoïétiques : fibres nerveuses sympathiques, cellules de Schwann, adipocytes, ostéoblastes, ostéocytes, fibroblastes. Ces cellules et structures coopèrent pour assurer un environnement régulant et organisant l'hématopoïèse³⁷. Les populations cellulaires de la niche, hématopoïétiques ou non, fournissent aux CSH une association de facteurs clés (**Figure 4**).

2. Organisation spatiale

La MO étant très vascularisée, les CSH se retrouvent en relation étroite avec les cellules endothéliales et stromales périvasculaires associées. Ces dernières se distinguent par leur expression du récepteur à la Leptine (LepR) et leur forte production de SCF et CXCL12 ; elles en sont la source principale dans la MO.

Sinusoides et artérioles semblent constituer des niches distinctes : on considère à l'heure actuelle que les CSH actives (en division) résident dans des niches périnusoidales, tandis que celles artériolaires seraient occupées plutôt par des CSH en quiescence^{38,39}.

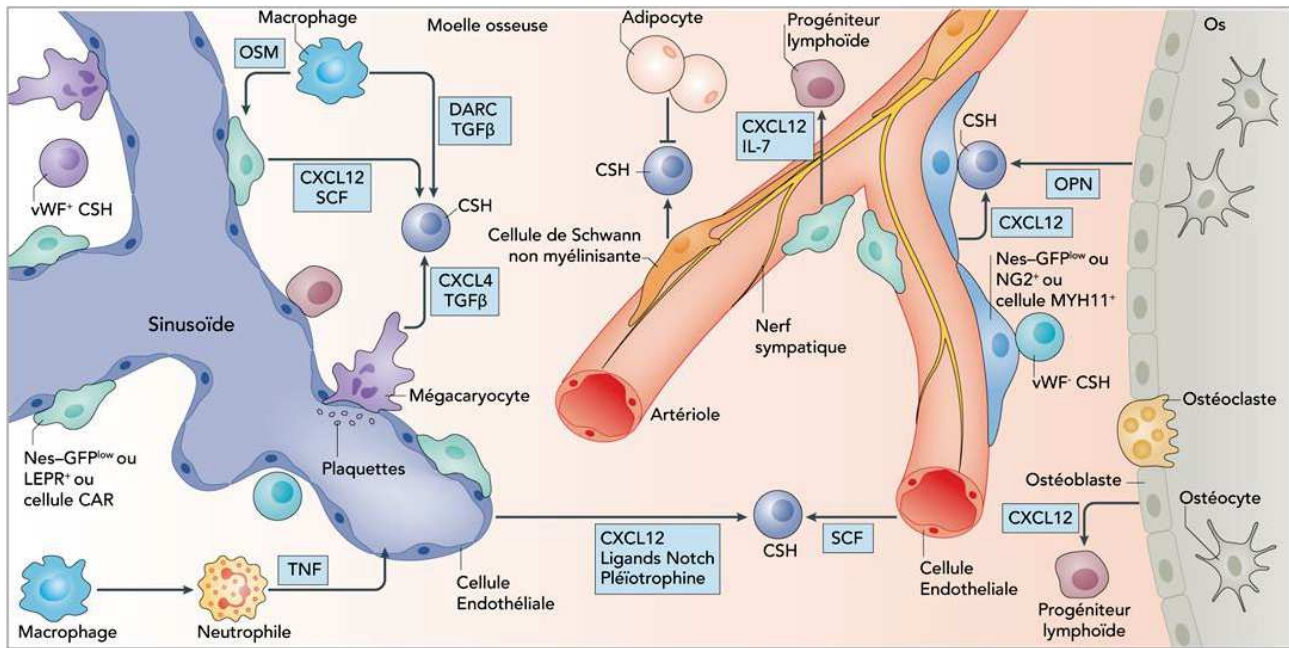


Figure 5. Organisation schématique de la cavité médullaire

Représentation schématique de la niche médullaire de cellules souches hématopoïétiques (CSH), où différents types cellulaires et facteurs régulent directement ou indirectement leur activité. De plus en plus de preuves définissent la vascularisation et cellules stromales associées comme des régulateurs clés de la maintenance des CSH. Le système nerveux sympathique (SNS) régule la mobilisation des CSH, et les cellules de Schwann non myélinisantes pourraient contribuer à leur quiescence. Les mécanismes moléculaires impliquant les ostéoblastes dans la régulation des CSH n'ont pas été clairement identifiés, mais ils pourraient réguler les progéniteurs lymphoïdes. Les adipocytes pourraient négativement affecter la maintenance des CSH. Les cellules hématopoïétiques comme les macrophages à ploïdie faible (générant moins de plaquettes), neutrophiles et mégacaryocytes peuvent également participer à la maintenance ou mobilisation des CSH. Cellules CAR : Cellules Réticulaires Abondantes en CXCL12; SCF : Stem Cell Factor; CXCL4 : C-X-C-Chemokine ligand 4; CXCL12 : CXC Chemokine Ligand 12; DARC : Duffy Antigen Receptor for Chemokines; IL-7 : Interleukin-7; LEPR : Leptin Receptor; MYH11 : Myosin Heavy chain 11; NG2 : Neuron/Glial antigen 2; OPN : Ostéopontine; OSM : Oncostatine M; TGF- β : Transforming Growth Factor- β ; TNF : Tumour Necrosis Factor; vWF : von Willebrand Factor. Figure adaptée de Pinho et. al, 2019.

Les niches artériolaires contiennent notamment des cellules de Schwann non myélinisantes, capable d'activer le TGF-B latent produit par les cellules environnantes, contribuant à la quiescence⁴⁰. L'impact respectif de ces deux niches sur la quiescence des CSH reste controversé^{41,42} (**Figure 5**).

Les fibres nerveuses nociceptives et sympathiques innervant la moelle osseuse régulent la libération circadienne de CSH dans le sang périphérique, cela directement ou par l'intermédiaire des cellules stromales respectivement^{43,44}. Au-delà de leur mobilisation vers l'extérieur de la niche, les CSH présenteraient d'ailleurs une mobilité insoupçonnée en interne, leur permettant de varier leur interactions cellulaires⁴⁵.

L'organisation spatiale de la MO assure que des environnements distincts régulent également différents types de cellules progénitrices. Le lignage ostéoblastique soutient les progéniteurs davantage engagés, et en particulier ceux du lignage lymphoïde^{39,46}. Les CSH les plus actives localisées au niveau des niches périartériolaires seraient biaisées envers le lymphoïde⁴⁷. A l'inverse, les progéniteurs myéloïdes sont enrichis au niveau des sinusoïdes⁴⁸. L'érythropoïèse requiert notamment le SCF produit par les cellules périsinusoïdales LepR⁺⁴⁹. Ces lignages ne colocalisent pourtant pas avec les CSH, ce qui suggère l'existence de différents domaines au sein des niches sinusoïdales, restreintes à certains types cellulaires⁴².

3. Contribution des populations hématopoïétiques

Certaines populations hématopoïétiques persistent dans la moelle osseuse et contribuent à la maintenance des CSH. Les plus étudiées sont les macrophages et les mégacaryocytes. Les macrophages présents se localisent à proximité de sinusoïdes et sont des acteurs clés de la quiescence : l'interaction de leur marqueur membranaire CD234 avec CD82 à la surface des CSH, inhibe leur progression dans le cycle cellulaire⁵⁰. Ils régulent également la rétention des CSH en modulant la production de CXCL12 par les cellules stromales périvasculaires⁵¹. Le CXCL12 environnant stimule également la maturation des mégacaryocytes (thrombopoïèse) et la libération subséquente de plaquettes. Mégacaryocytes et plaquettes sont l'unique source de CXCL4 dans la moelle osseuse, lequel promeut la quiescence des CSH⁵². Leur production de TGF-B1 renforce cet effet^{53(p201)}. Enfin, la perte des mégacaryocytes est associée à une augmentation du cyclage cellulaire des CSH, et en conséquence leur nombre³⁹.

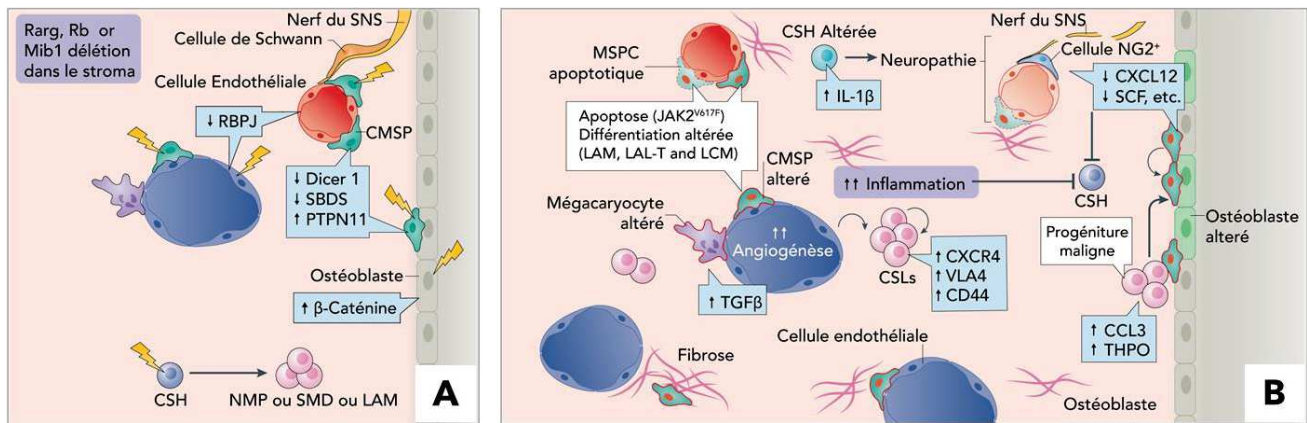


Figure 6. Altérations de la niche hématopoïétique dans les leucémies

(A) Différentes lésions génétiques et épigénétiques (éclairs) peuvent cibler les cellules régulatrices de la niche hématopoïétique, et promouvoir l'apparition d'hémopathies malignes.

(B) Les hémopathies malignes peuvent à l'inverse induire un remodelage de la niche hématopoïétique de sorte à ce qu'elle supporte la leucémogénèse, au détriment de l'hématopoïèse. Les altérations produites par les hémopathies malignes peuvent générer un environnement pro-inflammatoire, caractérisé par de la fibrose, un remodelage vasculaire, une neuropathie et la réduction de sécrétion de facteurs de soutien des CSHs par les cellules stromales. Les cellules souches leucémiques (CSL) peuvent aussi augmenter l'expression de certains de leurs récepteurs membranaires pour favoriser leur survie. LAM : Leucémie Aiguë Myéloïde; CCL3 : CC Chemokine Ligand 3; LCM : Leucémie Chronique Myéloïde; CXCL12 : CXC Chemokine Ligand 12; CXCR4 : CXC Chemokine Receptor type 4; IL-1 β : Interleukin-1 β ; SMD : Syndrome myélodysplasique; MIB1 : Mind Bomb 1; NMP : Néoplasme Myéloprolifératif; NG2 : Neuron/Glial Antigen 2; PTPN11 : Protein Tyrosine Phosphatase, Non-receptor type 11; RB : Retinoblastoma Protein; RBPJ : Recombination signal Binding Protein for immunoglobulin- κ J region; SCF : Stem Cell Factor; SNS : Système nerveux sympathique; TGF β : Transforming Growth Factor- β ; THPO : Thrombopoietin; VLA-4 : Very Late Antigen 4. Figures adaptées de Pinho et. al, 2019.

Les niches hématopoïétiques restent à l'heure actuelle des environnements hétérogènes trop peu compris. Les CSH ne peuvent exister que dans ces microenvironnements physiques les soutenant, mais l'impact de ceux-ci sur leur activité et différenciation présente encore des zones d'ombre. Ne serait-ce que la raison pour laquelle la moelle osseuse reste le site privilégié de l'hématopoïèse, alors que les facteurs clés CXCL12 et SCF sont produits à large échelle dans différents organes, hématopoïétiques ou non.

B. Perturbation de la niche

1. Implication dans les hémopathies

Un nombre croissant d'études suggère que la niche hématopoïétique pourrait être remodelée pour soutenir la leucémogénèse, ou encore être elle-même à l'origine d'hémopathies³⁹ (**Figure 6**).

Les cellules souches leucémiques (CSLs) tendent à remodeler leur environnement en une unité oncogénique favorisant la croissance tumorale au détriment de l'hématopoïèse normale^{54,55}. Les CSLs sécrètent effectivement des facteurs de croissance altérant la différenciation des cellules souches et progénitrices, qui en conséquence s'accumulent et ne sont plus capables d'assurer une hématopoïèse efficace.

Le dysfonctionnement des composants de la niche peut également prédisposer ou contribuer à l'initiation d'hémopathies malignes. Dans les cellules stromales de patients atteints de MDS et LAM, des anomalies génétiques et épigénétiques, différentes de celles des cellules hématopoïétiques, ont été observées⁵⁶⁻⁵⁸. Chez 38% des patients atteints de MDS ou LAM, une activation constitutive de la β -caténine dans les ostéoblastes et une hausse de la signalisation Notch dans les CSH sont observées, et pourraient être responsables de la leucémogénèse. Un modèle murin vient appuyer cette observation, puisque l'introduction chez celui-ci d'une mutation activatrice de la B-caténine dans les ostéoblastes, altère la différenciation myéloïde et lymphoïde et conduit au développement d'une LAM⁵⁹. Le phénomène s'illustre au travers de diverses autres mutations opérées dans les cellules stromales. Il a par exemple été observé que la délétion chez la souris de l'endonucléase Dicer1 dans le lignage ostéoblastique endigue la différenciation des ostéoblastes, et déclenche un syndrome myélodysplasique pouvant évoluer

en LAM⁶⁰. De même, la délétion du gène Rb1 génère un néoplasme myéloprolifératif, qu'elle soit effectuée dans les cellules stromales ou myéloïdes⁶¹.

2. Hématopoïèse extramédullaire

Des sites anatomiques autres que la moelle osseuse possèdent un potentiel hématopoïétique. Comme mentionné précédemment, une hématopoïèse hépatique se met transitoirement en place pendant le développement fœtal. En postnatal, chez l'Homme, en cas de stress sévère ou d'altération médullaire, une hématopoïèse splénique peut s'initier dans la pulpe rouge. Des cellules endothéliales et périvasculaires stromales y produisent le CXCL12 nécessaire à la rétention des CSH, tandis que des macrophages leur prodiguent un ancrage via la molécule d'adhésion VCAM-1^{62,63}. Cette hématopoïèse extramédullaire peut se déclencher par exemple lors d'un besoin accru en érythrocytes pendant la grossesse ou en cas d'anémie, ou encore lors d'infection, de fibrose ou de néoplasie.

Chapitre 2

La leucémie aiguë myéloïde

Avant de rentrer dans le vif du sujet, voici quelques notions clés pour resituer les leucémies aiguës myéloïdes (LAM) par rapport aux hémopathies malignes.

Les hémopathies malignes sont des proliférations tumorales clonales, qui se développent à partir de cellules hématopoïétiques, dans la moelle osseuse, le thymus ou les organes lymphoïdes secondaires. Elles désignent un groupe hétérogène de maladies, qui se distinguent les unes des autres par la lignée hématopoïétique dont est issue la cellule tumorale, ainsi que de son état de différenciation et la localisation de la prolifération. Il devient question de « leucémie » lorsque des cellules hématopoïétiques cancéreuses sont retrouvées dans le sang circulant des patients.

Les leucémies peuvent se développer à partir de cellules issues de la lignée lymphoïde ou de la lignée myéloïde. Celles développées à partir de la prolifération d'un progéniteur lymphoïde ou myéloïde, situé dans la moelle osseuse et dont la différenciation est bloquée à un stade précoce ont une évolution rapide ; il s'agit de leucémies dites « aiguës ». Les leucémies développées à partir de cellules lymphoïdes ou myéloïdes matures (après sortie de la moelle ou du thymus) ont un développement généralement plus lent ; il s'agit de leucémies dites « chroniques ».

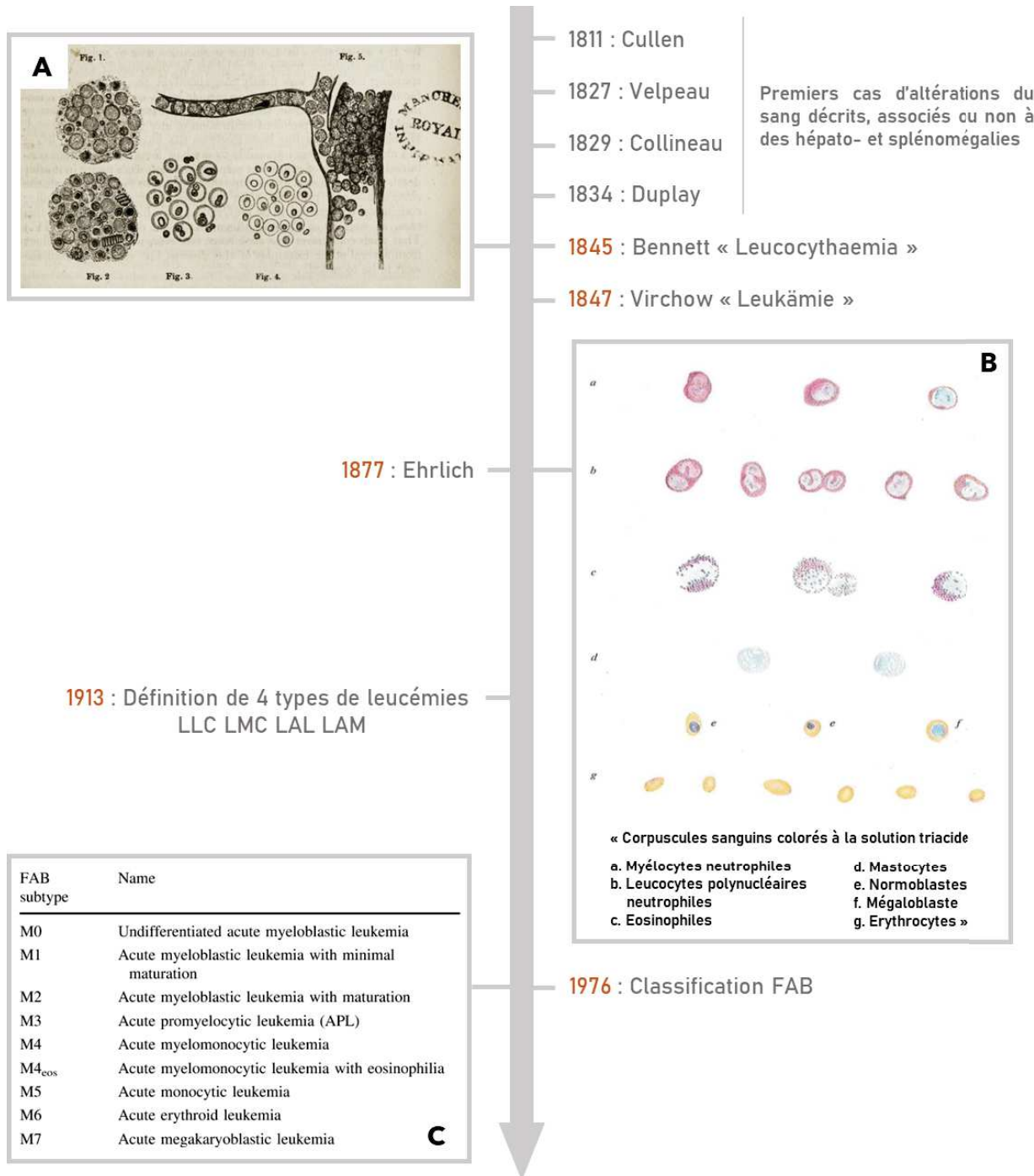


Figure 7. Des prémices de l'hématologie...

(A) Illustration de la « leucocythaemia » décrite par J.H. Bennett en 1845 (Bennett 1851). (B) La coloration triacide dévoilée en 1877 par Paul Ehrlich permet la distinction des cellules basophiles, éosinophiles et neutrophiles sur la base de l'affinité des colorants (Ehrlich 1880). (C) Classification FAB des leucémies aiguës myéloïdes. Elle tient compte des caractéristiques cytologiques présentées par les cellules immatures, appelées « blastes », qui renseignent sur leur lignée d'origine et leur degré de maturation. Les LAM sont alors répertoriées en 8 groupes. (Bennett 1976).

I. Historique et Classification des LAM

A. Découverte des leucémies

La première description de la leucémie est faite en 1811 par le chirurgien Peter Cullen, confronté à un patient de 35 ans présentant des douleurs abdominales et une forte fièvre. Au-delà d'une « *splenitis acutus* », il relève la coloration laiteuse de son sang lors des saignées qui lui sont prodiguées, mais l'associera davantage à une absorption graisseuse qu'à une atteinte hématopoïétique⁶⁴. Il s'agissait là probablement du premier cas de leucémie décrit dans la littérature.

Les descriptions de sang « laiteux » par diverses personnalités se répètent dans la littérature les années suivantes. La morphologie des cellules observées gagne en précision, grâce aux progrès de la microscopie, et est illustrée pour la première fois par John Hughes Bennett en 1845⁶⁵ (**Figure 7A**). La même année, Rudolf Virchow décrit une perturbation de l'équilibre entre les populations de globules blancs et globules rouges des patients⁶⁴. Deux ans plus tard, il baptise le phénomène « leukämie » (leukos = blanc / aima = sang).

B. Classification des leucémies

Avant de chercher à diagnostiquer, traiter et étudier une pathologie, celle-ci doit être décrite, définie et nommée. Les pionniers de l'analyse anatomoclinique ont rapidement cherché à classer les pathologies décrites, sur la base d'observations morphologiques. L'élaboration d'une coloration histochimique par Paul Ehrlich en 1877 permet, quelques années plus tard, la distinction des lignées myéloïdes et lymphoïdes, et de leurs sous-populations (granulocytes éosinophiles, neutrophiles et basophiles)⁶⁶ (**Figure 7B**). Cette avancée technique mène en 1913 à la reconnaissance de 4 types de leucémies, selon les populations cellulaires concernées : la leucémie lymphocytaire chronique (LLC), la leucémie myéloïde chronique (LMC), la leucémie aiguë lymphoblastique (LAL) et la leucémie aiguë myéloïde (LAM)⁶⁷.

Un consortium d'experts de la communauté internationale d'hématologie met en place en 1976 la classification « French-American-British » (FAB) (**Figure 7C**), reposant sur des critères cliniques et cytologiques. Cette classification n'a qu'une faible valeur pronostic, mais sera utilisée pendant 25 ans⁶⁸.

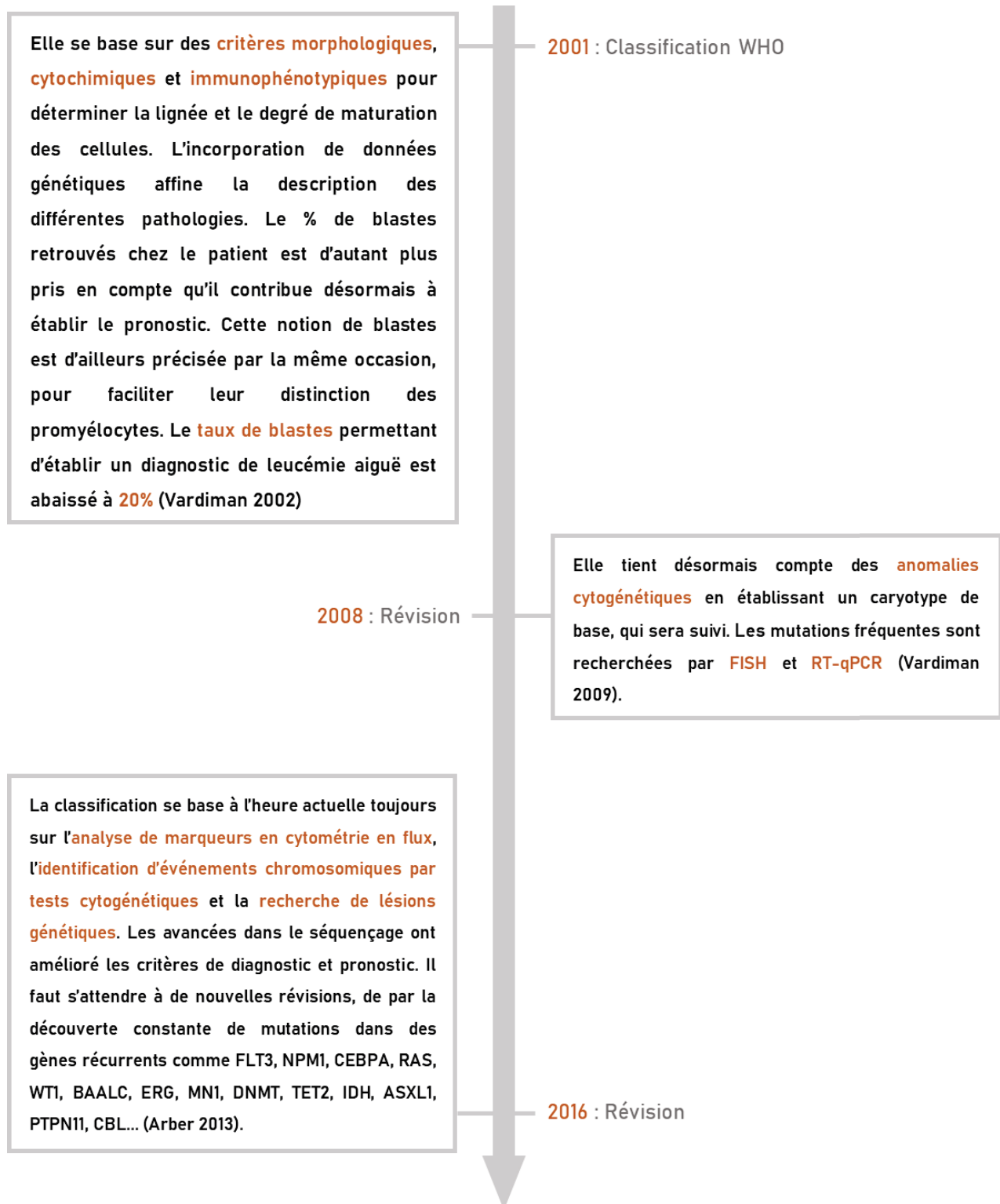


Figure 8. ... A l'avènement de l'analyse cytogénétique

L'OMS publie en 2001 la 3^e version d'une classification proposée par un comité d'experts (*Myeloid and Acute Leukemia Clinical Advisory Committee*). Elle constitue le premier vrai consensus mondial quant à la définition et la nomenclature des hémopathies malignes⁶⁹. (**Figure 8A**). La révision de 2008 apporte des précisions dans la description des hémopathies connues, et distingue quelques nouveaux groupes : cela toujours dans cette optique d'améliorer la terminologie utilisée dans les essais cliniques et la recherche⁷⁰. Une approche multiparamétrique devient progressivement possible, grâce à la contribution de cliniciens et généticiens (**Figure 8B**). La révision de 2016 intègre de nouvelles données immunophénotypiques, cytogénétiques et surtout moléculaires afin de préciser encore ces classifications⁷¹. (**Figure 8C**).

Ces classifications s'étoffent de nouvelles informations à mesure des avancées technologiques et scientifiques. Leurs révisions tentent de rendre les catégories cliniquement plus pertinentes pour faciliter la gestion des patients et aiguiller vers une meilleure prise en charge.

II. Mécanismes physiopathologiques des LAM

Quelle que soit leur origine, les cancers partagent certaines propriétés permettant leur développement : insensibilité aux signaux anti-croissance, indépendance vis-à-vis des facteurs de croissance, résistance à l'apoptose, prolifération illimitée non contrôlée, induction de l'angiogenèse, invasion tissulaire, et capacité à se disséminer.

La LAM n'échappe pas à la règle et se caractérise par une perte de capacité à répondre aux signaux régulateurs normaux de la prolifération et de la différenciation normale⁷². Les cellules reprogrammées qui prolifèrent ne sont plus capables de poursuivre leur maturation et envahissent la moelle osseuse, conduisant à un défaut de production des cellules sanguines saines.

A. Paysage mutationnel des LAM

Bien que les LAM s'illustrent particulièrement par leurs anomalies chromosomiques structurales, des études génomiques à grande échelle, à partir de cellules leucémiques de patients, ont révélé l'existence d'anomalies génétiques récurrentes. Celles-ci sont globalement groupables par fonctions dans lesquelles elles sont impliquées.

Parmi les plus courantes figure la mutation de la protéine chaperonne nucléolaire NPM1, dans 30% des LAM, et 45-66% des LAM avec un caryotype normal. Sa mutation entraîne son export anormal dans le cytoplasme, et l'empêche notamment de remplir sa fonction stabilisatrice de P53. On suppose également que sa localisation cytoplasmique aberrante entraîne la séquestration de PU.1, régulateur de gènes hématopoïétiques comme MEIS1 et les HOX^{73,74}.

Le gène suppresseur de tumeur *P53* est muté dans moins 10% des LAM. Il s'agit généralement de mutations faux-sens, aboutissant à une perte de son activité, et donc favorisant l'instabilité génomique, comme dans la majorité des cancers^{73,74}.

Parmi les plus répandues (30% des LAM), la mutation du gène *FLT3*, codant pour un récepteur à activité tyrosine kinase, conduit quelle que soit la mutation (*FLT3-ITD* dans 20-25% des LAM / *FLT3-TKD* dans 5-10%) à son activation constitutive. Celle-ci promeut la prolifération et la survie cellulaire, au détriment de la différenciation. Il en est de même pour les protéines GTPase *NRAS* et *KRAS* (10 à 25%) classiquement touchés par des mutations ponctuelles activatrices, ou encore le récepteur *KIT* (ou *CD117*) (10%) concerné par des mutations de gain de fonction^{73,74}.

Il n'est pas rare que des régulateurs épigénétiques soient mutés. C'est le cas par exemple d'enzymes clés de la déméthylation de l'ADN comme *DNMT3A* (30% des LAM), *IDH1* (5-15%), *IDH2* (10-15%) et *TET2* (9-23%), ciblées par des mutations inactivatrices. La méthyltransférase *EZH2* est elle aussi ciblée par des mutations de perte de fonction dans les LAM⁷³⁻⁷⁵. Cette catégorie sera étayée au chapitre 3.

Des facteurs de transcription hématopoïétiques clés sont aussi touchés. C'est le cas de *RUNX1*, muté ponctuellement dans 5 à 15% des LAM, et de *CEBPA*, facteur de transcription gouvernant la différenciation myéloïde, touché par des mutations de perte de fonction dans 10% des LAM^{73,74}. Certains gènes *HOX* sont également mutés ou impliqués dans des translocations, comme par exemple la fusion *NUP98-HOXA9*^{76,77}

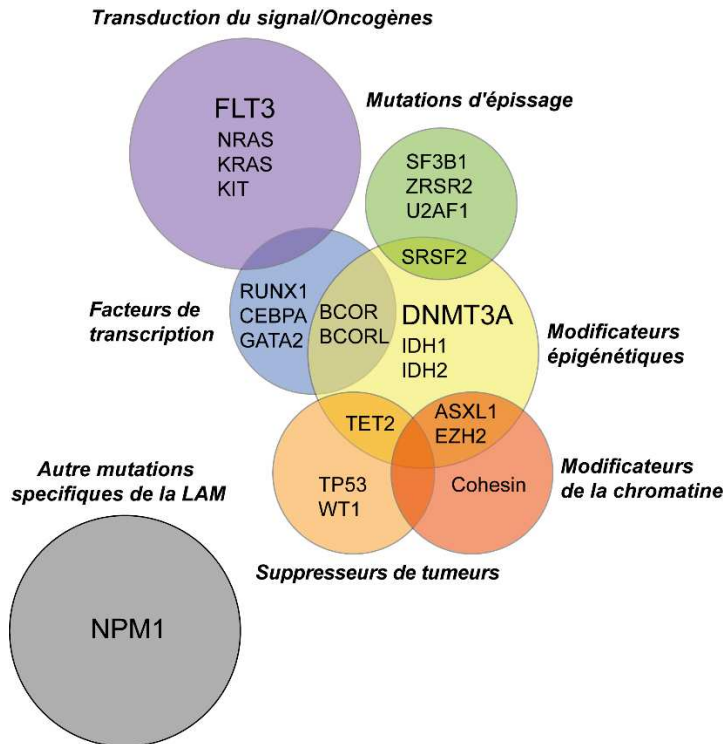


Figure 9. Chevauchement fonctionnel des mutations drivers dans les LAM.

Chevauchement fonctionnel des mutations drivers dans les LAM. La taille des cercles reflète leur fréquence. Les 3 plus communes sont en gras. Extrait de Kishtagari et. al, 2020.

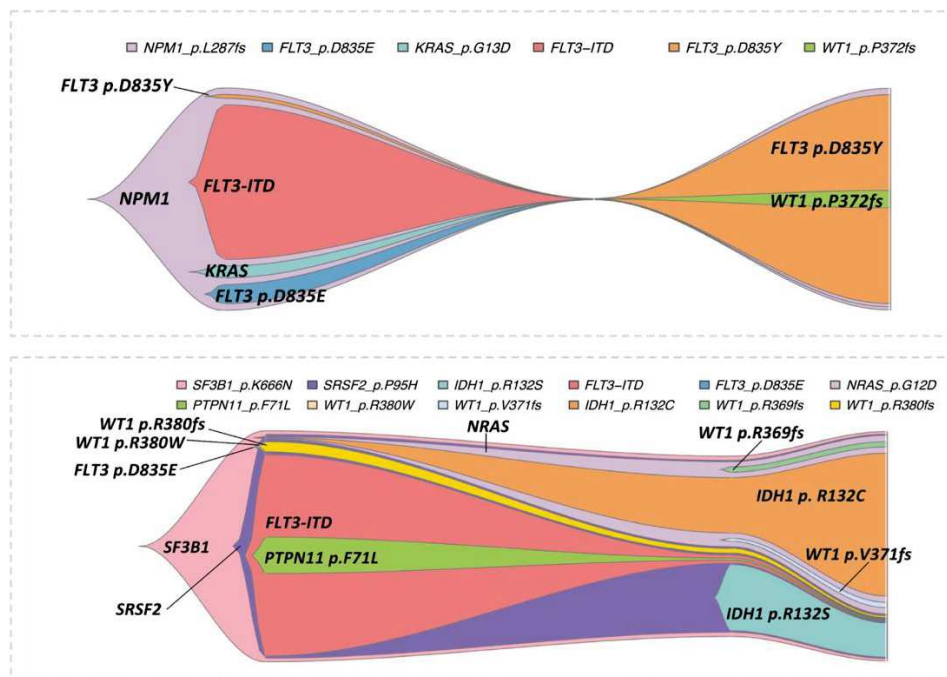


Figure 10. Modèles d'évolution clonale de la LAM.

Ces Fish plots montrent des modèles d'évolutions clonales de leucémies aiguës myéloïdes, en réponse à des thérapies déduites de données d'analyse en single-cell. Les mutations drivers sont représentées à gauche (NPM1 et SF3B1), on constate par la suite l'émergence et la fluctuation des mutations passagers. Figures extraites de Morita et. al, 2020.

Les mutations qui participent à l'initiation de la leucémogénèse (*drivers*) se répartissent en 7 catégories, en fonction de leurs conséquences fonctionnelles (**Figure 9**). Ces mutations partagent des co-occurrences ou dissociations exclusives qui aident à identifier le chemin emprunté par la LAM. Des mutations *passengers* vont ensuite apparaître au hasard, du fait de la fidélité limitée des mécanismes de réparation de l'ADN⁷³.

B. Modèles de pathogenèse

1. Notion de cellules souches leucémiques

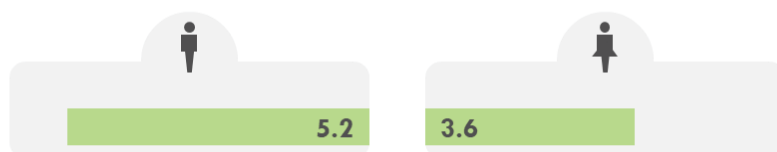
A l'émergence stochastique de la LAM (combinaison de facteurs intrinsèques et extrinsèques qui mènent à la pathologie), s'ajoute un modèle hiérarchique basé sur l'existence de cellules souches leucémiques (CSL). Ces clones dotés d'une capacité d'autorenouvellement, sont capables d'initier et de maintenir la leucémie, mais aussi de la réinitier après transplantation chez un modèle murin immunodéficient. L'origine de ces CSL est débattue puisqu'il pourrait très bien s'agir d'une CSH ayant accumulé des mutations successives lui conférant un potentiel prolifératif, ou d'un progéniteur multipotent ayant réacquis la capacité d'autorenouvellement. On parvient à identifier les CSL sur la base de marqueurs qui leur sont relativement spécifiques (CD123, CD96, CLEC12A, TIM3). De par leur état de quiescence et leur localisation dans la région endostéale de la niche hématopoïétique, les CSL sont particulièrement résistantes à la chimiothérapie, et contribuent à la rechute⁷⁸.

2. Notion d'hétérogénéité et évolution clonale

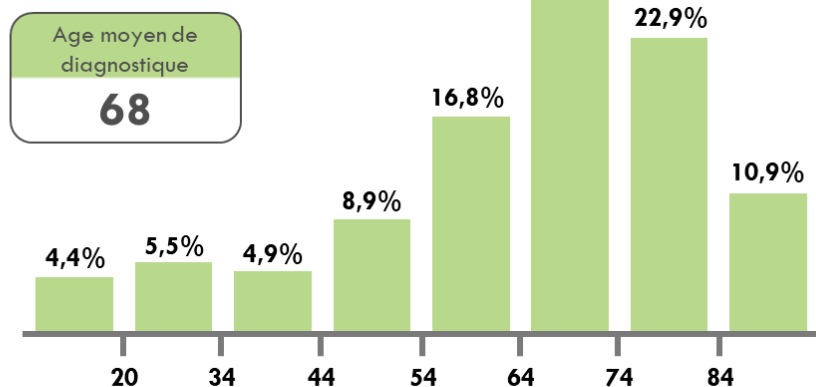
La LAM est une maladie polyclonale qui évolue, notamment à travers la sélection de clones par le traitement mis en place. Au sein de la population tumorale, trois sous-populations de cellules leucémiques peuvent être distinguées : la population principale (cluster fondateur), la population éliminée par le traitement, et la population qui est à l'inverse sélectionnée. Le traitement peut en effet conduire à une évolution dynamique : elle est dite « longitudinale » si le clone pathologique est éliminé, et « branchée » s'il s'en suit l'acquisition de nouvelles mutations ou la sélection de clones résistants. Les leucémies secondaires sont souvent provoquées par la sélection d'un clone prenant le dessus⁷⁹.

Les techniques de séquençage de cellules uniques (*single cell sequencing*) ont permis de mettre en évidence cette hétérogénéité cellulaire et architecture clonale. L'analyse de 37 mutations *drivers* courantes a ainsi permis de modéliser l'évolution clonale des LAM, et de confirmer l'acquisition séquentielle de mutations au cours de la leucémogénèse. Ces données confirment aussi la co-occurrence (DNMT3, TET2, WT1 par exemple) ou l'exclusion mutuelle (IDH1 et IDH2 par exemple) de certaines d'entre elles ⁸⁰ (**Figure 10**).

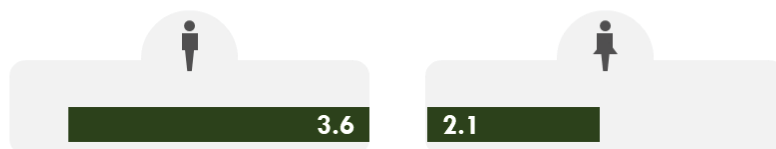
▪ **Taux de nouveaux cas pour 100 000 personnes**



▪ **Nouveaux cas par tranche d'âges**



▪ **Taux de décès pour 100 000 personnes**



▪ **Décès par tranche d'âges**

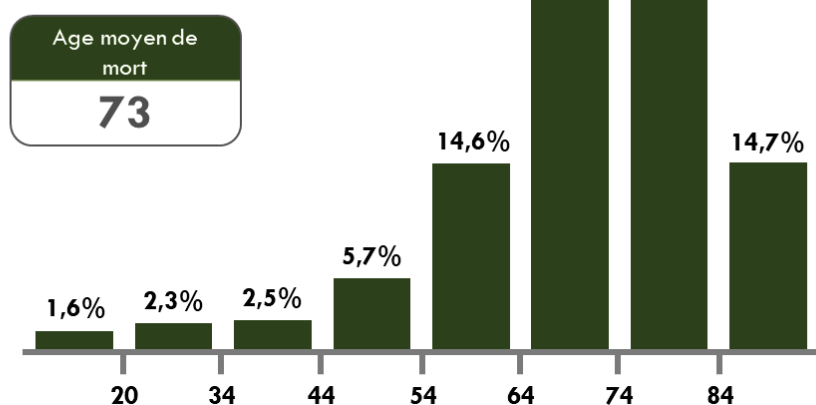


Figure 11. Épidémiologie

Adapté de *The Surveillance, Epidemiology, and End Results (SEER) Program, NIH*

III. La leucémie aiguë myéloïde, à l'échelle du patient

A. Epidémiologie des LAM

1. Statistiques

Chez l'adulte 80% des leucémies aiguës sont de forme myéloïde (LAM) alors qu'elles ne représentent que 20% des cas chez l'enfant qui développe essentiellement des leucémies aiguës lymphoblastiques (LAL). En 2020, 20 240 nouveaux cas de LAM ont été répertoriés aux Etats-Unis, ainsi que 11 400 décès associés à ces hémopathies. L'incidence est croissante avec l'âge, et on constate une prédominance aussi bien en termes d'incidence que de décès chez les patients de sexe masculin (1,5x). Le taux de survie à 5 ans de ces patients est actuellement de 29,5% et est en constante augmentation (*The Surveillance, Epidemiology, and End Results (SEER) Program, NIH*) (**Figure 11**).

2. Facteurs de risques

Un facteur de risque est une composante personnelle, génétique ou de l'environnement qui augmente les chances de développer une maladie.

a. Prédispositions héréditaires

Une fréquence plus importante de LAM est décrite dans certaines maladies génétiques constitutionnelles^{81,82}. Parmi les plus notables se trouvent le syndrome de Down⁸³ et l'anémie de Fanconi⁸⁴, augmentant tout deux l'incidence de la maladie de près de 20 fois. La prédisposition conférée par certains syndromes s'explique souvent par les mutations génétiques qui y sont associées, comme celles de NRAS et KRAS dans les rasopathies⁸².

b. Facteurs iatrogènes

Les agents de chimiothérapie induisent un risque accru de développer une leucémie, qu'elle soit associée ou non avec une radiothérapie^{85(p199)}. La LAM est la forme de cancer secondaire la plus répandue à la suite de traitements chimiothérapeutiques, 10-15% des malades la développent suite à l'exposition à des agents alkylants ou des inhibiteurs de topoisomérases II, qui induisent l'apparition de mutations. Les LAM secondaires présentent

globalement une mauvaise réponse aux chimiothérapies intensives, et une faible survie globale des patients⁸⁶.

c. Agents leucémogènes

Aussi, certaines activités professionnelles exposent quotidiennement à des agents cancérigènes de diverses formes. Benzène et formaldéhyde figurent parmi les plus connus pour augmenter la survenue de leucémies^{87,88}. Viennent ensuite pesticides, fertilisants, et nombre de substances chimiques répertoriées de cancérigènes et recensés par le Centre International de Recherche sur le Cancer (CIRC ou IARC). Une exposition indirecte à ces derniers peut également conduire au développement d'une leucémie.

Certains virus sont reconnus comme leucémogènes, que leur action soit directe (intégration de matériel viral qui altère la transcription de gènes ou expression d'oncogènes viraux) ou indirecte (inflammation chronique ou immunosuppression). HTLV1, premier rétrovirus humain découvert, prédispose par exemple aux leucémies via l'insertion de son ADN/ARN dans les cellules hôtes^{89,90}. Certaines toxines fongiques sont également carcinogènes, comme l'aflatoxine, possiblement via leur action immunosuppressive, mais peu de données sont à disposition quant à leur possible mécanisme d'action⁹¹.

Par ailleurs, une exposition à des radiations peut être à l'origine de leucémies. Seules les radiations ionisantes démontrent avoir une incidence, mais pas les rayons X et les radiations non ionisantes⁹¹.

d. Facteurs individuels

La fréquence de survenue des LAM augmente avec l'âge⁹². Aussi, le tabagisme - pointé du doigt depuis maintenant plusieurs décennies quant à son implication dans le cancer⁹³ - accroît le risque de LAM de 25 à 40%, s'il est toujours d'actualité⁹⁴. L'obésité est elle aussi reconnue depuis peu comme facteur de risque, son impact sur le pronostic étant encore controversé^{95,96}.

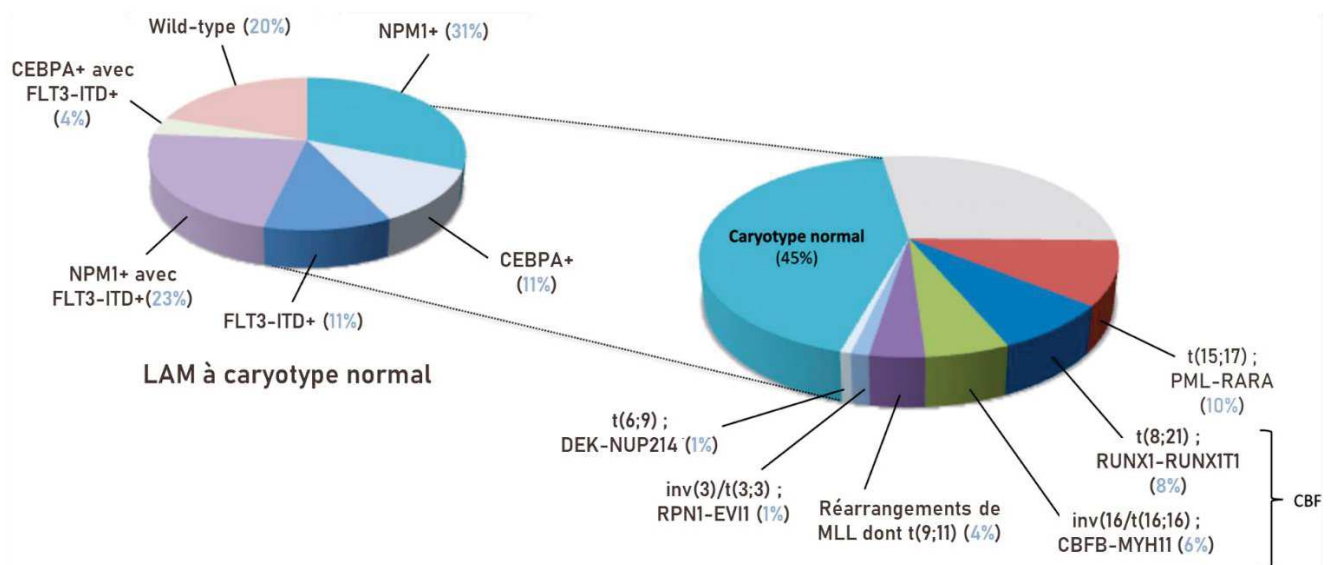


Figure 12. Anomalies moléculaires récurrentes chez les patients LAM

Les translocations récurrentes sont un mécanisme privilégié d'activation d'oncogènes dans les leucémies, et leur importance est maintenant admise. 55% des leucémies aiguës myéloïdes sont concernées. Figure extraite de Duployez et. al, 2015.

B. Diagnostic et classification des LAM

1. Symptomatologie

La symptomatologie est large et variée, et peut se présenter sous forme d'anémie (fatigue, pâleur cutanéomuqueuse, dyspnée), thrombopénie (syndrome hémorragique), neutropénie (syndrome infectieux avec fièvre et infections à répétition), voir éventuellement d'hypertrophie gingivale, de leucémides, de douleurs osseuses et d'hypertrophie des ganglions, de la rate ou du foie⁹⁷. Ces symptômes reflètent l'insuffisance médullaire liée à l'envahissement par des cellules malignes.

2. Examens de dépistage et diagnostic

Le diagnostic s'initie par un hémogramme : une analyse du sang périphérique qui décèle l'insuffisance médullaire au travers de cytopénies, et/ou la présence de blastes en quantité élevée, dont la présence est confirmée par l'étude cytologique d'un frottis sanguin coloré au MGG. Le myélogramme permet d'affirmer le diagnostic de leucémie aiguë lorsque le pourcentage de cellules immatures blastiques est supérieur à 20% dans la moelle osseuse. La lignée d'origine des blastes est ensuite déterminée par analyses cytochimiques et immunocytochimiques. Une analyse par cytométrie en flux vient préciser leurs caractéristiques phénotypiques.

L'analyse du caryotype recherche des anomalies de nombre ou de structure des chromosomes en lien avec la pathologie (retrouvées dans 50-60% des cas). Cette analyse peut être complétée par une étude à l'aide de sondes fluorescentes spécifiques de certaines régions chromosomiques (FISH). Ces analyses cytogénétiques mettent en évidence des inversions, translocations, duplications ou délétions spécifiques de certaines LAM. Des analyses de biologie moléculaire recherchent des équivalents moléculaires de translocations (gènes de fusion) par PCR, et des anomalies d'expression de gènes ou de transcrits de fusion par PCR quantitative (RT-qPCR)^{73,98}. Enfin, des analyses par séquençage haut débit recherchent des variants délétères (mutations, insertions, délétions) dans des panels de gènes (**Figure 12**).

Table 1
Cytogenetic and molecular prognostic factors in acute myelogenous leukemia (AML).

Risk category	Cytogenetics	Molecular abnormalities
Favorable risk	t(8;21)(q22;q22.1) inv(16)(p13.1;q22) or t (16;16)(p13.1;q22)	RUNX1-RUNX1T1 CBFB-MYH11 Mutated NRM1 w/o FLT3- ITD or with FLT3-ITD ^{low} Biallelic mutated CEBPA
Intermediate risk	T(9;11)(p21.3;q23.3) Cytogenetic abnormalities not in other category	Mutated NPM1 and FLT3ITD ^{high} Wild-type NPM1 w/o FLT3-ITD or with FLT3- ITD ^{low}
Poor risk	T(6;9)(p23;q34.1) T(v;11q23.3) T(9;22)(q34.1;q11.2) inv(3)(q21.3;q26.2) or t(3;) (q21.3;q26.2) -5 or del(5q); -7; -17/abn (17p)	MLLT3-KMT2A DEK-NUP214 KMT2A rearranged BCR-ABL1 GATA2, MECOM(EVI1) Complex karyotype Wild-type NPM1 and FLT3-ITD ^{high} Mutated RUNX1 Mutated ASXL1 Mutated TP53#

Data from Estey,¹⁴ Rockova et al,¹⁵ and O'Donnell.¹⁶

Tableau 1. Stratification du risque encouru par les patients - Classification Leukemia Net 2017

3. Etablissement du pronostic

Ces différentes analyses permettent de préciser le diagnostic, et de rechercher des mutations dans des gènes actionnables par des thérapies ciblées. Elles permettent également de rechercher les anomalies moléculaires permettant le suivi de la maladie résiduelle après traitement. Publiée pour la première fois en 2010, la classification *European Leukemia Net (ELN) 2017*³ stratifie le risque encouru par les patients en 3 niveaux, selon les anomalies cytogénétiques ou moléculaires identifiées^{73,99} (**Tableau 1**).

C. Traitements de la LAM

1. Thérapie conventionnelle

a. Chimiothérapie d'induction

La prise en charge classique de la LAM est une chimiothérapie intensive, dite d « induction ». Un ensemble de facteurs sont pris en compte pour déterminer l'éligibilité d'un patient à ce traitement. A l'échelle de l'individu, non seulement l'âge, mais la condition physique, les comorbidités ou encore les objectifs thérapeutiques visés et les préférences du patient sont observés. A cela s'ajoutent les caractéristiques de la maladie, prises en compte au travers des anomalies cytogénétiques dépistées, qui stratifient la réponse au traitement¹⁰⁰. Le traitement conventionnel, employé depuis maintenant plusieurs décennies, consiste en l'infusion continue de cytarabine (100 à 200mg/m²) pendant 7 jours, et d'anthracycline (daunorubicine ou idarubicine) pendant 3 jours. On s'y réfère communément en tant que « 7+3 »¹⁰¹. Malgré les avancées dans les systèmes de délivrance, la limite principale des chimiothérapies reste leurs comorbidités et haute toxicité¹⁰².

La chimiothérapie d'induction est une étape décisive et vise à atteindre le stade dit de « rémission complète » qui se traduit par la présence de moins de 5% de blastes dans la moelle osseuse, et une numération de neutrophiles >1000/uL et de plaquettes >100 000/uL. Son taux est très bon chez les enfants (90% / 90% chez enfants LAL-B), puis décroît avec l'âge : 45-65% chez les < 65 ans et 30-40% > 65 ans¹⁰⁰.

b. Thérapie de consolidation

Lorsque la rémission est atteinte, la maladie minimale résiduelle (MRD) est évaluée. Elle traduit la présence subsidiaire de cellules malignes ayant résisté au traitement, et se manifeste au travers de translocations et mutations présentes chez celles-ci. L'investigation des transcrits anormaux (PML-RAR α pour les APL par exemple) se fait en routine par PCR quantitative, mais aussi par cytométrie en flux multi-couleurs. Détecter une MRD est associé à un taux de rechute élevé. Une rechute est diagnostiquée lorsque l'on observe à nouveau plus de 5% de blastes dans la moelle osseuse. Bien qu'une seconde rémission soit possible, la rechute représente la première cause de décès par la maladie. Pour la prévenir, une thérapie de consolidation est mise en place, son but est de prolonger la rémission, la probabilité d'une rechute diminuant au fil des années lorsque cet état est maintenu⁷⁴. Pour assurer le maintien de cette phase de rémission, il y a l'option de poursuite du traitement initial. Une formulation orale de l'azacitidine (CC-486) est d'ailleurs désormais accessible pour cette phase de maintenance¹⁰³. Les patients à risque modéré recourent, si un donneur est trouvé, à une transplantation allogénique de CSH. A l'issue de celle-ci, un traitement de maintenance est également mis en place pour limiter le risque de rechute¹⁰⁴.

2. Ciblage de protéines mutées

Les caractéristiques hétérogènes de la maladie ont motivé, ces dernières années, le développement de traitements davantage spécifiques que systémiques.

Les mutations récurrentes mentionnées précédemment sont des cibles privilégiées des nouvelles formes de thérapie. Il a par exemple été montré que les inhibiteurs d'IDH1/2 induisaient la différenciation des cellules malignes chez 10 à 20% des patients. Des inhibiteurs de FLT3 sont également en essais cliniques depuis de nombreuses années, ils consistent globalement en des inhibiteurs se fixant au récepteur pour tenter de diminuer la signalisation en découlant. Des inhibiteurs de KRAS et NRAS sembleraient sensibiliser aux 2 catégories de traitements précédentes¹⁰⁵.

3. Ciblage de la voie apoptotique

La voie apoptotique, à laquelle les cellules leucémiques parviennent à échapper, est également une cible de choix. Les protéines BCL2 et MCL1 sont des protéines anti-apoptotiques fréquemment surexprimées dans les LAM et dont l'inhibition fait l'objet d'études cliniques¹⁰⁵.

4. Thérapies immunitaires

a. Anticorps monoclonaux

Des anticorps monoclonaux, dirigés contre les antigènes tumoraux désirés, sont produits *in vitro*, puis utilisés comme vecteurs. Couplés à des composés chimiques ou des toxines, ou bien radiomarqués (par des radionucléides, particules α ou β), ils permettent la délivrance du composé aux cellules leucémiques ou leur irradiation¹⁰⁶.

Des anticorps bispécifiques sont en voie de développement : ils allient le rôle de vecteur et de modulateur de l'immunité. Ils possèdent deux régions variables : l'une fixant la sous-unité CD3- ϵ d'un récepteur de cellules T, et l'autre dirigée vers un antigène à la surface des cellules leucémiques. Ils rapprochent les lymphocytes T des cellules cancéreuses pour faciliter leur élimination, via l'établissement entre eux d'une synapse cytolytique conduisant à la lyse. Plusieurs variantes, parmi lesquelles DARTs, BiKEs et TriKEs, sont en essais cliniques¹⁰⁷.

b. Réactivation du système immunitaire

La résistance des cellules leucémiques aux thérapies conventionnelles pourrait être favorisée par leur échappement au système immunitaire. Elles parviennent notamment à cet échappement en surexprimant des protéines inhibitrices de l'activation des lymphocytes (CTLA-4 et PD1 notamment), constituant des « points de contrôles immunitaires ». Sur le long terme, le système immunitaire finit par s'épuiser face à la persistance d'antigènes tumoraux. L'immunothérapie s'illustre par la stimulation/réactivation du système immunitaire. A ses prémices, elle consistait simplement en l'apport d'interleukine-2 (IL-2) et d'interféron- γ (IFN- γ), stimulant respectivement les lymphocytes infiltrant les tumeurs et les cellules dendritiques¹⁰⁸.

Afin d'éviter d'atteindre l'état d'épuisement immunitaire, il est à l'heure actuelle possible d'agir sur les points de contrôle immunitaires. Les traitements les plus utilisés sont ceux associant à la chimiothérapie l'administration d'anticorps ciblant les protéines CTLA-4 ou PD-1, pour maintenir l'activation des lymphocytes T. CD47 est également une cible étudiée, puisque sa surexpression permet aux cellules leucémiques d'échapper à la phagocytose par les macrophages¹⁰⁵.

La vaccination thérapeutique a pour but d'induire une réponse humorale et cellulaire, dirigée contre les antigènes associés aux tumeurs (TAA), en activant des lymphocytes T spécifiques. Les TAA ciblés doivent (i) être fortement exprimés sur les cellules/blastes leucémiques en comparaison avec les autres tissus, (ii) jouer un rôle pivot dans la survie de la cellule leucémique, (iii) induire une forte réponse CD4 et CD8. Les 2 catégories principales de vaccins sont basées sur l'injection de peptides (WT1 / PR-1 et PR3 / PRAME / RHAMM) ou sur l'infusion de cellules dendritiques pré-activées in vitro par des antigènes tumoraux¹⁰⁶.

Les thérapies cellulaires se basent sur l'injection de cellules T et NK adoptives, mais aussi sur l'exposition à des cytokines augmentant la cytotoxicité des cellules NK (IL-2, -12, -15 et -18). L'administration d'agents hypométhylants augmentent également la réponse immunitaire¹⁰⁹. Aussi, les cellules T peuvent être modifiées génétiquement pour exprimer des récepteurs (CAR : *Chimeric Antigen Receptor*) les redirigeant contre les cellules leucémiques, et ne nécessitant pas les prérequis d'activation habituels. Enfin, des virus oncolytiques sont utilisés pour lyser les cellules leucémiques et induire la production de DAMPs (damage-associated molecular patterns). **(Figure 13)**

5. Radiothérapie externe

Un traitement par rayon X peut être mis en place, afin d'éliminer les cellules qui se sont disséminées en dehors de la moelle osseuse et du sang circulant. Certaines parviennent en effet à se loger dans le cerveau, le liquide spinal, les testicules... Sa visée peut aussi être palliative, afin de réduire les douleurs osseuses ou soulager l'hyperplasie hépatique et splénique.

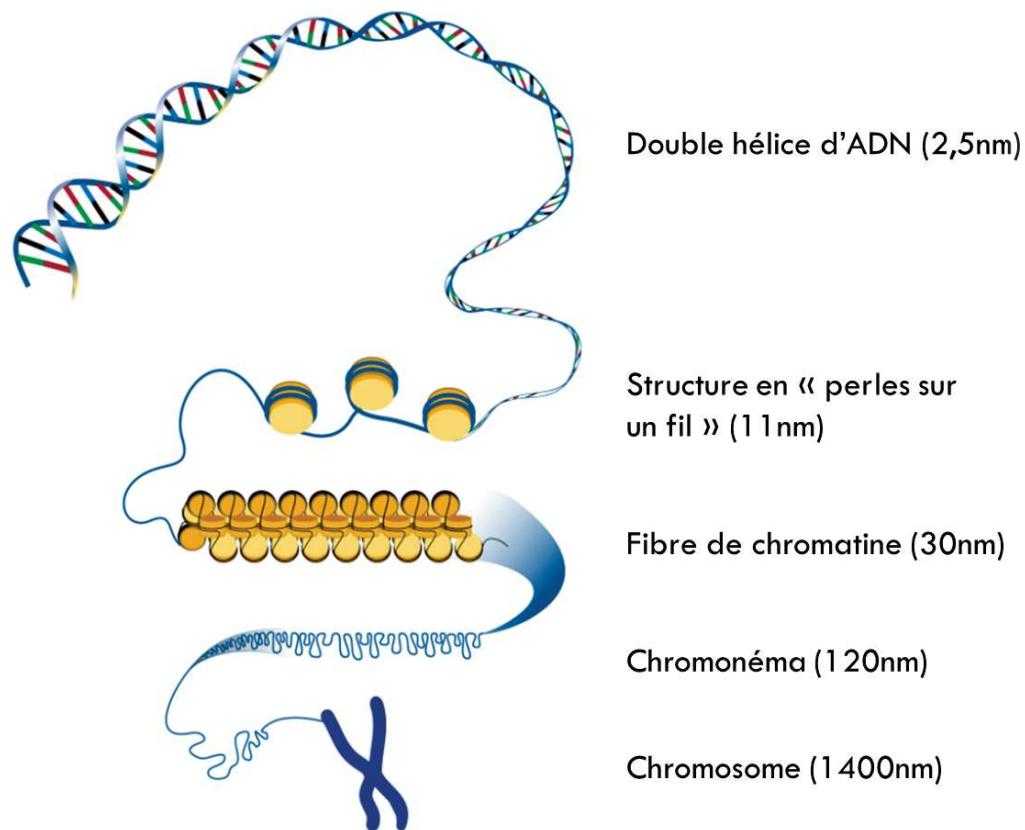


Figure 14. Organisation de la chromatine.

L'ADN est une double hélice de brins antiparallèles formés de bases puriques (A/G) et pyrimidiques (T/C), attachées à un squelette phosphate-sucre. L'ADN enroulé sur les nucléosomes forme une structure dite en « perles sur un fil », qui s'organise en une fibre de 30nm, puis en chromonéma de 120nm, en chromatine de 300-700nm, avant de former le chromosome. Figure adaptée de Studywrap.com.

Chapitre 3

Notions d'épigénétique

*La chromatine est un complexe associant acides désoxyribonucléiques (ADN), protéines et acides ribonucléiques (ARN). Elle se trouve dans un état plus ou moins compacté, lui permettant de faire tenir ses 2 mètres approximatifs de longueur dans un noyau cellulaire d'en moyenne 6µM. L'unité de base de la compaction de la chromatine est le nucléosome, constitué d'un enroulement de 147 paires de bases d'ADN autour d'un octamère d'histones. Les histones sont des protéines très conservées chez les eucaryotes. Elles se composent de 3 hélices α connectées par 2 courts repliements. Les protéines chaperonnes empêchent l'agrégation des histones solubles, les assemblent pour former le nucléosome, et participent au dynamisme de celui-ci¹¹⁰. Les histones comportent beaucoup d'arginines et lysines, ce qui les charge fortement positivement et leur garantit une interaction stable avec l'ADN, chargé négativement. Tous les 6 nucléosomes (environ 1kb d'ADN), la structure effectue un tour complet de rotation autour d'un axe, ce qui crée une structure en solénoïde de 30nm de diamètre. Cette fibre s'organise en chromonéma puis en chromatine^{111,112} (**Figure 14**).*

La compaction de la chromatine n'a pas qu'un rôle structural, il est aussi transcriptionnel : le nucléosome impose en effet une barrière physique aux facteurs de transcription (FTs) et enzymes. La chromatine se doit de rester flexible et dynamique pour répondre aux besoins transcriptionnels en temps réel. La régulation de la transcription se fait ainsi par une action synergique de FTs, de la machinerie transcriptionnelle, de remodeleurs de la chromatine et de modifications d'histones.

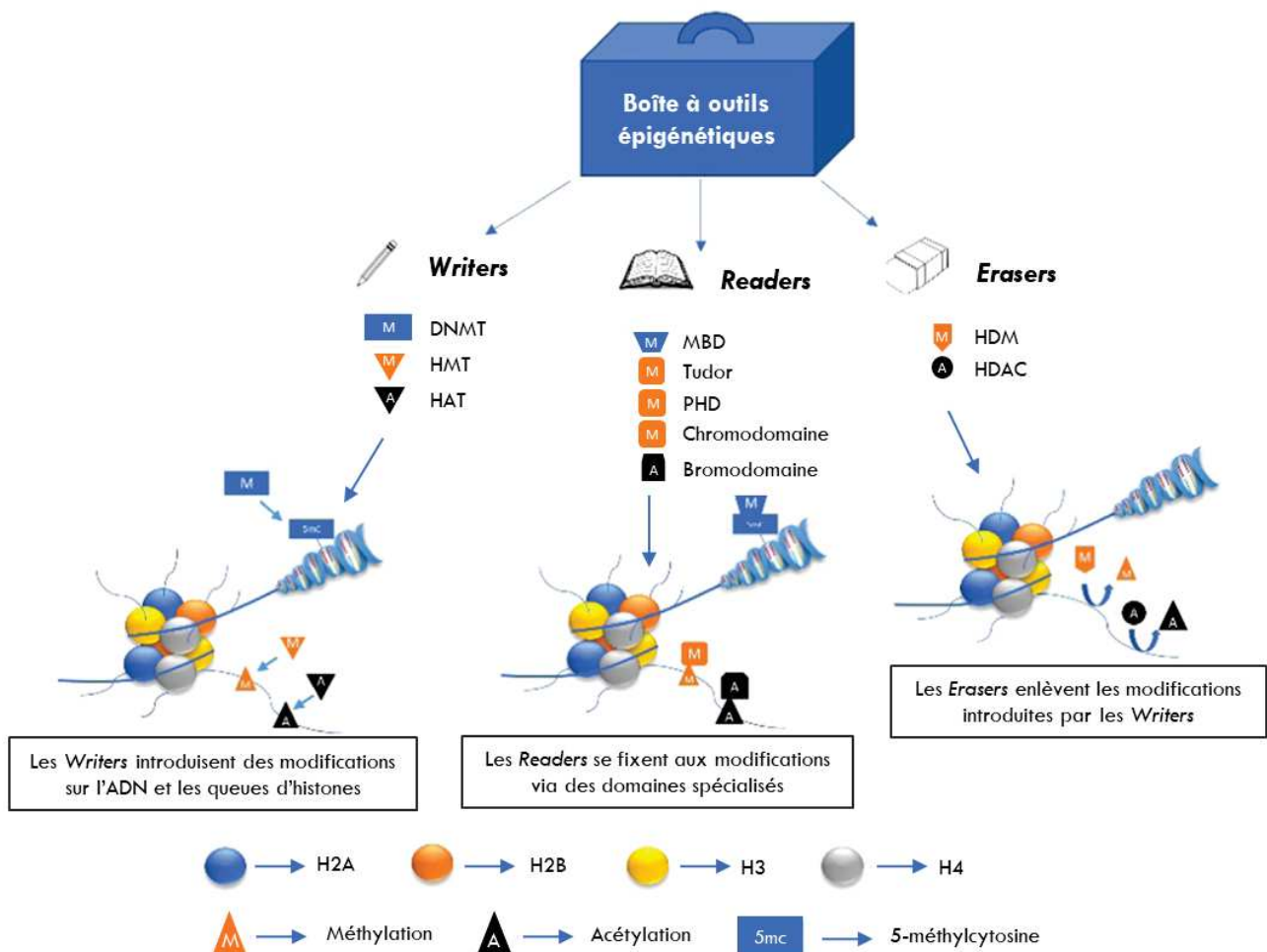


Figure 15. Writers, Readers et Erasers.

DNMT : DNA-Methyltransferase; HMT : Histone-Lysine-Methyltransferase; HAT : Histone Acetyltransferase; MBD : Methyl-Binding Domain; PHD : Plant Homeodomain; HDM : Histone Demethylase; HDAC : Histone Deacetylase. Figure adaptée de Biswas et. al, 2018.

I. Régulation épigénétique de la transcription

La transcription est le processus de formation d'ARN messager (ARNm) à partir d'ADN. L'ARNm exporté dans le cytoplasme est ensuite traduit en protéines, qui exercent diverses fonctions biologiques. Le contrôle de la transcription est le premier niveau de régulation d'expression d'un gène. La complexité d'un organisme est d'ailleurs davantage reflétée par la variété de ses mécanismes de régulation de l'expression génique, que par son nombre de gènes. La transcription implique la participation d'ARN polymérases, de FTs à l'action activatrice ou inhibitrice, et d'enzymes de modification des nucléosomes. Pour qu'un promoteur soit accessible aux FTs et à la machinerie transcriptionnelle, la chromatine doit être décompactée.

Alors que la génétique étudie l'hérédité des séquences d'ADN et de leurs altérations, l'épigénétique s'intéresse aux changements n'impliquant pas de modification de la séquence génique elle-même, mais sa disponibilité pour la machinerie de transcription. L'épigénome comprend l'ensemble des modifications covalentes réversibles au niveau des nucléotides ou des histones, susceptibles d'entraîner une reprogrammation de la cellule sans altérer la séquence génomique. Ces modifications médient une diversité dans les profils d'expression, et sont mitotiquement et méiotiquement héréditaires. Trois mécanismes épigénétiques dynamiques et régulés sont répertoriés : la méthylation de l'ADN, les modifications post-transcriptionnelles des histones et la régulation par des ARNs non codants. Ci-dessous seront désignés de « *writers* » les acteurs introduisant les modifications épigénétiques, de « *readers* » ceux capables de les identifier et les interpréter, et de « *erasers* » ceux les supprimant (**Figure 15**).

A. Mécanismes épigénétiques

1. Méthylation de l'ADN

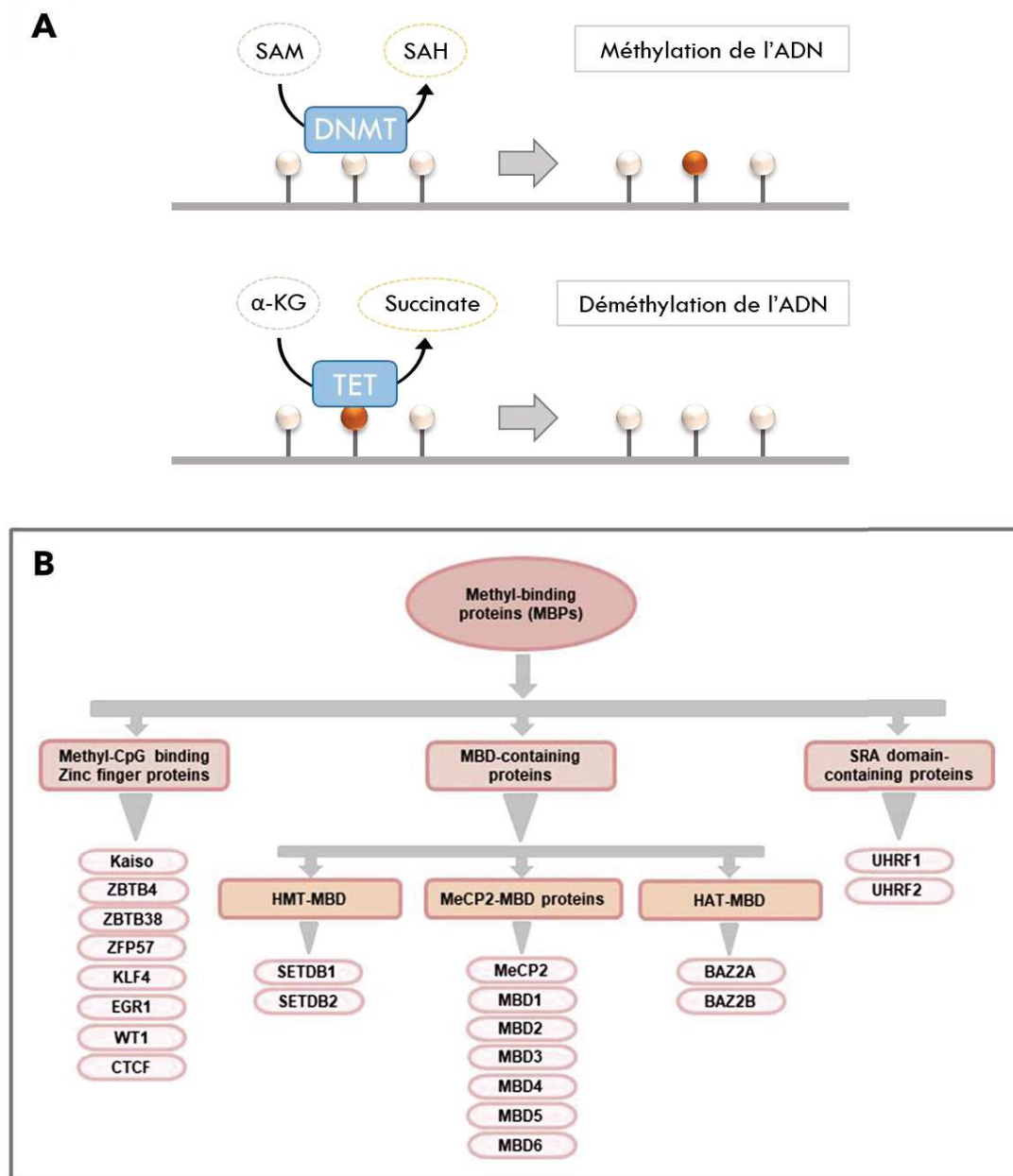


Figure 16. La méthylation de l'ADN.

(A) La méthylation de l'ADN est catalysée par des méthyltransférases de l'ADN (DNMT), à partir de S-adenosyle-méthionine (SAM), qui sert de source de méthyle et est transformé en S-adenosyle homocystéine (SAH). La déméthylation active de l'ADN est catalysée par la famille de protéines dioxygénases TET (Ten-eleven Translocation). Les protéines TET oxydent le 5mc en 5-hydroxy-méthylcytosine (5hmc), à partir d'a-kétoglutarate (α -KG), et produisent du succinate.

(B) Les protéines dotées d'une capacité de fixation aux CpG méthylés (*readers*) sont classifiées en 3 familles, selon le domaine fonctionnel qu'elles utilisent pour cette interaction. Figure extraite de Mahmood et. al, 2019.

La méthylation de l'ADN consiste en l'ajout d'un groupement méthyl sur le 5^e carbone d'une cytosine (C). Le processus est catalysé par deux types de *DNA-methyltransferases* (DNMTs, les *writers*). Les méthylations de novo, sur de l'ADN nu, sont introduites par DNMT3A et DNMT3B, avec l'assistance de DNMT3L^{113,114}. La maintenance des méthylations préexistantes est assurée par DNMT1 qui, lors de la réplication et de la réparation de l'ADN, reproduit le motif de méthylation du brin parental sur le brin néosynthétisé¹¹⁵. La déméthylation de l'ADN peut être passive, à l'issue de la réplication en cas d'absence ou dysfonction de DNMT1, ou active via différents processus enzymatiques impliquant les protéines TET¹¹⁶⁻¹¹⁸ (**Figure 16A**).

Les cytosines méthylées sont classiquement celles des dinucléotides 5'-C-phosphate-G-3' (CpG), que l'on retrouve abondamment dans le génome. 60 à 80% d'entre eux sont méthylés¹¹⁹. On les retrouve au niveau de séquences répétées, et des régions intra- et intergéniques. Là, leur méthylation restreint l'accessibilité de la chromatine, et permet la répression de rétrotransposons, la prévention de transcription cryptique, et régule la fixation de FTs. Elle réprime donc, en d'autres termes, l'expression d'éléments potentiellement délétères^{120,121}. 70% des promoteurs contiennent des régions abondantes en dinucléotides CpG, appelées « îlots CpG ». Elles sont typiquement dépourvues de méthylation, et situées à proximité du site d'initiation de la transcription (TSS). Ces domaines conservés au cours de l'évolution semblent favoriser l'accessibilité de la chromatine et la fixation de FTs. La méthylation des îlots CpG réduit l'interaction de l'ADN avec les FTs et la machinerie transcriptionnelle, ou bien recrute des protéines capables de s'y fixer. Les *readers* de la méthylation de l'ADN, appelés *Methyl-binding proteins* (MBP), se répartissent en 3 familles, en fonction du domaine fonctionnel de liaison à l'ADN méthylé qu'ils possèdent (**Figure 16B**). Ces protéines vont à leur tour recruter des complexes de remodelage de la chromatine, pour restreindre son accessibilité localement¹²²⁻¹²⁴.

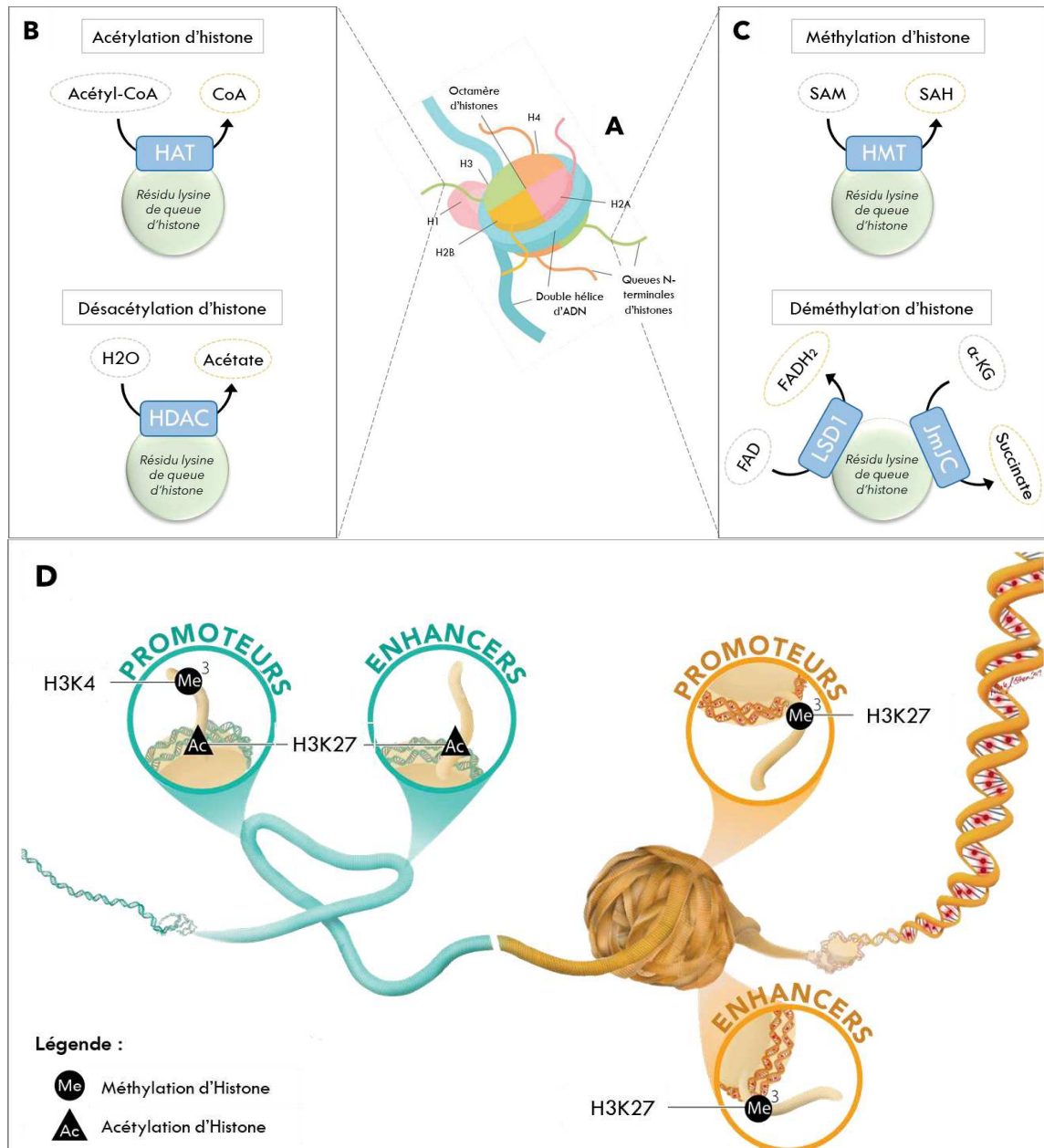


Figure 17. Modifications post-traductionnelles des histones.

(A) Organisation d'un nucléosome. (B) L'acétylation des histones est catalysée par des histone-acétyltransférases (HAT), à partir d'acétyl-CoA, et produit du CoA. Leur désacétylation fait classiquement intervenir des histone-désacétylases (HDAC) de classe I et II. (C) La méthylation des histones est catalysée par des histone-méthyltransférases (HMT). Leur déméthylation peut être médiée par la déméthylase-1 spécifique des lysines (LSD1), à partir de flavine-adénine-dinucléotide (FAD), ou par la déméthylase Jumonji C domain (JmJC) à partir d'α-kétoglutarate (α-KG) et produisant du succinate. (D) Localisation préférentielle des modifications post-traductionnelles des histones les plus répandues : H3K27ac, H3K4me₃, H3K27me₃ (Heintzman 2007). Figures adaptées de Birch et. al, 2020, Héloïse Newman (Behance).

2. Modifications post-traductionnelles des histones

L'extrémité N-terminale des histones est non structurée et dépasse du nucléosome ; elle est sujette à des modifications post-traductionnelles (**Figure 17A**). Ces modifications ouvrent la chromatine et/ou recrutent des effecteurs : elles servent de plateforme pour les *readers*, et/ou modulent la force des interactions entre l'ADN et l'octamère d'histones en altérant leur charge. Ces modifications sont souvent réversibles, mais capables d'être maintenues au cours des divisions cellulaires. En fonction des types de modifications et de leur combinaison, elles activent ou inhibent la transcription. Les plus fréquentes sont les acétylations et méthylations, elles s'appliquent principalement au niveau de résidus lysine (K) et arginine (R)¹²⁵.

a. Acétylation des histones

L'acétylation des histones consiste en l'ajout d'un groupe acétyle (Ac) par une *histone acetyltransferase* (HAT) à partir d'acétylcoenzyme A, le plus souvent au niveau d'une lysine. Elle neutralise les charges positives des lysines et donc de l'histone, ce qui réduit son affinité avec l'ADN et décompacte la chromatine : ces modifications sont associées à une activation de la transcription. L'histone H3 est par exemple hyperacétylée dans les régions fortement transcrites du génome, et est à l'inverse hypoacétylée dans des régions silencieuses comme les télomères. Le processus est inversé par des *histones-deacetylases* (HDACs). L'acétylation est reconnue par des protéines dotées d'un bromodomaine extra-terminal (BET) ou de *tandem PHD finger*^{126,127} (**Figure 17B**).

b. Méthylation des histones

La méthylation consiste en l'addition d'un ou plusieurs groupe(s) méthyle(s). Elle est catalysée par une *histone methyltransferase* (HMT) qui utilise de la S-adénosyl méthionine (SAM) comme substrat pour transférer le groupe méthyle sur les résidus lysine des queues d'histones. Le processus s'effectue le plus souvent sur des lysines de l'histone H3 ou H4, et consiste en des mono-, di- ou triméthylations. Selon le résidu d'acide aminé modifié, l'action est activatrice ou répressive sur la transcription. Une abondance de modifications H3K27me3 dans une région codante ou au niveau d'enhancers est par exemple corrélée à une répression de la transcription. Ces marques très répandues sont mises en place spécifiquement par la méthyltransférase PRC2, et s'organisent en domaines étendus¹²⁸. A l'inverse, H3K4me3 est

enrichie de façon canonique au niveau du promoteur et d'enhancers de gènes activement transcrits. Tous les promoteurs dotés d'îlots CpG présentent ces marques, leur abondance étant corrélée à l'activité du gène¹²⁹. L'apposition de cette marque peut se faire par 6 méthyltransférases distinctes (SETD1A, SETD1B, MLL1, MLL2, MLL3, MLL4). De manière plus générale, les méthylations des H3K4, H3K36 et H3K79 sont considérées comme activatrices de l'expression des gènes, celles touchant H3K27, H3K9 et H4K20 sont à l'inverse connues pour inhiber l'expression des gènes. Ces modifications ne sont pas exclusives : certains domaines sont bi ou trivalents, et des marques peuvent cohabiter au sein d'un même loci. La méthylation des histones est reconnue par une longue liste de domaines, qui ne cesse de croître, parmi lesquels *chromo-*, *tudor-*, *double/tandem tudor-* ou encore *PHD domains*¹²⁶. Celle-ci est effacée par des déméthylases¹²⁵ (**Figure 17C&D**).

3. Interactions entre ces mécanismes

a. Co-occurrences et exclusions

Il existe une certaine interdépendance entre les modifications épigénétiques : la présence de certaines peut stimuler la mise en place, le renforcement ou la maintenance de motifs, ou au contraire l'inhiber. Le phénomène de rétroaction positive s'illustre par exemple par la forte corrélation observée entre la méthylation des histones et celle de l'ADN : en présence de méthylation sur H3K9, une méthylation de l'ADN se met en place pour renforcer la compaction de la chromatine et limiter la transcription, cela au niveau d'éléments répétés et de rétrotransposons¹³⁰. Autre exemple, la présence de marques H3K4me3 semblerait favoriser le recrutement de HATs, expliquant sa cohabitation avec les marques H3K27ac^{131,132}.

A l'inverse, la présence de méthylations sur H3K4, au niveau de promoteurs actifs, endigue l'apposition de méthylation de novo sur les CpG¹³³. Ces mêmes marques empêchent également l'apposition de triméthylation sur H3K27, en inhibant l'activité du *writer* de cette modification (PRC2)¹³⁴. Cette observation se corrèle à l'exclusion mutuelle entre marques H3K27me3 d'une part et H3K4me3 et H3K27ac qui cohabitent d'autre part. Aussi, bien que toutes deux répressives, la méthylation de l'ADN et de H3K27 s'excluent mutuellement au niveau des îlots CpG¹³⁵.

b. Notion de bivalence

Malgré ces phénomènes de co-occurrences et exclusions mutuelles, il n'est pas rare que des marques « contradictoires » occupent transitoirement un même promoteur. C'est le cas des gènes développementaux dotés d'un promoteur bivalent, prêts à être rapidement activés ou réprimés, pour assurer une réponse transcriptionnelle efficace¹³⁶⁻¹³⁸. Lors du développement de l'épithélium intestinal, on observe par exemple la cohabitation des marques H3K4me3 et H3K27me3 au niveau du promoteur du gène *Cdx2*, lui permettant cette bivalence. L'activation d'AMPK favorise la prédominance des marques activatrices H3K4me3, et donc l'expression de *Cdx2*¹³⁹.

B. Dynamisme de la chromatine

1. Remodelage des nucléosomes

La densité et la position des nucléosomes influencent l'accès à la chromatine des FTs et de la machinerie transcriptionnelle^{140,141}. Les promoteurs forts ont ainsi tendance à se localiser au niveau de régions où la densité des nucléosomes est faible, et leur turnover important. La plupart des FTs ne se fixent qu'à de l'ADN nu, à l'exception des FTs dits « pionniers », classiquement des régulateurs développementaux, dont la seule fixation au niveau d'ADN nucléosomal compacté suffit à initier des événements transcriptionnels¹⁴². Pour accéder à l'ADN nucléosomal, les FTs conventionnels ont besoin d'un remodelage préalable.

Les remodeleurs de la chromatine, répartis en 4 sous-familles, catalysent des réactions de restructuration des nucléosomes. Ils consistent en des translocases d'ADN qui interrompent le contact de l'ADN avec les histones, sous l'impulsion d'une hydrolyse d'ATP (activité ATPase). Le remodelage consiste en des coulissements, espacements, échanges ou évictions d'histones¹⁴³.

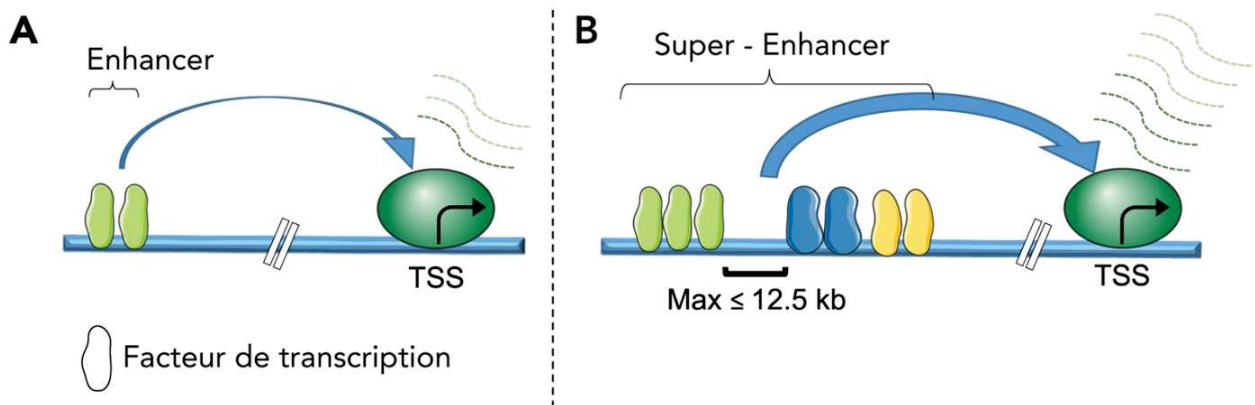


Figure 18. Concept d'enhancer et de super-enhancer.

(A) Les enhancers sont des éléments régulateurs situés à distance des sites d'initiation de la transcription (TSS) des gènes qu'ils régulent. (B) Les super-enhancers sont des regroupements d'enhancers espacés d'au maximum 12,5kb. Figure de Jia et. al, 2019.

2. Notion d'enhancer

Les enhancers sont décrits comme des éléments qui augmentent la transcription d'un gène donné, indistinctement de leur orientation respective et de leur distance génomique. Ils mesurent 200 à 1500pb, et se rapprochent de leur cible grâce à un mécanisme de repliement de la chromatine décrit ultérieurement. Enhancers et promoteurs partagent certains motifs au niveau de leur séquence, un environnement chromatinien et une activité similaire rendant leur distinction complexe. Comme les promoteurs, les enhancers actifs sont ainsi fixés de façon coordonnée et dépendante du contexte cellulaire par des FTs, qui recrutent la machinerie nécessaire au remodelage de la chromatine (modificateurs d'histones, ...) et à l'amorce de la transcription (protéines du complexe d'initiation de la transcription, ...)^{144,145}. Les super-enhancers sont des clusters d'enhancers, ils diffèrent en taille, densité de FTs, et capacité à activer la transcription^{145,146} (**Figure 18**).

Les enhancers sont particulièrement requis en contexte de régulation complexe, de par leur capacité à contrôler la transcription de manière spatio-temporelle et spécifique par rapport à un stimulus, à un type cellulaire ou à un état développemental. Les interactions enhancer-promoteur sont hautement spécifiques des types cellulaires. Elles sont très dynamiques et corrélées aux profils d'expression génique spécifiques des types cellulaires^{147,148}.

Les enhancers peuvent être transcrits en ARN non codants (eRNA). Ces derniers sont relativement courts (50-2000 nucléotides) et transcrits bi-directionnellement. On leur attribue un rôle dans la stabilisation de la fixation des FTs, et dans les repliements de la chromatine^{149,150}.

3. Organisation tridimensionnelle et topologique

a. Que sont les domaines d'associations topologiques (TADs) ?

A plus grande échelle, celle du chromosome, la chromatine s'organise en domaines d'« interactions chromatinienne préférentielles », appelés domaines d'associations topologiques (TADs). Ces domaines sont stables à travers les divisions, relativement invariants entre les types cellulaires, et conservés évolutionnellement. Ils mesurent en moyenne 185kb, mais cette longueur oscille en réalité entre 40kb et des millions de bases de longueur. Ce sont des unités structurales et fonctionnelles, enrichies en gènes actifs¹⁵¹.

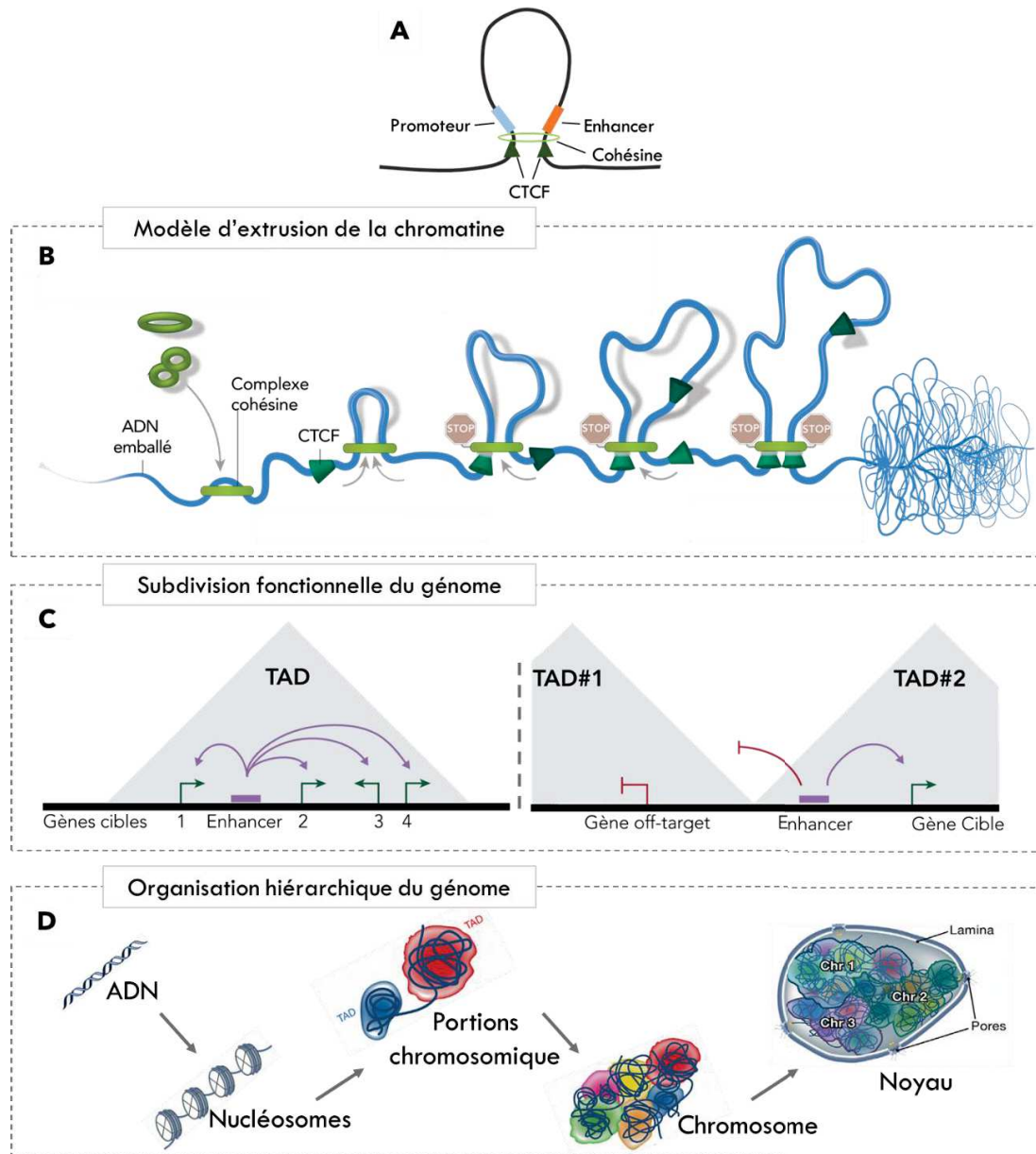


Figure 19. Organisation tridimensionnelle de la chromatine.

(A) Schéma d'une interaction enhancer-promoteur par repliement de la chromatine. (B) L'extrusion de la chromatine fait intervenir la cohésine et les protéines CTCF comme effecteurs de la formation des boucles d'interaction. Selon ce modèle, une boucle d'ADN franchit l'anneau formé par la cohésine, puis se déverse à travers jusqu'à ce que des protéines CTCF convergentes fixées sur la chromatine atteignent cet anneau. Adapté de Dolgin et. al, 2017. (C) Au sein d'un même TAD, un enhancer peut réguler l'expression de plusieurs gènes de manière coordonnée. Un enhancer ne peut cependant influencer la régulation que des promoteurs appartenant à la même unité génomique. (D) Les nucléosomes constituent le premier degré d'organisation de la chromatine, suivi de la ségrégation entre TADs, puis entre régions intrachromosomiques, et enfin la chromatine est ségrégée entre compartiments actifs et inactifs à l'échelle du noyau. Figure adaptée de Sexon et. al, 2015.

b. Structure des TADs

Les extrémités des TADs, qui sont plus ou moins nettes, se caractérisent par la présence à minima de complexes cohésines et -moins systématiquement- de protéines insulatrices *CCCTC-binding factor* (CTCF) (**Figure 19A**). Le complexe cohésine est constitué de 4 sous-unités : RAD21, SMC1, SMC3 et STAG1 ou STAG2. Il est très conservé et est capable de lier 2 segments d'ADN via sa structure annulaire. Il est notamment connu pour maintenir les chromatides sœurs après la réplication de l'ADN, en attendant leur ségrégation entre les cellules filles¹⁵². CTCF est la principale protéine insulatrice décrite chez les vertébrés. C'est une protéine contenant un motif d'interaction avec l'ADN à 11 doigts de zinc¹⁵³. Elle est classiquement impliquée dans la liaison des chromatides sœurs lors de la mitose, de la méiose, et de la réparation par recombinaison homologue. Il a été observé que plus il y avait de fixation de CTCF, plus stable était le TAD. CTCF fait office d'isolateur entre 2 TADs, et empêche leur fusion. La déplétion de CTCF entraîne ainsi des interactions avec les TADs voisins¹⁵⁴. Les deux frontières d'un TAD présentent des motifs de fixation de CTCF (5'-CCACNAGGTGGCAG-3') chacun dans 2 orientations possibles, mais très souvent convergentes¹⁵⁵. Un TAD contient une boucle de chromatine, formée par une extrusion de la chromatine à travers le complexe cohésine. La boucle s'étirerait jusqu'à ce que la cohésine rencontre les CTCF fixés au niveau de leurs sites^{156,157} (**Figure 19B**).

c. Fonctions des TADs

Tout d'abord considérée comme une simple unité de repliement du chromosome, cette partition de la chromatine génère en fait des subdivisions fonctionnelles du génome. Au niveau de ces domaines, les régions génomiques présentent une fréquence d'interaction entre elles plus importante que la moyenne. Une de leurs fonctions principales est de contraindre les interactions enhanceur-promoteur à l'intérieur d'un même TAD, et de bloquer celles avec des éléments extérieurs voisins. Au-delà de la chromatine, les TADs structurent ainsi également les interactions entre gènes et régulateurs. Aussi, enhanceurs et promoteurs à proximité partagent les FTs, permettant une régulation coordonnée des gènes¹⁵⁸. Les TADs divisent ainsi le génome en unités génomiques insulaires, régulées de manière « autonome »¹⁵⁹ (**Figure 19C**).

Ce n'est pas nouveau de conférer une relation spatiale aux gènes. De manière générale, la répartition de la chromatine dans le noyau est déjà corrélée avec son statut transcriptionnel : les régions actives sont plus souvent vers l'intérieur et les régions réprimées vers la périphérie¹⁶⁰. Et à plus grande échelle, les chromosomes se répartissent également dans des territoires distincts¹⁶¹ (**Figure 19D**).

II. Épigénétique et leucémie aiguë myéloïde

La dérégulation épigénétique, plus fréquente chez l'adulte et corrélée avec l'âge, est maintenant reconnue comme un mécanisme contribuant activement à la pathogenèse des hémopathies. Les gènes régulant les mécanismes épigénétiques sont parmi les plus fréquemment mutés dans ces pathologies. Plus de 70% des patients porteurs d'AML présentent des mutations dans des gènes liés à la méthylation de l'ADN ou l'acétylation des histones¹⁶².

A. Hématopoïèse physiologique

Le paysage chromatinien joue un rôle décisif dès le début du processus hématopoïétique¹⁶³. Une réorganisation dynamique de la chromatine prédispose à l'engagement envers un lignage donné, en facilitant les programmes transcriptionnels associés. Les CSH exprimant le gène SATB1 (AT-rich sequence-binding protein 1), acteur des complexes de remodelage de la chromatine, s'engagent par exemple préférentiellement vers le lignage lymphoïde. DNMT1 semble également favoriser cette orientation, puisque sa perte d'expression dans les CSH prédispose à un engagement érythroïde et myéloïde, et s'accompagne d'une diminution des populations lymphoïdes. La perte d'expression de DNMT3A déclenche quant à elle des programmes de différenciation érythroïde, en augmentant l'accessibilité de sites de fixation de facteurs de différenciation érythroïde clés (Klf1, Gata1, Tal1)¹⁶⁴.

Deux modèles s'opposent quant à l'implication de la régulation épigénétique. Le modèle « permissif » suppose que les CSH possèdent une permissivité / facilité d'engagement envers un type cellulaire. Il s'appuie sur l'observation que les sites de fixation de FTs spécifiques d'un

lignage sont d'emblée davantage accessibles dans les CSH qui évolueront vers ce lignage. Le modèle « instructif » propose à l'inverse une activation progressive des gènes spécifiques des différents lignages. Cette hypothèse est supportée par le fait que certaines régions génomiques gagnent en accessibilité au cours de la différenciation. On considère à l'heure actuelle que ces 2 mécanismes influencent conjointement le devenir des CSH¹⁶⁵.

Autre phénomène supportant cette théorie : Alors que les progéniteurs CD34+ présentent une chromatine dépourvue de marques répressives H3K27me3, on observe pendant la spécification des lignages une accumulation massive de celles-ci. Ces marques répressives vont néanmoins être perdues ponctuellement au contact de cytokines, et l'accessibilité nouvellement acquise de la chromatine y permettrait le recrutement de facteurs spécifiques de lignées^{35,166}.

B. Mécanismes perturbés

Les lésions de l'épigénome sont variées : motifs anormaux de méthylation de l'ADN, perturbation des modifications d'histones, altération de l'organisation de la chromatine... Un nombre croissant d'entre elles sont associées aux LAM, et portent atteinte à des facteurs essentiels à l'hématopoïèse. Les mutations dans des gènes comme DNMT3A, TET1, TET2, IDH2, et ASXL1, mentionnées dans le chapitre précédent, sont des événements précoces dans la leucémogénèse¹⁶⁷.

1. Méthylation de l'ADN

Les cancers dans leur ensemble sont associés à une hypométhylation générale : la déméthylation contribue à l'instabilité génomique, puisque l'exposition de séquences répétées prédispose à la recombinaison, et augmente les aneuploïdies. Elle s'accompagne d'une activation de la transcription notamment au niveau d'éléments répétés ou transposables, et d'oncogènes¹⁶⁸. Cette perte de méthylation peut s'expliquer par une perte de fonction de DNMTs. L'exemple typique est celui de DNMT3A, principale responsable de la méthylation dans le compartiment hématopoïétique, qui se retrouve muté dans 30% des LAM. La mutation a classiquement lieu dans son domaine catalytique, au niveau du résidu R882. La protéine

mutante agit comme un dominant négatif, qui antagonise la tétramérisation de l'enzyme et réduit son activité méthyltransférase de près de 80%¹⁶⁹. Il s'agirait d'un événement précoce de la leucémogénèse, privilégiant l'autorenouvellement des CSH et donc leur expansion¹⁷⁰. On constate aussi une hyperméthylation locale des îlots CpG au niveau du promoteur de gènes suppresseurs de tumeurs^{171,172}. L'hyperméthylation est courante dans les LAM, en comparaison avec les cellules hématopoïétiques saines, mais pas toutes ne contribuent à la leucémogénèse. La perte de fonction de TET2, par décalage du cadre de lecture ou mutations non-sens, dans 9-23% des LAM en est un exemple. Elle perturbe la différenciation et augmente l'autorenouvellement des CSH et progéniteurs, mais ne suffit pas à elle seule pour induire une transformation leucémique¹⁷³. On observe également des mutations mutuellement exclusives affectant IDH1 ou IDH2 dans 17% des LAM. Puisque ces enzymes inhibent l'activité de TET2, leur régulation à la hausse résulte également en une hyperméthylation, qui bloque la différenciation^{172,174}.

2. Modifications d'histones

La méthylation des histones peut être perturbée par des mutations dans les writers. On retrouve parmi les plus notables, des mutations touchant les membres des complexes PcG, et plus particulièrement EZH2¹⁷⁵. Sa perte de fonction, de mauvais pronostic, diminue l'apposition de marques H3K27me3 et conduit notamment à une régulation à la hausse des gènes Hox¹⁷⁶. Un exemple emblématique est également le gène MLL, fréquemment muté par des duplications en tandem ou des réarrangements chromosomiques. De ses nombreuses fusions - on lui dénombre près de 70 partenaires- résulte une perte de son domaine SET à activité méthyltransférase. Ses variants chimériques soutiennent la leucémogénèse par la dérégulation transcriptionnelle de ses cibles, parmi lesquelles MEIS1 et HOXA9, dont il est en charge du maintien d'expression^{177,178}.

Des altérations génétiques peuvent aussi conduire au recrutement anormal de modificateurs d'histones. La fusion AML1-ETO induit par exemple le recrutement d'un complexe contenant une HDAC au lieu d'une HAT au niveau du promoteur de gènes comme IL3, GM-CSF ou MPO, réprimant leur transcription, et la différenciation myéloïde¹⁷⁹.

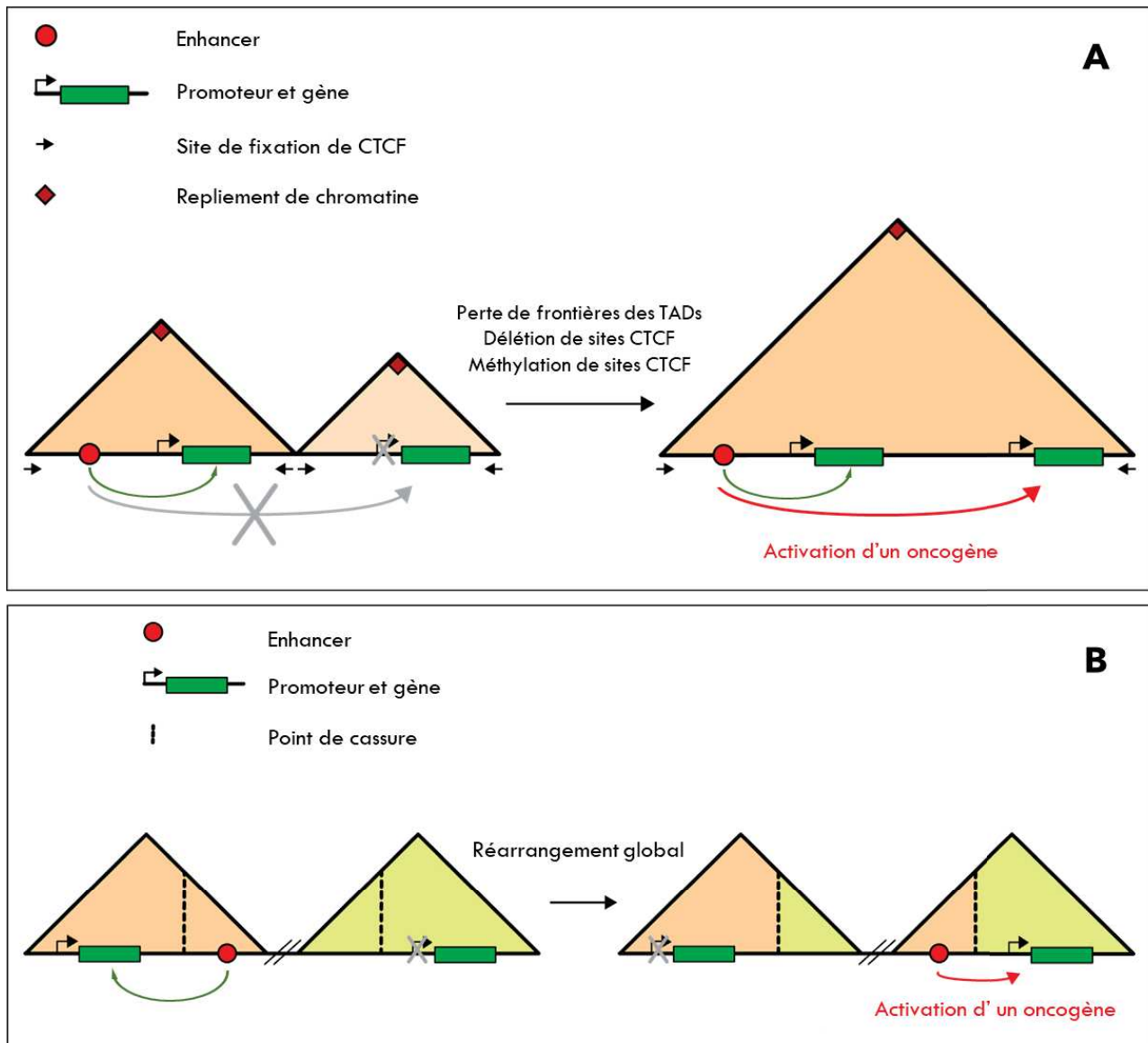


Figure 20. Perturbation des TADs

(A) Activation oncogénique par perte de frontières entre 2 TADs. Un nouveau TAD se crée et permet au enhancer de l'un d'interagir de manière aberrante avec un gène de l'autre. (B) Activation oncogénique par réarrangement de TADs. L'intervention de 2 fractions de TADs crée 2 TADs chimériques. Le gène du premier n'est plus exprimé, tandis que l'oncogène du second est maintenant exprimé par l'activation du enhancer relocalisé. Figures adaptées de Valton et. al, 2016.

3. Organisation de la chromatine

La perte de remodeleurs des nucléosomes a un impact inévitable sur l'hématopoïèse, puisque l'accessibilité de la chromatine définit les programmes transcriptionnels des progéniteurs hématopoïétiques¹⁸⁰.

Concernant les enhancers, l'augmentation de leur nombre de copies est un mécanisme classique d'activation oncogénique, impliqué par exemple lors de la surexpression du gène MYC dans 3% des LAM¹⁸¹. Des mutations somatiques ponctuelles peuvent également y perturber la fixation des FTs, dans les LLC l'expression du gène PAX5 est par exemple diminuée à la suite de mutations dans l'un de ses enhancers¹⁸². Aussi, ils peuvent au même titre que les promoteurs être touchés par le dépôt erroné de modifications épigénétiques, qui modifient leur accessibilité¹⁸³. La méthylation des enhancers peut d'ailleurs servir de biomarqueur de réponse au traitement, et d'outils diagnostic¹⁸⁴.

La structure et les constituants des TADs sont aussi sujets à des anomalies en contexte de LAM. L'interruption de TADs peut causer l'activation d'oncogènes. Elle se produit soit par altération de leurs frontières, via par exemple une méthylation excessive aberrante à leur niveau, soit par des réarrangements chromosomiques, comme par exemple une inversion ou une délétion. Dans certaines LAL-B, le proto-oncogène *FLT3* se retrouve par exemple surexprimé à la suite de délétions ciblant le locus 13q12.2, qui conduisent à l'interruption d'un TAD et modifient les interaction promoteur-enhancers de la région : on parle de « détournement d'enhancer »¹⁸⁵ (**Figure 20**). Aussi, les mutations affectant le complexe cohésine constituent une nouvelle classe de mutations drivers¹⁸⁶. Elles semblent intervenir précocement dans la pathogenèse, et contribuer à l'aggravation des SMDs. Toutes les sous-unités du complexe peuvent être mutées de manière exclusive, STAG2 étant la cible la plus fréquente¹⁸⁷. Enfin, on observe une hausse de fixation de CTCF sur la chromatine, dans les LAM, inversement corrélée à la méthylation de l'ADN¹⁸⁸.

C. Thérapies épigénétiques pour LAM

Les modifications épigénétiques étant dynamiques et réversibles, elles constituent des cibles thérapeutiques intéressantes. Les cibles les plus prometteuses sont la méthylation de l'ADN et la désacétylation des histones. Les inhibiteurs utilisés, en monothérapie ou thérapie combinée, visent à induire la différenciation, l'apoptose, l'inhibition de croissance, l'arrêt du cycle cellulaire et la mort cellulaire¹⁸⁹.

1. Inhibiteurs de méthyltransférase de l'ADN

L'inhibition des DNMTs se fait via l'utilisation d'agents hypométhylants (HMAs) comme l'azacitidine et la décitabine d'ores et déjà approuvés cliniquement. L'azacitidine s'incorpore à l'ARN, la décitabine à l'ADN, et tout deux entraînent la dégradation des DNMTs, et donc une hypométhylation de l'ADN. Les inhibiteurs de DNMTs (DNMTis) ont pour but principal de restaurer l'expression de gènes passés sous silence, mais ils stimulent également une réaction immunitaire anti-leucémique suite aux dommages qu'ils causent à l'ADN, et à l'expression d'antigènes associés aux tumeurs qu'ils induisent. Aussi, puisque l'azacitidine phosphorylée a tendance à se maintenir dans l'ARN, elle perturbe la synthèse protéique et induit l'apoptose¹⁸⁹. Le traitement est cytotoxique à forte dose, de nouvelles molécules sont en développement pour y pallier, comme par exemple une version orale des composés¹⁸⁹.

Les DNMTs sont des cibles intéressantes car l'hypométhylation favorise l'apoptose des cellules cancéreuses, tout en restant surmontable pour les cellules saines. La piste est d'autant plus pertinente que la méthylation semble jouer un rôle important dans l'évolution de MDS en LAM.

2. Inhibiteurs d'acétylases

Les inhibiteurs d'HDACs (HDACis) visent à induire une hyperacétylation des régions réprimées (acétylation globale des histones). Ils induisent un arrêt du cycle cellulaire et l'apoptose. Les HDACis touchent surtout les marques H2AK9ac, H3K14ac et H3K23ac. Leur action semble pour l'instant faible dans un contexte de LAM, des combinaisons de molécules sont en cours d'essai. Ces composés ont également une toxicité sur le long terme non négligeable du fait de l'hypo- ou hyperacétylation générée¹⁹⁰.

La combinaison DNMTis et HDACis permet généralement une réexpression de suppresseurs de tumeur et régulateurs du cycle cellulaire mis sous silence épigénétiquement. Souvent déjà utilisées pour traiter les tumeurs solides, ces nouvelles molécules sont encore au banc d'essai auprès des patients atteints d'hémopathies. L'obstacle thérapeutique majeur est encore une fois la non-sélectivité des inhibiteurs. Aussi, l'efficacité des agents seuls, sans chimiothérapie concomitante, reste limitée...

Chapitre 4

Le gène homéotique *CDX2*

La recherche d'autres altérations par RNA-seq, WES, CGH-array, à partir de cohortes de patients, a mené à l'observation d'une expression ectopique du gène CDX2 dans 80% des leucémies aiguës.

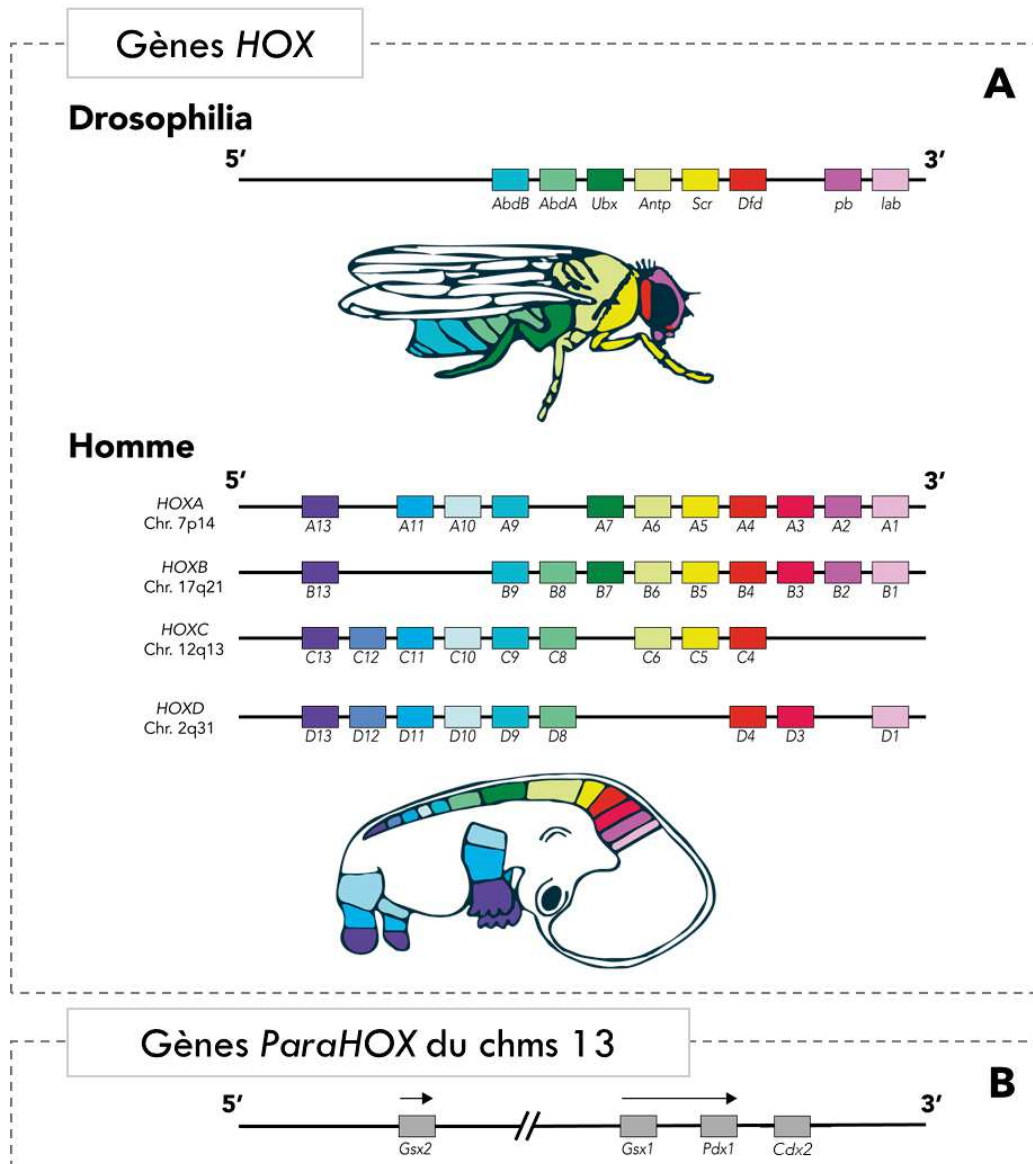


Figure 21. Les gènes HOX et ParaHox

(A) Les gènes HOX présentent une colinéarité spatiale d'expression : chaque cluster est rangé sur le chromosome dans le même ordre que celui de leur zone d'expression le long de l'axe antéropostérieur. Ils présentent également une colinéarité temporelle d'expression : ceux situés en 5' d'un cluster sont les plus tardifs à s'exprimer. Figure adaptée de Durston et. al, 2011. (B) Les gènes du locus Para-Hox, ici du chromosome 13, présentent également une expression colinéaire spatiale et temporelle (Brooke et al., 1998).

I. Cdx2 : gène homéotique à homéoboîte

A. Origine, structure et fonction

Caractérisés pour la première fois chez la *Drosophile* en 1978, les gènes homéotiques sont en charge de la mise en place du plan d'organisation de l'organisme au cours du développement embryonnaire. Ils spécifient l'identité des différents segments qui composent l'embryon. L'altération d'un gène homéotique provoque ainsi la transformation partielle ou totale du ou des segments placés sous son contrôle, on parle d'homéose¹⁹¹. Au sein de cette grande famille se distinguent les gènes à homéoboîte (homeobox genes). Ils se caractérisent par la présence d'une séquence conservée de 183 nucléotides, codant pour un homéodomaine de 61 acides aminés. Ce domaine consiste en 3 hélices *alpha* organisées en une structure hélice-boucle-hélice, leur permettant d'interagir avec l'ADN au niveau de séquences consensus riches en nucléotides A et T.

On distingue 2 classes de gènes homéotiques chez l'Homme. Présents au nombre de 39 et répartis en 4 clusters distincts, les gènes homéotiques de classe I sont des facteurs de transcription, plus connus sous le nom de « gènes *HOX* ». Pendant le développement embryonnaire, leur domaine d'expression restreint le long de l'axe antéro-postérieur confère une identité positionnelle (**Figure 21A**). Les gènes homéotiques de classe II, facteurs de transcription également, sont davantage dispersés dans le génome et forment de plus petites familles. Parmi eux se distingue notamment le cluster ParaHox, localisé chez l'Homme sur le chromosome 13, composé de *GSX1*, *PDX1* et *CDX2*¹⁹² (**Figure 21B**). La similarité de séquence entre l'ensemble de ces gènes homéotiques suggère que ces clusters résultent de duplications d'un locus protoHox initial^{193,194}.

B. Caudal Type 2 Homeobox gene 2

Mis en évidence chez la *Drosophile* par mutagenèse aléatoire, le gène Caudal (Cad) est en charge de la spécification de l'extrémité postérieure de l'embryon, et de l'identité des tissus qui s'y mettent en place. Son invalidation conduit à une troncation postérieure¹⁹⁵.

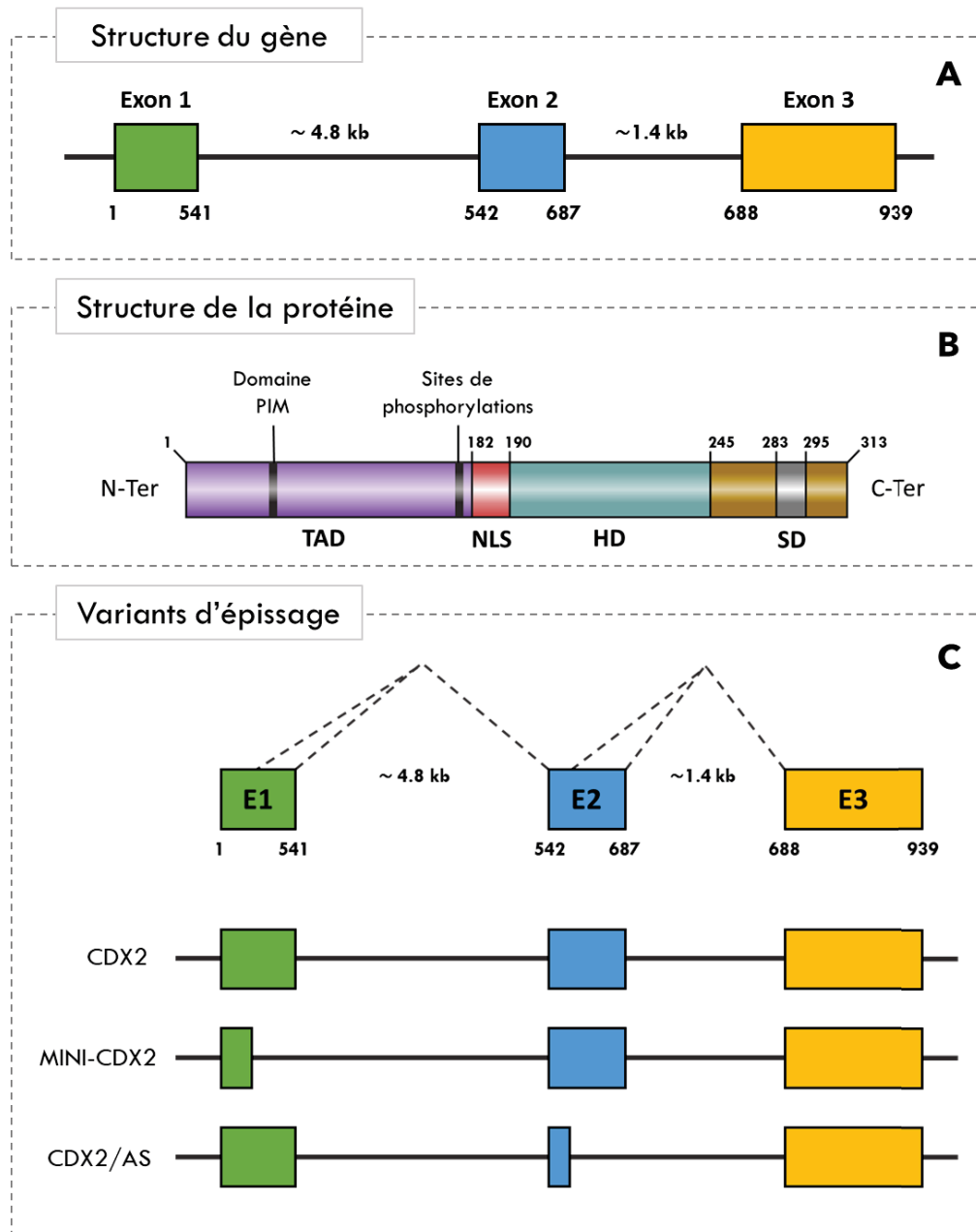


Figure 22. Le facteur de transcription homéotique CDX2

(**A**) La séquence codant pour le gène CDX2 se compose de 3 exons. Cette représentation schématique n'est pas à l'échelle. (Yagi et. al, 1999). (**B**) Le facteur de transcription se compose de 313 acides aminés, formant un domaine de transactivation (TAD), un signal de localisation nucléaire (NLS), un homéodomaine (HD), un domaine de stabilisation (SD) et un domaine d'interaction avec les PBX (PIM). Figure adaptée de Rawat et. al, 2012. (**C**) Le variant d'épissage MINI-CDX2 est dépourvu de domaine de transactivation, de par la troncation du 1er exon (Balbinot et., 2017). Le variant d'épissage CDX2/AS possède un homéodomaine tronqué, de par la délétion de 4pb dans le 2e exon, introduisant un décalage du cadre de lecture et modifiant et tronquant en conséquence la région C-terminale de la protéine (Witek et. al, 2014).

Chez les mammifères, le gène Cdx2, « Caudal type homeobox 2 » par homologie de séquence avec ce gène *Caudal*, se compose d'environ 7,2kb répartis en 3 exons et 2 introns (**Figure 22A**). Une protéine de 313 acides aminés en découle ; elle possède 3 domaines fonctionnels lui permettant d'exercer sa fonction principale de facteur de transcription. A l'extrémité N-terminale se trouve son domaine de transactivation qui, pourvu d'un site de phosphorylation, régule son activité transcriptionnelle. Se présente ensuite l'homéodomaine, riche en aminoacides basiques, qui assure sa liaison aux promoteurs des gènes cibles au niveau de séquences C/AATAAAA/Y. Un domaine de stabilisation présentant des sites de phosphorylation occupe l'extrémité C-Terminale ; il régit sa dégradation par le protéasome et donc sa demi-vie¹⁹⁶ (**Figure 22B**).

Cdx2 possède 2 variants d'épissage. « Cdx2 A/S » résulte d'un décalage du cadre de lecture et possède un domaine homéotique tronqué l'empêchant de se lier à l'ADN de ses cibles transcriptionnelles. Le décalage enrichit par ailleurs son extrémité C-Terminale en résidus sérine et arginine, lui permettant une activité d'épissage au niveau de minigènes¹⁹⁷. Le second, nommé Mini-Cdx2, mis en évidence au laboratoire, est dépourvu de domaine de transactivation et fait office de dominant-négatif en encombrant les sites de fixation de Cdx2 au niveau de ses gènes cibles¹⁹⁸ (**Figure 22C**).

Cdx2 possède 2 paralogues, Cdx1 et Cdx4, localisés respectivement chez l'Homme sur le chromosome 5 et le chromosome X. Ils possèdent de nombreuses cibles transcriptionnelles communes, et donc une certaine redondance dans leurs actions, au cours de processus développementaux¹⁹⁹. Les gènes Cdx sont notamment des régulateurs directs des gènes Hox, via les sites consensus de fixation présents au niveau de leurs clusters²⁰⁰⁻²⁰².

L'intérêt tout particulier porté à Cdx2 se justifie par ses nombreux patterns d'expression, cela dans divers contextes cellulaires, aussi bien physiologiques que pathologiques.

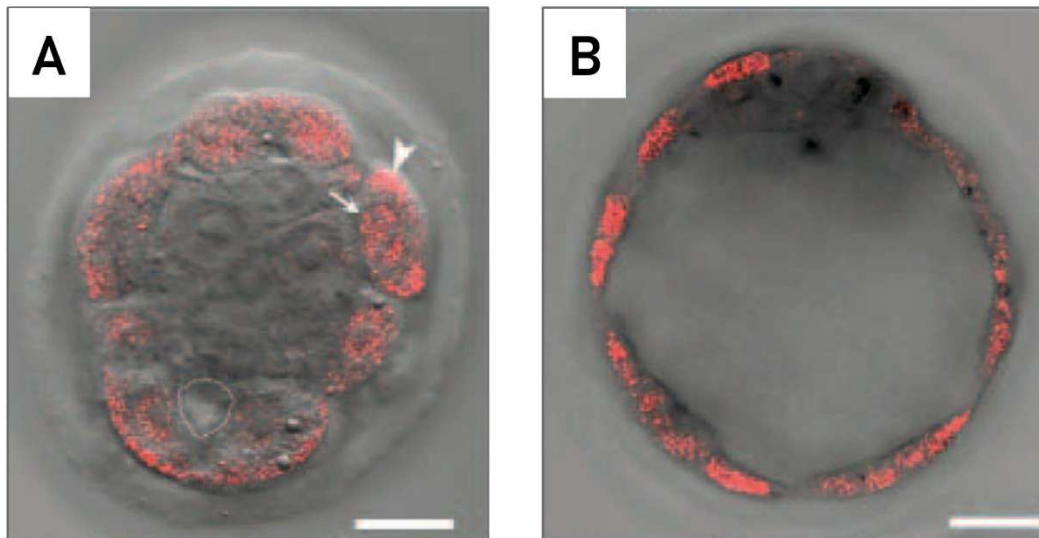


Figure 23. Expression de *Cdx2* dans les cellules externes de la morula

Immunohistochimie avec anticorps polyclonal ciblant *Cdx2*, couplé à un fluorochrome rouge. (A) Blastocyste de souris à 3,25 jours post-coïtum. (B) Blastocyste de souris à 3,5 jours post-coïtum. Barre d'échelle : 15µm. Photos extraites de Strumpf et. al, 2005.

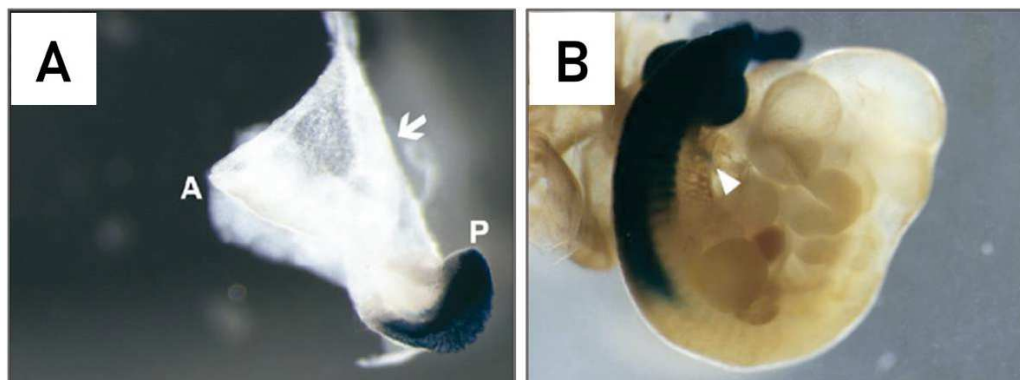


Figure 24. Expression de *Cdx2* dans la partie caudale de l'embryon

(A) Le gène X-gal, placé sous la dépendance du promoteur du gène *Cdx2*, s'exprime dans la partie postérieure (P) de l'embryon de souris (coloration bleue) à E8,5, mais pas dans sa partie antérieure (A) et dans les annexes embryonnaires (flèche). (B) A E10,5 la partie postérieure de l'embryon de souris est toujours colorée. La tête de flèche blanche indique le tube digestif. Photos extraites de Benahmed et. al, 2008.

II. Fonctions physiologiques de Cdx2...

A. ... au cours de l'embryogenèse

1. Établissement du trophoderme

Les fonctions de Cdx2 au cours du développement embryonnaire sont particulièrement bien identifiées chez la souris. Il intervient tout d'abord au stade précoce de 8 cellules, à E3.5, lors de la blastulation, pour la détermination de la lignée trophodermique. Il s'agit du premier événement de différenciation, nécessaire à la nidation et à la nutrition de l'embryon.

Son expression est stimulée au niveau des cellules extérieures de la morula, notamment par les facteurs de transcription Tead4, Rbpj, Oct1, et Gata3, et par un mécanisme d'autorégulation²⁰³⁻²⁰⁹ (**Figure 23**). Dans ces cellules, il réprime l'expression de gènes de pluripotence comme Nanog, Sox2 et Oct4²¹⁰⁻²¹³. Les cellules internes de la morula, où son expression est réprimée par Oct4, conservent leur pluripotence et évoluent en cellules de la masse interne^{206,214,215}. Cette régulation différentielle entre cellules intérieures et extérieures de la morula est dépendante des voies de signalisation Notch et Hippo²⁰⁴⁻²⁰⁶. Une absence d'expression de Cdx2 (dans les souris Cdx2^{-/-}) est létale à ce stade du développement, les embryons ne parvenant pas à s'implanter dans la paroi utérine. À l'inverse, son expression forcée dans des cellules ES in vitro stimule la différenciation trophoblastique²⁰⁹.

2. Élongation postérieure de l'embryon

Toujours chez la souris, après la gastrulation, autour de E8,5, il est exprimé dans la partie caudale de l'embryon, au niveau des 3 feuillets embryonnaires, où il contribue à son élongation postérieure²¹⁴ (**Figure 24**).

À ce stade, Cdx2 et ses paralogues sont des régulateurs directs essentiels des gènes Hox, participant au processus d'élongation et de spécification des différents segments²¹⁶⁻²²⁰. Sa perte de fonction, induite post-implantation pour contourner sa létalité précoce, est associée à une antériorisation de leurs domaines d'expression. Ce phénomène, nommé « *shift* homéotique antérieur » se traduit par des troncations postérieures et des perturbations de la mise en place des vertèbres et membres inférieurs^{221,222}. L'expression d'autres acteurs clés de l'élongation axiale est amoindrie en absence de Cdx2, dont notamment Wnt3a^{220,223}.

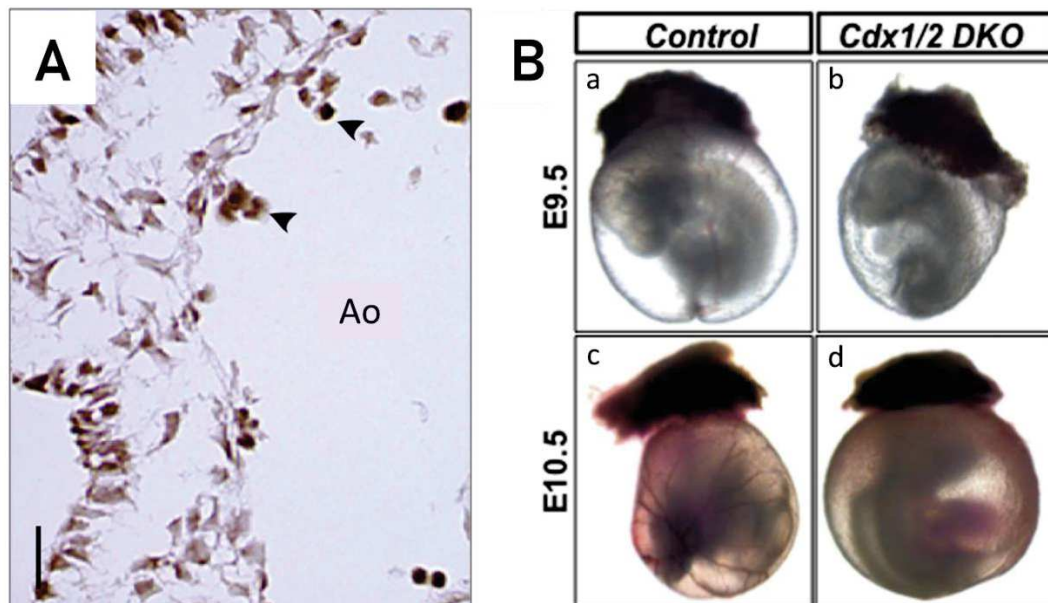


Figure 25. Expression de *Cdx2* pendant l'émergence hématopoïétique

(A) Section au niveau de l'aorte (Ao) d'un embryon de souris à E10,5. Immunomarquage anti-CDX2. Les têtes de flèches noires indiquent des clusters hématopoïétiques intra-aortiques. Barre d'échelle : 10µm. Photo extraite de El Omar et. al, 2021. (B) Les embryons KO pour *Cdx1* et *Cdx2* présentent moins de vascularisation et de globules rouges dans le sac vitellin. Photos extraites de Brooke-Bisschop et. al, 2017.

Cette interaction témoigne du rétrocontrôle positif de Cdx2 sur la voie WNT, en charge de son expression pendant l'embryogenèse²²⁴.

3. Hématopoïèse primitive

CDX2 participe au même moment à la mise en place de l'hématopoïèse primitive et à la vascularisation du sac vitellin (**Figure 25**).

Sa perte de fonction (Cdx2^{-/-}) post-implantatoire avait effectivement déjà permis de soupçonner des défauts dans le développement du mésoderme extra-embryonnaire, à l'origine des tissus hématopoïétiques²²². Bien qu'elle soit associée à la diminution d'expression de facteurs de transcription hématopoïétiques (*Gata1*, *Runx1*, *Scl*, ...), son impact n'est pas clair chez la souris. Sans pour autant être indispensables à l'initiation des premiers événements hématopoïétiques, Cdx2 et ses paralogues (Cdx1 et Cdx4) contribuent conjointement à leur maintien. Leur action y semble particulièrement redondante, tous étant des régulateurs des gènes *HOX*. Leur perte individuelle serait ainsi compensée par le maintien des 2 autres²²⁵⁻²²⁷. Au niveau de l'endothélium hémogénique, SOX17 semble favoriser l'accessibilité du promoteur de Cdx2, et donc son expression²²⁸. Pendant ce processus, Cdx2 est à nouveau indispensable à l'expression des gènes *Hox*, qui sont des acteurs clés de l'hématopoïèse chez l'embryon puis tout au long de la vie. Ils régulent la prolifération et différenciation des progéniteurs hématopoïétiques, leur expression faiblissant fortement par la suite^{229,230}.

Par ailleurs, il a été observé que la perte de fonction de CDX2 et -surtout- de CDX4 altère l'expansion des vaisseaux allantoïques et leur branchage dans l'ectoderme chorionique, ce qui en conséquence empêche le développement du placenta et du cordon ombilical²¹⁷. Leur perte de fonction conjointe perturbe également l'hématopoïèse primitive et la vascularisation du sac vitellin en réduisant notamment l'expression du facteur de transcription SCL (TAL1)²³¹. Il a également été montré récemment que CDX2 se fixait sur le promoteur de l'ACE (*Angiotensin Converting Enzyme*), facteur clé de l'émergence des CSH²³².



Figure 26. Expression de Cdx2 dans l'endoderme postérieur

Le gène X-gal, placé sous la dépendance du promoteur du gène Cdx2, s'exprime le plus fortement au niveau du tube digestif, issu de l'endoderme postérieur, de l'embryon de souris à E13 (tête de flèche blanche). Photo extraite de Benahmed et. al, 2008.

4. Spécification du tissu intestinal

Le tube digestif dérive de l'endoderme définitif, qui se différencie progressivement et chronologiquement de son extrémité antérieure à son extrémité postérieure. L'expression de *CDX2* se restreint, à partir du stade E12,5 à l'endoderme postérieur, auquel il confère une identité intestinale (**Figure 26**).

Il a été observé que la voie WNT stimulait l'expression de *CDX2* au niveau de l'endoderme précisément entre E7,5 et E8,5²³³. Son expression est également modulée par des facteurs tels HNF- α , GATA6 et TCF4²³⁴. La perte conditionnelle de son expression dans l'endoderme postérieur à E9 perturbe la morphogenèse de l'intestin, et est létale peu après la naissance. Elle engendre une antériorisation du tube digestif via l'activation d'un programme de différenciation plus antérieur. On remarque en conséquence l'apparition de cellules kératinisées, caractéristiques de l'œsophage, et l'obstruction du tube digestif. Y est associée une baisse d'expression de facteurs de transcription pro-intestinaux tels HNF4- α et ISX²³⁵. Une délétion plus tardive de *CDX2*, à E13,5, transforme l'épithélium intestinal en une structure gastrique, avec apparition de marqueurs épithéliaux gastriques comme BARX1, SOX2, Claudine-18, MUC1, GHRL et des gastrokinases^{236,237}.

L'hypothèse est que *CDX2* puisse agir sur ses gènes cibles en fonction de l'accessibilité de la chromatine, qui fluctue au cours de la spécification de l'endoderme. Il ne possède donc qu'une fenêtre de temps restreinte, où l'environnement chromatinien est permissif, pour imposer une identité tissulaire intestinale. C'est pourquoi dans les modèles pré-cités, le timing de délétion de *CDX2* semble influencer le phénotype observé : plus celle-ci sera précoce, plus antériorisée sera l'identité acquise. Autrement dit, en absence d'un facteur imposant un destin intestinal, l'endoderme reste susceptible à une transformation homéotique antérieure²³⁸.

B. ... tout au long de la vie adulte

Chez l'adulte, *CDX2* n'est exprimé plus que dans l'épithélium intestinal, dont il est le célèbre facteur d'engagement clé. Son domaine d'expression crée une frontière nette entre estomac et duodénum. L'expression de *CDX2* augmente progressivement le long de l'intestin proximal, atteint son maximum au niveau du colon proximal, où on la retrouve tout le long de

l'axe crypto-villositaire. Elle décroît ensuite dans le colon distal en se localisant dans les cellules différenciées, cela jusqu'au rectum^{239,240}. Ses régulateurs essentiels sont *HNF4-α*, *GATA6*, *TCF4/β-caténine*. Ils ont démontré la capacité à activer l'expression endogène de *CDX2* dans des cellules non intestinales (Hela), via une régulation synergique²³⁴. *CDX2* est également capable de s'autoréguler, par fixation sur sa propre région promotrice, chez l'adulte²⁴¹. Dans le compartiment intestinal, *CDX2* maintient l'identité tissulaire et assure une homéostasie en régulant notamment la prolifération et la différenciation des cellules intestinales.

1. Fonctions transcriptionnelles dans l'intestin

CDX2 possède une activité majeure de facteur de transcription, qu'il exerce grâce à la capacité de son homéodomaine à se lier à l'ADN. Il peut se fixer à plus de 16 000 sites dans l'épithélium intestinal, au niveau de régions régulatrices aussi bien proximales que distales²⁴²⁻²⁴⁴. *CDX2* présente une occupation contextuelle de la chromatine en fonction de l'état de différenciation de la cellule, contexte cellulaire qui influence également ses partenaires transcriptionnels. Il interagit par exemple davantage avec *GATA6* dans les cellules non différenciées et proliférantes, et avec *HNF4-α* une fois leur différenciation achevée²⁴⁵. *CDX2* favoriserait d'ailleurs leur fixation au niveau de certaines régions, en étant capable d'influencer l'accessibilité de la chromatine^{246,247}. Il le ferait notamment via son interaction avec Brg1, une sous-unité du complexe de remodelage des nucléosomes SWI/SNF²⁴⁸.

CDX2 est exprimé dans toutes les cellules de l'épithélium intestinal monostratifié, sauf les cellules de Paneth, et notamment dans les cellules souches où il contribue à la détermination de leur identité intestinale et à leur détermination : en activant l'expression des facteurs de transcription *Klf4* et *Math1*, il les oriente vers un le lignage sécréteur^{249,250}, ou les dirige à l'inverse vers le lignage absorbant en stimulant l'expression de *Dll1*²⁵¹. Dans les cellules souches, la fixation précoce de *Cdx2* au niveau des séquences régulatrices de ses gènes cibles semble prédictive de leur expression²⁵². Dans les cellules intestinales matures, *CDX2* régule l'expression de nombreux gènes essentiels à leurs fonctions. On compte parmi ses gènes cibles : ceux codant pour des enzymes digestives comme la saccharase isomaltase²⁵³, la lactase²⁵⁴ ou la phospholipase A²⁴² ; pour des transporteurs comme *GLUT2*²⁵⁵ ou *ASBT*²⁵⁶ ; pour des mucines comme *MUC2*²⁵⁷ ou *MUC4*²⁵⁸ ; pour des récepteurs membranaires comme *FXR*²⁵⁹ ou *IRS2*²⁶⁰ ; et enfin ceux codant pour des protéines d'adhésion cellulaire comme la LI-cadhérine²⁶¹, la E-cadhérine²⁶² ou la Claudine-2²⁶³.

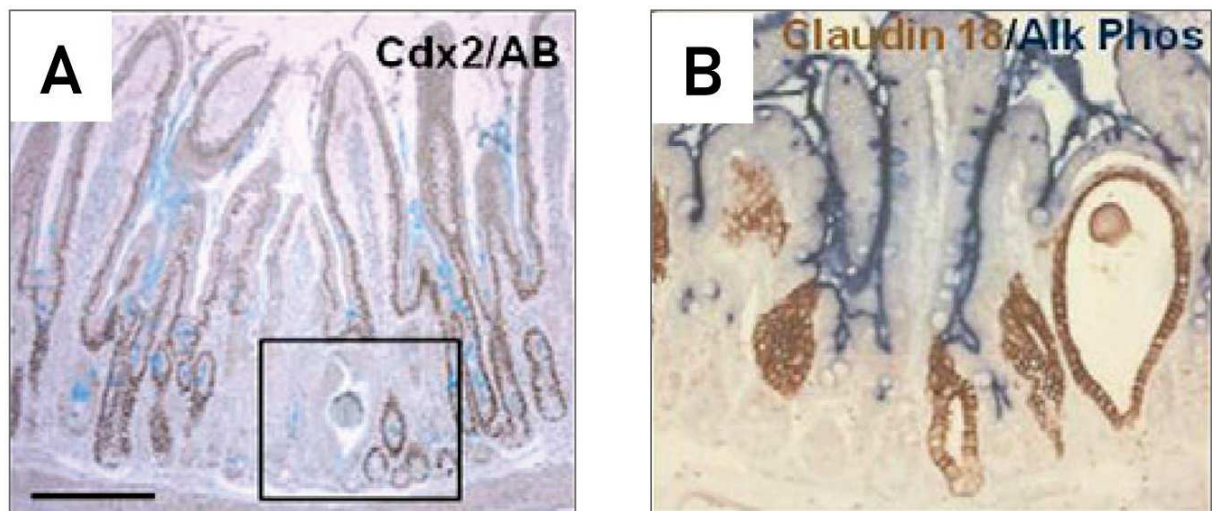


Figure 27. Conséquences de la perte de fonction mosaïque de Cdx2 dans l'épithélium intestinal

Chez la souris, sa perte de fonction mosaïque dans l'épithélium intestinal provoque l'apparition d'un phénotype tubulo-glandulaire typique de l'estomac. **(A)** Les cellules ne l'exprimant pas ne migrent plus le long de l'axe crypto villositaire et forment des vésicules cystiques (encadrées en noir). **(B)** On remarque chez ces cellules l'expression de Claudine18 (marqueur gastrique, en brun) et la perte d'expression de la phosphatase alcaline (marqueur intestinal, en bleu). Barre d'échelle : 100µm. Photos extraites de Stringer et. al, 2012.

CDX2 aiguillerait également la fixation du facteur de transcription TCF4, effecteur de la voie Wnt avec lequel il partage un grand nombre de cibles²⁶⁴.

La perte mosaïque de son expression dans l'épithélium intestinal n'altère pas son architecture comme on pouvait l'observer lors du développement embryonnaire. Elle perturbe en revanche localement la différenciation des cellules, et y induit un phénotype tubulo-glandulaire typique de l'estomac, soulignant ainsi son rôle essentiel dans le maintien d'une identité intestinale²³⁶(**Figure 27**).

CDX2 refreine la prolifération in vivo en diminuant directement ou indirectement l'activité de la voie β -caténine, cela en interagissant avec la β -caténine ou via la stimulation de l'expression d'antagonistes de la voie comme Mucdhl, Axine-2 et Apc²⁶⁵⁻²⁶⁷. In vitro, il a été montré que CDX2 stimulait directement l'expression de P21, un inhibiteur de Cdk²⁶⁸. Dans la lignée SW480, sa délétion augmente la prolifération cellulaire et l'apparition de marqueurs d'immaturité chez les cellules, alors que dans celle Caco2/TC7 sa surexpression est associée à un arrêt de la prolifération et stimule la différenciation^{269,270}.

CDX2 participe également à l'homéostasie plus globale de la muqueuse intestinale, notamment en contrôlant la population de cellules immunitaires. Il régule l'expression de H2-T3, membre du complexe majeur d'histocompatibilité. Sa perte réduit le nombre et la dégranulation des cellules CD8- α induites, et augmente l'infiltration des macrophages au niveau de l'épithélium intestinal, conduisant en une réaction inflammatoire chronique²⁷¹.

L'action transcriptionnelle de CDX2 est antagonisée par son variant d'épissage MiniCdx2. La surexpression de MiniCdx2 conduit à une perte d'identité intestinale et une transdifférenciation de type gastrique, soit au même phénotype que la perte de CDX2¹⁹⁸.

2. Fonctions non-transcriptionnelles dans l'intestin

Au-delà de son activité transcriptionnelle, CDX2 est capable d'interagir avec des partenaires protéiques diversifiant ses fonctions.

En interagissant avec le complexe protéique Ku70/Ku80, il réduit par exemple l'activité de la machinerie de réparation de cassures d'ADN double-brin de type NHEJ. L'efficacité réduite de la réparation conduit les cellules à l'apoptose, et leur évite l'introduction de potentielles mutations délétères par ce mécanisme de réparation peu fidèle²⁷². CDX2 interagit également avec le facteur de transcription HoxB7, qui à l'inverse soutient la réparation de type NHEJ, de sorte à empêcher sa fixation sur le promoteur de ses gènes cibles²⁷³. CDX2 est également un régulateur non-transcriptionnel de la machinerie d'épissage¹⁹⁸.

Deux mécanismes l'impliquent à nouveau dans la régulation de la prolifération cellulaire. Il stabilise d'une part p27Kip1, inhibiteur d'E-CDK2, ce qui bloque les cellules en phase G0/G1 de leur cycle²⁷⁴. En plus de son contrôle déjà transcriptionnel sur la voie WNT, CXD2 parvient d'autre part, via son domaine N-terminal, à interagir directement avec la β -caténine. Il perturbe son association avec le complexe protéique TCF, et inhibe ainsi l'activité de la voie²⁷⁵. Ces mécanismes lui confèrent une fonction de suppresseur de tumeur.

III. Fonctions de Cdx2 dans les cancers

L'expression de *CDX2* fluctue en contexte pathologique : elle peut aussi bien disparaître localement dans l'épithélium intestinal, que s'activer de manière aberrante en dehors de ce compartiment. On parle alors d'expression ectopique.

A. Perte d'expression dans les CCR

Les multiples casquettes de *CDX2* dans l'épithélium intestinal ont naturellement conduit à questionner les conséquences de sa perte. Les premières hétéroplasies de type gastro-œsophagien observées chez les souris hétérozygotes *Cdx2*^{+/-} laissaient penser à une initiation tumorale, mais n'évoluaient cependant pas en adénomes ou adénocarcinomes intestinaux^{221,276}.

Pourtant, la perte d'un allèle de *CDX2* au niveau de l'épithélium intestinal est retrouvée chez 30% des patients avec un CCR. Elle est associée à un haut grade tumoral (faible différenciation et fort taux de métastases entre autres), et à pronostic défavorable (survie globale plus faible).

L'expression de *CDX2* tend effectivement à être diminuée plus fortement dans les cancers intestinaux les plus agressifs, comme ceux de type festonné (sous-type C4)²⁷⁷. Elle est à l'inverse davantage maintenue dans ceux présentant les cellules les plus différenciées^{278,279}. *CDX2* a de ce fait été proposé comme biomarqueur pronostic de routine. Cet impact pronostic semblerait en fait dépendre du statut sous-jacent de la réparation des mésappariements chez les patients, indicative de la présence de métastases²⁸⁰⁻²⁸⁴. De par la complexité de sa régulation et de ses modes d'actions, il reste difficile d'utiliser l'expression de *CDX2* comme un marqueur pronostic ou prédictif.

B. Expression ectopique dans l'épithélium gastrique

Les métaplasies intestinales surviennent à la suite de traumatismes tissulaires chroniques associés à de la régénération. Ce sont des lésions pré-cancéreuses caractérisées par un « *switch* » phénotypique du tissu gastrique vers une différenciation intestinale. Les métaplasies intestinales de l'estomac impliquent donc une répression du programme transcriptionnel gastrique, associée notamment à la perte d'expression du facteur clé *SOX2*, et à un gain de différenciation intestinale, gouvernée dans 90% des cas par l'expression ectopique de *CDX2*. Paradoxalement, bien que son expression ectopique initiale favorise la transdifférenciation et soit donc plutôt oncogénique, les patients *CDX2+* présentent une survie globale supérieure^{285,286}.

Une expression forte de *CDX2* permettant une meilleure différenciation des cellules cancéreuses, son maintien devient alors protecteur, conformément à son rôle suppresseur de tumeur intestinal^{287,288}. Avec l'avancée de la maladie, la perte progressive de l'expression de *CDX2* est associée à un moins bon pronostic, et est corrélée avec la perte de *p53*²⁸⁹. Dans ces métaplasies dites « incomplètes », l'absence de différenciation terminale, aussi bien gastrique qu'intestinale, rend les cellules cancéreuses davantage instables. *CDX2* constitue de ce fait un biomarqueur intestinal pronostique²⁹⁰.

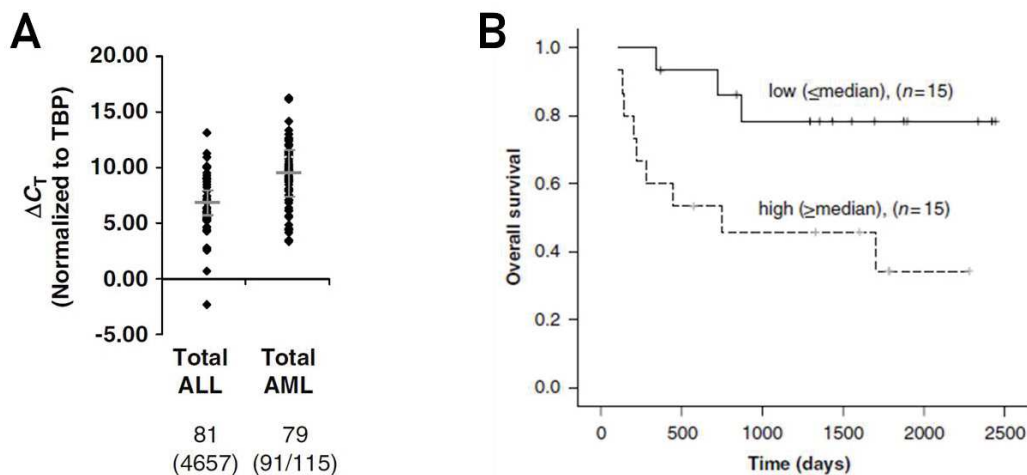


Figure 28. Expression aberrante du gène Cdx2 dans la moelle osseuse

Une expression aberrante de Cdx2 est retrouvée dans la moelle osseuse de 80% des patients atteints de leucémie aiguë, aussi bien myéloïde que lymphoïde. Cette expression représente un marqueur de mauvais pronostic. **(A)** Analyse de l'expression de Cdx2 dans la moelle osseuse de patients atteints de leucémie aiguës lymphoïdes (LAL) et myéloïdes (LAM), par RT-qPCR. Les valeurs de ΔC_T sont normalisées par rapport au gène de ménage TATA binding protein. Chaque point représente un patient. Les barres correspondent à l'expression médiane. **(B)** Courbe de Kaplan-Meier représentant la survie globale des patients atteints de LAL exprimant plus ou moins fortement Cdx2. Figures extraites de Thoene et. al, 2009.

C. Expression ectopique dans le compartiment hématopoïétique

Le gène *CDX2* est important lors de l'hématopoïèse primitive, mais n'est normalement plus exprimé dans le système sanguin chez l'adulte. Son expression ectopique est néanmoins observée dans 80% des cas de leucémie aiguë, au niveau de la moelle osseuse des patients. Son expression est mono allélique²⁹¹. Le mécanisme moléculaire menant à son activation reste inconnu, mais celle-ci représente un marqueur de mauvais pronostic chez les patients souffrant de leucémie aiguë myéloïde (LAM) ou lymphoïde (LAL), son niveau d'expression étant corrélé à l'agressivité de la pathologie. En revanche, les patients atteints de leucémies chroniques ne présentent pas d'expression de *CDX2*²⁹²⁻²⁹⁴ (**Figure 28**).

Le gène *CDX2* contribuant à l'hématopoïèse chez l'embryon, sa réactivation lors des leucémies aiguës rappelle le phénomène de résurgence de l'expression embryonnaire, que l'on observe avec certains gènes développementaux dans des contextes pathologiques^{295,296}.

D. Modélisation de l'expression dérégulée de Cdx2

La dérégulation de l'expression de *CDX2* est modélisée in vitro et in vivo, afin d'en analyser les conséquences, et surtout son impact sur la pathogenèse...

1. Perte d'expression dans l'épithélium intestinal

En contexte protumoral, stimulé par un traitement cancérigène à l'azoxyméthane, les souris *Cdx2*^{+/-} développent plus d'adénocarcinomes dans le colon distal que les souris contrôles. Sa diminution d'expression sensibilise à la tumorigenèse, les polypes bénins évoluant en adénocarcinomes²⁹⁷. Les mêmes observations sont faites chez les souris *Cdx2*^{+/-} sur fond de mutation *Apc*^{+/ Δ 14} : la faible expression de *Cdx2* et *Apc* au niveau du colon distal est associée à l'apparition de polypes adénomateux²⁹⁸. Ces deux études indiquent qu'une diminution de l'expression de *Cdx2* est décisive dans la tumorigenèse colique. On observe d'ailleurs une cooccurrence entre la perte de *Cdx2* et les mutations du proto-oncogène *Braf*. Ces événements, seuls ou combinés, sont associés à un mauvais pronostic et prévalents dans les CCR de forme

festonnée. La modélisation de cette combinaison d'altérations chez la souris *Cdx2^{null}/Braf^{V600E}* indique une réduction de la survie globale et démontre leur coopération²⁹⁹. La perte de CDX2 est donc un événement précoce de la progression tumorale.

In vivo toujours, il a été observé que dans un contexte protumoral (*Apc^{+/-Δ14}*) et de KO mosaïque de *Cdx2*, les cellules déplétées en *Cdx2* induisent une tumorigenèse dans les cellules adjacentes qui expriment encore *Cdx2*, sans devenir elles-mêmes cancéreuses. Le mécanisme à l'origine de cette distinction est inconnu, mais les cellules *Cdx2*⁻ déclenchent une activation chronique de la voie NFκB et induisent la production d'oxyde nitrique synthase (iNOS) par le microenvironnement stromal. Ceci favorise la perte d'hétérozygotie d'*Apc* dans les cellules voisines et leur transformation tumorale³⁰⁰. La suractivation de la voie Wnt, par inactivation d'APC, est un événement précoce fréquent dans les CCR. En absence de *Cdx2*, la levée d'inhibition de la voie accélère la formation tumorale en induisant la progression G1/S dans le cycle cellulaire^{267,301}. Aussi, *in vitro* comme *in vivo*, la diminution d'expression de CDX2 facilite la progression, l'invasion et la migration des cellules tumorales³⁰². Enfin, la réintroduction de son expression antagonise la transition épithélio-mésenchymateuse (EMT) et donc la métastase, dans les CCR, via la régulation de l'expression de *SNAIL* et la stabilisation de la β-caténine³⁰³.

In vitro cependant, il a été montré que l'amplification du locus de CDX2 dans des cellules cancéreuses colorectales, était associé à une augmentation de son expression, et contribuait à leur survie et la prolifération³⁰⁴. Ainsi, malgré son rôle de suppresseur de tumeur dans ce compartiment, CDX2 semble jouer un rôle oncogénique quand il y est surexprimé.

2. Gain d'expression dans l'épithélium gastrique

Des modèles murins d'expression de CDX2 dans l'épithélium gastrique ont été mis au point via l'introduction de transgènes plaçant sa séquence sous la dépendance du promoteur de *Foxa3* ou de la sous-unité beta de l'ATPase H⁺/K⁺. Ceux-ci ont montré que sa seule présence suffit à induire la transdifférenciation de l'épithélium gastrique en une structure épithéliale de type intestinale. Là, il régule l'expression de gènes intestinaux tels *MUC2*, la phosphatase alcaline ou encore la villine³⁰⁵. La régulation d'expression de la villine passerait d'ailleurs dans ce contexte également par le recrutement, au niveau de son promoteur, du complexe SWI/SNF par CDX2³⁰⁶. Les métaplasies intestinales développées évoluent ensuite de manière autonome en

carcinomes gastriques invasifs, ces observations confirmant qu'il s'agit de véritables lésions précancéreuses^{307,308}.

In vitro, il a été observé qu'une fois ectopiquement exprimé, *CDX2* réprime la voie de signalisation Sonic Hedgehog (SHH), en se fixant directement au promoteur du gène éponyme³⁰⁹. Sa présence promeut également la migration des lignées gastriques AGS et MKN45, via la stimulation de Reg IV, déjà connu pour son rôle dans la métastase et la migration des cellules cancéreuses gastriques³¹⁰. Une étude a également démontré son rôle dans la prolifération, l'invasion et l'apoptose des cellules cancéreuses gastriques³¹¹. Néanmoins, une fois la dysplasie stabilisée et lors de la progression du carcinome, la surexpression de *CDX2* témoigne, *in vitro* comme *in vivo*, de son rôle protecteur en antagonisant croissance, invasion et EMT des cellules cancéreuses gastriques³¹².

3. Gain d'expression dans le compartiment hématopoïétique

In vitro, la suppression de *CDX2* inhibe la prolifération des cellules de lignées de LAM et réduit leur potentiel de clonogénicité. Aussi, des expériences de greffe de cellules de moelle osseuse exprimant la protéine *CDX2* murine chez des souris irradiées ont montré que *CDX2* était capable de transformer les CSH^{291,313}. *In vitro*, l'expression de *CDX2* augmente le potentiel d'autorenouvellement et le potentiel leucémogène des progéniteurs hématopoïétiques, cela notamment via la dérégulation de certains gènes *HOX*, classiquement dérégulés au cours des LAM. On constate les mêmes transformations en cas de perte de *CDX2* ou des *HOX* dans certains modèles murins. Ses cibles notables sont *HOXA9* et *MEIS1*, essentiels à la progression de la leucémie^{229,291,314-316}. Il inhibe également le gène Kruppel-like factor 4 (*KLF4*), doté d'une fonction duelle dépendante du contexte cellulaire³¹⁷. Ces résultats suggèrent un rôle oncogénique de *CDX2* lorsque celui-ci s'exprime dans le compartiment hématopoïétique.

IV. Mécanismes de régulation de l'expression de Cdx2

A. Régulation post-transcriptionnelle et post-traductionnelle

En contexte physiologique, l'expression de *CDX2* est régulée par des mécanismes post-transcriptionnels impliquant des miRNAs. Son ARNm peut ainsi être dégradé par un ciblage par miR-9, miR-16 ou miR-22, sous l'impulsion de *CDX1*³¹⁸. La protéine MEX3A contrôle également la demi-vie de son ARNm dans les cellules souches intestinales, en liant sa région 3' UTR, et en l'adressant aux corps P, en charge de la dégradation des ARNm^{319,320}.

La protéine *CDX2* est ensuite sujette à des modifications post-traductionnelles influençant sa fonction, sa stabilité et sa demi-vie. La phosphorylation de la sérine 60 de son domaine de transactivation inhibe par exemple son activité transcriptionnelle dans les cryptes intestinales³²¹. Dans les cellules différenciées des villosités, sa phosphorylation par la MAPK p38 stimule à l'inverse son activité³²². La phosphorylation de son domaine C-Terminal par la kinase CDK2 (sérine 281 et 293) ou par ERK1/2 (site inconnu) induit son ubiquitination puis sa dégradation par le protéasome^{196,323}. A l'inverse, l'inhibiteur p27Kip1 antagonise sa phosphorylation ; son profil d'expression se superpose d'ailleurs à celui de *CDX2* le long de l'axe crypto-villositaire³²⁴.

B. Régulation transcriptionnelle chez l'embryon précoce

De la gastrulation à la mi-gestation, une région de 5kb de chromatine active en amont du premier exon de *CDX2* constitue sa région promotrice. Aussi, un enhancer de *CDX2* intervenant spécifiquement au stade blastocyste a été mis en évidence. Situé environ 5,5kb en amont du TSS, il est fixé par les facteurs de transcription Tead4 et Rbpj, respectivement effecteurs des voies de signalisation Hippo et Notch²⁰⁴. *CDX2* possède un second enhancer situé dans son premier intron, actif au cours du développement, dans le mésoderme postérieur³²⁵.

A ce stade, *CDX2* est connu pour être un gène bivalent. Son expression est permise à la fois par la présence de modifications épigénétiques activatrices, et la levée de modifications inactivatrices, au niveau de son promoteur^{138,326}. *Cdx2* est notamment réprimé par la présence

de méthylations sur H3K9 au niveau de son promoteur, dans les cellules de la masse interne du trophoblaste^{327,328}.

C. Régulation transcriptionnelle dans l'intestin et les CCR

Le maintien de son expression dans l'épithélium intestinal requiert par la suite des éléments de régulation plus en amont du TSS, sur environ 9kb. Vers -8,5kb en amont du TSS, un segment de 250pb, qui consiste en un cluster de sites fixation de facteurs de transcription, est fixé par les protéines mentionnées précédemment : HFN4- α , GATA6, TCF4/ β -caténine²³⁴.

L'instabilité génomique est une caractéristique cellulaire typique des CCR et est essentielle à la progression tumorale. Le gène *Cdx2* se situe sur un locus (13q12.2) à risque en termes de réarrangements et amplifications³²⁹. Comme mentionné précédemment, son amplification lors de CCR contribue à la survie et la prolifération des cellules tumorales. Mutations somatiques et polymorphismes ne semblent pas participer à sa baisse d'expression^{330,331}.

Dans les CCR, le promoteur de *Cdx2* est sujet à une méthylation de ses îlots CpG. Les premiers traitements avec des agents déméthylants ne restaurent pas son expression *in vitro*, cette piste thérapeutique visant à rétablir son expression avait été écartée³³¹⁻³³³. Plus récemment cependant, l'inhibition de DNMTs dans les lignées SW620 et COLO205, exprimant faiblement *CDX2*, a démontré un fort rehaussement de son expression. L'expression de *CDX2* serait donc différentiellement impactée par la méthylation de son promoteur en fonction des lignées cellulaires. Dans ces 2 lignées, l'inhibition de HDAC4 et HDAC5 conduit au même résultat, HDAC5 s'avérant d'ailleurs recruté au niveau de la région promotrice de *CDX2*³³⁴.

Tout comme dans les cellules trophoblastiques, le facteur de transcription Oct1 est capable de recruter des complexes de remodelage de la chromatine au niveau du locus de *CDX2*. Des observations *in vitro*, montrent qu'Oct1 est à la fois capable de stimuler et d'inhiber la transcription de gènes, via le recrutement respectivement de NuRD et Jmjd1a/KDM3A. NuRD est responsable de la maintenance de la méthylation de l'ADN, tandis que Jmjd1a catalyse l'effacement des marques H3K9me2, ce qui induit la déméthylation de l'ADN voisin³³⁵.

D. Régulation transcriptionnelle dans les carcinomes gastriques

Les mécanismes en amont de son expression ectopique au niveau de l'épithélium gastrique sont encore globalement incompris. Deux stimuli environnementaux majeurs ont néanmoins été identifiés : l'exposition à *Helicobacter pylori*, phénomène le mieux caractérisé, et aux acides gastriques.

Plus de 90% des patients avec un cancer gastrique sont infectés par *Helicobacter pylori*³³⁶. L'infection provoque une gastrite chronique pouvant évoluer en métaplasie intestinale et ultimement en carcinome gastrique, ces deux états étant irréversibles même après éradication de l'infection³³⁷. L'expression de *CDX2* se perd néanmoins lors de la disparition d'*Helicobacter pylori*^{338,339}. *Helicobacter pylori* stimule la signalisation par la voie BMP, qui contribue au switch des profils d'expression de *CDX2* et *SOX2*^{285,340,341}. On remarque notamment la surexpression d'*ACVR1*, récepteur de la voie des BMP, lors de l'infection³⁴². Aussi, en sa présence, la signalisation découlant du récepteur intracellulaire NOD1 est diminuée. Celle-ci étant en charge de l'inhibition de la voie NFκB via TRAF3, sa diminution suractive la voie, et en conséquence contribue à l'expression de *CDX2*³⁴³. *Helicobacter pylori* favorise également la signalisation intracellulaire par les SMADs, dont *CDX2* est une cible transcriptionnelle, en stimulant l'expression de BMP2 et SMAD4³⁴⁴.

D'autre part, il a été observé, *in vitro*, que l'exposition prolongée des cellules gastriques à de l'acide désoxycholique (DCA) favorise l'expression de *CDX2* et ses cibles transcriptionnelles, cela de manière dose dépendante^{345,346}. Ces traitements avec des dérivés d'acide biliaire stimulent la voie NFκB, dont *CDX2* est une cible. Cela par exemple en augmentant l'expression de miR-92a-1-5p, capable de cibler les ARNm de FOXD1, un inhibiteur de la voie^{347,348}. Dans les lignées cancéreuses gastriques *in vitro*, il est la cible de nombreux régulateurs transcriptionnels, parmi lesquels HNF4α, TCF4, ISX, GATA6^{234,349}.

La méthylation de l'ADN est l'une des altérations génétiques les plus importantes des cancers gastriques. Bien que celle-ci fluctue au niveau du promoteur de *Cdx2* en conditions pathologiques, la différence avec le tissu sain n'est pas suffisante pour expliquer son expression ectopique au niveau des métaplasies^{350,351}. Le facteur de transcription progastrique Sox2 semblerait cependant avoir un rôle dans le maintien de la méthylation du promoteur de *CDX2*

dans les tissus gastriques. Les deux gènes sont connus pour exercer une répression réciproque entre eux, et il a été montré que le KD de *SOX2* stimulait la déméthylation du promoteur de *CDX2*^{352,353}. Aussi, la forte expression de DNMT1 et l'hyperméthylation du promoteur de *CDX2* contribuent à son inactivation transcriptionnelle dans la lignée MKN-45³⁵⁴.

E. Régulation transcriptionnelle dans les LAM

Dans le compartiment hématopoïétique, *CDX2* n'est pas impliqué dans des translocations ou remaniements notables décrits dans les leucémies aiguës, et ne présente pas d'hypométhylation ou mutations au niveau de son promoteur^{291,293}. La première piste investie était son implication dans le réarrangement chromosomique t(12 ;13)(p13 ;q12), générant le gène de fusion *ETV6-CDX2*. Son association avec un régulateur de la survie des CSH laissait supposer une contribution à la leucémogénèse³⁵⁵. Un modèle murin basé sur l'expression du gène de fusion a néanmoins permis d'invalider cette hypothèse, les souris ne développant qu'une myéloprolifération³⁵⁶.

Le facteur de transcription OCT1 se fixe également au niveau du promoteur de *CDX2* dans le compartiment hématopoïétique, en contexte physiologique comme pathologique. Il a été observé à partir d'échantillons humains qu'OCT1 ne va recruter JMJD1A que dans les cellules leucémiques³⁵⁷. Ce recrutement s'accompagne d'une diminution des marques H3K9me2, et d'une déméthylation totale des îlots CpG voisins. Les sites de fixation de OCT1 sont connus pour être associés à des changements dynamiques dans la méthylation de l'ADN au cours de l'hématopoïèse. Des méthyltransférases s'associent à des complexes reconnaissant la méthylation de H3K9, qu'il module avec ses cofacteurs³⁵⁸.

OBJECTIFS

CDX2 est un gène polyvalent, présentant diverses facettes selon l'environnement cellulaire auquel il est confronté. Malgré les connaissances amassées, les événements en amont et aval de son activation ectopique dans le compartiment hématopoïétique restent encore globalement incompris. Les pièces du puzzle se rassemblent, mais ces événements n'ont pas encore pu être replacés au sein de réseaux fonctionnels, dans ce contexte cellulaire. L'investigation bat son plein, grâce aux alliés non négligeables que sont modèles murins et nouvelles technologies d'analyse génomique...

Mon objectif premier était de modéliser son expression ectopique dans le compartiment hématopoïétique. Cela de sorte à pouvoir identifier les outils et chemins qu'il emprunte pour induire une leucémogénèse afin de comprendre les facteurs décisifs à sa dualité : lesquels peuvent bien peser dans la balance pour l'orienter vers une fonction d'oncogène ou de suppresseur de tumeur ?

L'origine de l'activation de l'expression de *CDX2* dans les leucémies aiguës a également déjà fait l'objet d'études, et celle-ci n'est pas due à un changement de méthylation de son promoteur, l'amplification de son locus ou la présence de mutations l'affectant. Son induction pourrait résulter d'une activation de sa transcription par des modifications épigénétiques, cela au niveau de son promoteur et/ou de régions enhancers. Le but était ici d'identifier des régions régulatrices de *CDX2*, cela via un travail comparatif entre 3 contextes cellulaires : intestinal, gastrique et hématopoïétique. A nouveau, le but de cette étude est d'étoffer le panel de données transversales le concernant.

RÉSULTATS

Partie 1

Contexte de la mise au point du modèle murin

A. Un premier modèle d'induction *ex vivo*

Une expression ectopique de CDX2 a été mise en évidence dans 80% des leucémies aiguës. Mais le rôle oncogénique de CDX2 restait à démontrer. En 2007, un premier modèle murin d'expression ectopique du gène Cdx2 dans les CSHs a été proposé²⁹¹. L'induction de son expression consistait en 2 étapes. Une surexpression s'effectuait dans un premier temps *in vitro*, par transduction d'un lentivirus contenant la séquence codant pour le gène hCdx2, placée sous la dépendance du promoteur du Cytomégalo virus (CMV). Les CSHs transduites étaient ensuite transplantées à des souris irradiées, par injection dans la veine de la queue. Les souris receveuses développaient des LAM, de phénotype immature, démontrant le rôle oncogénique de Cdx2 *in vivo*. Certains aspects du modèle étaient cependant critiquables : une induction effectuée *ex vivo*, plaçant qui plus est la séquence de Cdx2 en aval d'un promoteur à forte activité, puis l'irradiation au préalable des souris receveuses pour assurer la prise de greffe...

Notre idée était alors d'optimiser la modélisation, en garantissant une induction d'expression *in vivo*, d'une intensité modérée, cela dans un contexte de niche hématopoïétique intacte. Le but était de faire intervenir son expression ectopique dans l'environnement le plus physiologique possible, cela en respectant la compétition entre cellules l'exprimant ou non.

B. Modèle d'induction *in vivo*

Le modèle mis au point au laboratoire se base sur la technologie Cre-Lox, et est issu du croisement de 2 souches de souris transgéniques. La souche Rosa26-LSL-hCdx2 possède un transgène plaçant la séquence de l'ADNc codant pour la protéine CDX2 humaine en aval d'un stop de transcription floxé, dans le premier exon du locus ubiquitaire Rosa26. Les séquences murine et humaine du gène Cdx2 présentent une identité de séquence de 85% au niveau ARNm et 92% au niveau protéique. De nombreuses études *in vitro* ont montré que les deux protéines avaient les mêmes fonctions. Pour notre modèle d'expression ectopique, le gène Cdx2 a été choisi dans sa version humaine afin de pouvoir distinguer une éventuelle activation

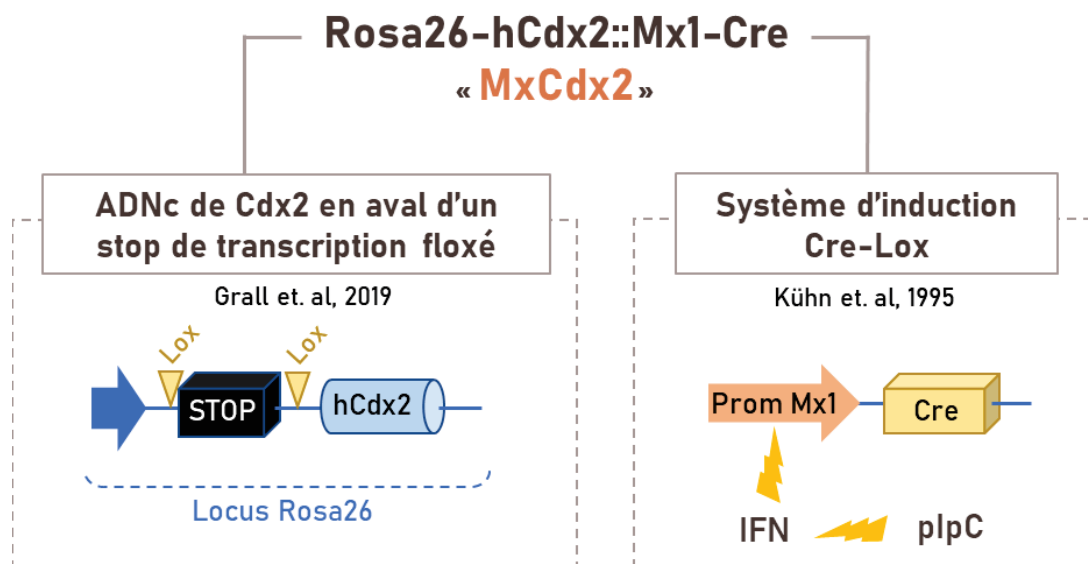


Figure 29. Induction de l'expression de Cdx2 dans le compartiment hématopoïétique

Le modèle murin Rosa26-hCdx2::Mx1-Cre est issu du croisement entre 2 lignées de souris transgéniques : les Rosa26-hCdx2 et les Mx1-Cre. Chez les Rosa26-hCdx2, la séquence codant pour hCdx2 est placée en aval d'un stop de transcription floxé, à l'intérieur du locus ubiquitaire Rosa26. Et chez le Mx1-Cre, le gène codant pour la recombinaise Cre est placé en aval du promoteur Mx1, spécifique des CSH et inductible à l'interféron de type 1 et de type 3. Pour déclencher la production d'IFN, nous injectons de l'acide polyinosinique-polycytidylique (pl-pC). Le pl-pC consiste en des petits ARN double brins, qui miment une infection virale. Grall et. al, 2021, Kühn et. al, 1995.

endogène de la forme murine, par un mécanisme d'autorégulation²⁴¹. La souche de souris Mx1-Cre possède un transgène plaçant la séquence de la recombinaise Cre sous la dépendance du promoteur Mx1. La recombinaise Cre est une intégrase issue du bactériophage P1, elle permet initialement la cyclisation de son génome, en catalysant la recombinaison entre sites LoxP via une activité topoisomérase de type I³⁵⁹. Le gène Mx1 code pour une GTPase à activité antivirale, dont l'expression est déclenchée dans les CSHs par la sécrétion d'IFN de type I et II, en présence d'ARN viral et quelques ADN viraux. La souche Mx1-Cre est classiquement utilisée pour la création de modèles murins d'expression inducible, particulièrement dans les CSH³⁶⁰. Les lignées murines du laboratoire ne sont pas de fond génétique pur. Ce parti-pris se justifie par une volonté de préserver l'hétérogénéité interindividuelle inhérente aux cohortes de patients, une consanguinité excessive pouvant biaiser les phénomènes physiologiques et pathologiques.

L'induction d'expression de Cdx2, chez les souris double transgéniques Rosa26-hCdx2 ::Mx1-Cre, dites «Mx-Cdx2», consiste en l'injection, à 3 reprises avec 1 journée d'intervalle, d'acide polyinosinique-polycytidylique (pl-pC). Il s'agit de petits ARN doubles brins, mimant une infection virale et déclenchant la production d'interféron α et β . Les injections déclenchent l'activité du promoteur Mx1, et donc la production de recombinaise Cre dans les CSHs. Dans ces cellules, le stop de transcription en amont de la séquence de Cdx2 est excisé, et Cdx2 est libre de s'y exprimer (**Figure 29**).

Ce modèle murin développe une LAM en moyenne dans les 90 jours suivant l'induction. Il nous a permis de nous intéresser aux conséquences cellulaires et moléculaires de l'expression de Cdx2....

CDX2 expression in the hematopoietic lineage promotes leukemogenesis via TGFb inhibition

Ava Galland¹, Victor Gourain² , Karima Habbas¹, Yonca G ler¹, Elisabeth Martin¹ , Claudine Ebel³, Manuela Tavian¹ , Laurent Vallat^{1,4} , Marie-Pierre Chenard⁵ , Laurent Mauvieux^{1,4} , Jean-No l Freund¹ , Isabelle Duluc¹  and Claire Domon-Dell¹ 

¹ Universit  de Strasbourg, Inserm, IRFAC / UMR-S1113, FHU ARRIMAGE, ITI InnoVec, FMTS, Strasbourg, France

² Institute of Biological and Chemical Systems, Karlsruhe Institute of Technology, Germany

³ Inserm, IGBMC, UMR-S 1258, Universit  de Strasbourg, Illkirch, France

⁴ Laboratoire d'H matologie, Centre Hospitalier Universitaire de Strasbourg, France

⁵ D partement de Pathologie, Centre Hospitalier Universitaire de Strasbourg, France

Keywords

acute monoblastic leukemia; BAMBI; ectopic expression; oncogene; tumor suppressor

Correspondence

J. -N. Freund, INSERM, UMR-S1113, 3 avenue Moli re, 67200 Strasbourg, France
Tel: +33 388 277727
E-mail: jean-no l.freund@inserm.fr

Ava Galland and Victor Gourain contributed equally to this work.

(Received 25 November 2020, revised 13 April 2021, accepted 15 April 2021, available online 26 June 2021)

doi:10.1002/1878-0261.12982

The intestine-specific caudal-related homeobox gene-2 (*CDX2*) homeobox gene, while being a tumor suppressor in the gut, is ectopically expressed in a large proportion of acute leukemia and is associated with poor prognosis. Here, we report that turning on human *CDX2* expression in the hematopoietic lineage of mice induces acute monoblastic leukemia, characterized by the decrease in erythroid and lymphoid cells at the benefit of immature monocytic and granulocytic cells. One of the highly stimulated genes in leukemic bone marrow cells was BMP and activin membrane-bound inhibitor (*Bambi*), an inhibitor of transforming growth factor-  (TGF- ) signaling. The *CDX2* protein was shown to bind to and activate the transcription of the human *BAMBI* promoter. Moreover, in a leukemic cell line established from *CDX2*-expressing mice, reducing the levels of *CDX2* or *Bambi* stimulated the TGF- -dependent expression of *Cd11b*, a marker of monocyte maturation. Taken together, this work demonstrates the strong oncogenic potential of the homeobox gene *CDX2* in the hematopoietic lineage, in contrast with its physiological tumor suppressor activity exerted in the gut. It also reveals, through BAMBI and TGF-  signaling, the involvement of *CDX2* in the perturbation of the interactions between leukemia cells and their microenvironment.

1. Introduction

Major developmental genes, such as caudal-related homeobox gene-2 (*CDX2*) [1], have emerged beyond ontogenesis as critical players in cancer. This homeobox gene has multiple functions during embryonic development including trophoblast formation, elongation, and patterning of the posterior body, and intestinal specification, before being selectively expressed in the gut epithelium throughout adulthood where it exerts a tumor suppressor role [2–6]. Conversely, it is ectopically

turned on in precancerous metaplastic lesions of the foregut [1] and, beyond the digestive tract, in a high proportion of acute leukemia associated with poor prognosis (see Ref [7] and references therein). Cellular studies have shown that *Cdx2* confers oncogenic properties to murine hematopoietic stem cells *in vitro* [8,9], while turning on its expression *in vivo* induces myelodysplastic lesions, a few of them evolving into overt leukemia [10]. However, the mechanism(s) associated with this pro-oncogenic activity remain largely elusive. Here, we developed a mouse model of conditional

Abbreviations

CDX2, caudal-related homeobox gene-2; BAMBI, BMP and activin membrane-bound inhibitor; TGF- , transforming growth factor- .

ectopic expression of human CDX2 in the hematopoietic lineage. This led to acute monoblastic-type leukemia involving the alteration of transforming growth factor- β (TGF- β) signaling. Interestingly, TGF- β is an important pathway in hematopoiesis, which is frequently altered in leukemogenesis [11]. Indeed, it negatively regulates cell proliferation and stimulates differentiation and apoptosis during normal hematopoiesis, whereas these effects are often impaired in hematologic malignancies due to deletions/mutations of members of the pathway or to the deregulation of cofactors by oncoproteins [12].

2. Materials and methods

2.1. Mice

RsCDX2 knockin mice having the human *CDX2* coding sequence preceded by a loxP-excisable transcriptional stop sequence inserted into the *Rosa26* locus [13], transgenic *Mx1Cre* mice containing the *Cre* coding sequence placed downstream of the promoter of the interferon-inducible gene *Mx1* [14] (Jackson Laboratory) and immunodeficient NSG mice (Charles River) were housed in the certified animal facility (#H-67-482-21). Protocols were approved by the French Ministry of Agriculture under the permit APA-FiS#833. Mice were genotyped by PCR using primers listed in Table S1. Three-month-old *RsCDX2::Mx1Cre* mice (designated as *MxCDX2*) and control littermates received 3 intraperitoneal injections of poly(I:C) (Sigma-Aldrich) at 10 mg kg⁻¹ of body weight every 2 days. For transplantation, 10⁵ bone marrow cells of *MxCDX2* or control mice were injected into the caudal vein of NSG mice.

2.2. Plasmids, siRNA, and luciferase assays

Plasmids pFlag (pFlag-CMV2; Sigma-Aldrich, Darmstadt, Germany), pFlag-CDX2 [15], pBAMBI-Luc [16], and pRL-null (Promega Inc., Charbonnières les Bains, France) have been described. siRNAs are listed in Table S1. Luciferase assays were performed using the Dual Reporter Luciferase Assay (Promega Inc.).

2.3. Cell line establishment, cell culture, transfection, and TGF- β treatment

Femoral bone marrow cells of *MxCDX2* mice were plated at 2 × 10⁶ cells per well in DMEM supplemented with 20% FBS, 14 ng·mL⁻¹ mIL3, 24 ng·mL⁻¹ mIL6, 112 ng·mL⁻¹ mSCF (Promokines, Camon, France), and

antibiotics for 1 week. The cell line AGK463 was established by progressive starvation of the three cytokines and reduction in the serum to 10%. Transfections used 10⁶ AGK463 cells, 50 nM siRNA alone, or 40 nM of these siRNAs with 10 nM siGLO RISK-free Control (Horizon Perkin-Helmer, Waterbeach, UK) and Viromer® BLUE Kit (Lipocalix, Halle, Sachsen-Anhalt, Germany). TGF- β treatment (20 ng·mL⁻¹) was performed 24 h after transfection during 48 h before analyses.

Human K562 myeloid leukemia cells were cultured as described [17]. For RNA analyses, 10⁶ cells were nucleofected with 1.5 μ g pFlag or pFlag-CDX2 following the Amaxa protocol [18]. For luciferase assays, 1.5 × 10⁵ cells were transfected using Lipofectamine™ 3000 (Invitrogen, Carlsbad, CA, USA). Analyses were performed 48 h later.

2.4. Serial replating assays

Femoral bone marrow cells of *MxCDX2* mice and control littermates were plated at 2 × 10⁴ cells per well in 1.1 mL methylcellulose (MethoCult GF M3434; Stem Cell Technologies, Grenoble, France) and cultured for 10 days. Colonies were counted, and cells were serially replated under the same conditions.

2.5. Histology, cytology, and flow cytometry

Bones were decalcified with formic acid. Histology and immunohistochemistry with anti-CDX2 antibody (EPR2764Y; Abcam, Paris, France) were performed as described [4]. Cytotyping of bone marrow cells used May Grünwald–Giemsa staining (MGG).

Flow cytometry was performed using a Fortessa cytometer (BD Biosciences, Le Pont de Claix Cedex, France) with antibodies listed in Table S2.

2.6. RNA preparation, RT-qPCR, RNAsequencing, and data analyses

RNA preparations were performed using TRI Reagent (Euromedex, Souffelweyersheim, France). The quality of total RNA was assessed on RNAchip with a Bioanalyzer (Agilent).

For RT-qPCR experiments, RNA was extracted from the indicated cell lines and analyzed with the probes listed in Table S1.

For RNAseq experiments, RNA was extracted from the bone marrow of 6 littermate mice: 3 *RsCDX2::Mx1Cre* mice and as control 1 *RsCDX2*, 1 *Mx1Cre*, and 1 wild-type mouse. All 6 mice were treated with poly(I:C), and RNA was prepared 6 weeks later. RNAseq data analyses were performed as described

[13] using STAR (v2.5) [19] for sequence alignment against the reference mouse genome GRCm38.90, and HTSeq (v0.6.1) [20] and DESeq2 (v1.10.1) [21] for reads counting and normalization. Differentially expressed genes were selected based on $|\log_2(\text{fold change})| > 1$ and $P\text{-value} < 0.05$, corrected with the false discovery rate (FDR) multiple testing method. Gene Ontology enrichment was done with the one-tailed exact Fisher's test, and $P\text{-values}$ were corrected with FDR multiple testing. Functional clustering used DAVID [22].

Transcription factor-binding motif analysis used HOMER (v4.10.1) [23] in the promoter gene regions extending from 2000 bp upstream to +50 bp with respect to the transcription starting site(s). Enriched motifs were manually curated by comparing with the most up-to-date version of JASPAR [24] and TomTom from the MEME suite [25].

For point mutation analysis, RNAseq reads were mapped against the mouse reference genome GRCm38.90 with the gap-aware aligner STAR as described above. Multimapped reads were not output, and alignments at splicing junctions were refined with a second pass based on the annotation of the reference genome. Duplicated reads were then flagged with MarkDuplicates from the Picard toolbox (<http://broadinstitute.github.io/picard/>). The remaining misaligned reads were removed with the tool SplitNCigarReads from the toolkit GATK (v4.0.9.0) [26]. single nucleotide polymorphisms (SNPs) were called with the tool HaplotypeCaller from GATK [27]. Soft-clipped bases were discarded during SNP calling, and the threshold for emitting the SNPs was set to 20 (Phred scale). SNPs were filtered based on their quality with the tool VariantFiltration from GATK with default parameters. SNPs were finally annotated with VEP from Ensembl [28]. Recurrent mutations were defined as SNPs present in all three *MxCDX2* mice but in none of the three control littermates. Mutational signatures were tested with SIGNAL (<https://signal.mutationsignatures.com/>) in comparison with published models of myeloid leukemia [29] and DNA repair mechanisms [30].

2.7. Western blots

Whole bone marrow or AGK463 protein extracts separated on SDS/PAGE were analyzed by western blot using mouse monoclonal anti-CDX2 (CDX2-88, BioGenex, Fremont, CA, USA; dilution 1/2000) and mouse monoclonal anti- β -actin (C4, Millipore, Dachstein, France; dilution 1/25 000) antibodies. HRP-conjugated secondary anti-mouse antibody was used for ECL detection (GE Company, Amersham, UK).

2.8. Chromatin immunoprecipitation

ChIP assays were performed using the Magna ChIP™ G Kit (Merck-Millipore, Dachstein, France) and anti-CDX2 antibody (EPR2764Y; Abcam) or IgG. qPCR amplification of the immunoprecipitated material used the primers listed in Table S1.

2.9. Statistics

Statistics used the two-tailed t -test for mean comparisons and log-rank test to compare the survival times between *MxCDX2* or engrafted NSG and control mice (GRAPHPAD, Prism, <https://www.graphpad.com/scientific-software/prism/>).

3. Results and Discussion

3.1. Oncogenic activity of CDX2 in the hematopoietic lineage *in vivo*

To address the oncogenic potential of ectopic expression of the human CDX2 homeoprotein in the hematopoietic lineage, *MxCDX2* mice were generated by intercrossing *RsCDX2* [13] and *Mx1Cre* [14] mice, and the resulting adult animals aged 2–3 months were treated with poly(I:C) to induce CDX2 expression. All the treated *MxCDX2* mice ($n = 18$) became moribund with a median survival time of 115 days, in contrast to treated control *RsCDX2* or *Mx1Cre* mice ($n = 13$; Fig. 1A). Comparing *MxCDX2* mice to controls 6 weeks after poly(I:C) administration revealed a pale bone marrow with reduced cellularity (Fig. 1B) and a monomorphic aspect (Fig. 1C) characterized by the accumulation of blasts expressing the *CDX2* mRNA (Fig. 1D) and protein visualized by western blot and immunohistochemistry (Fig. 1E,F). In addition, CDX2-positive blasts invaded several organs leading to hepatomegaly and splenomegaly and in most cases thymus atrophy (Fig. 1G,H).

The properties of bone marrow cells were first investigated using the *ex vivo* replating assay. As shown in Fig. 2A, bone marrow cells of control mice ($n = 4$) loosed replating potential at the 3rd step. On the contrary, cells of *MxCDX2* mice ($n = 4$), while producing twice less colonies than controls at the 1st plating step in line with the reduced bone marrow cellularity, maintained their growing activity throughout the five successive plating steps of this study (Fig. 2A,B). The second line of characterization of bone marrow cells was based on transplantation assays. Noteworthy, all recipient mice transplanted with bone marrow cells of

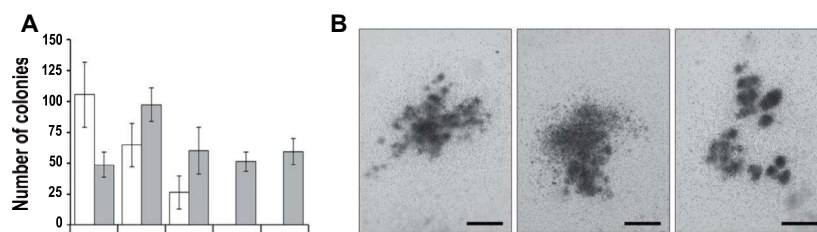
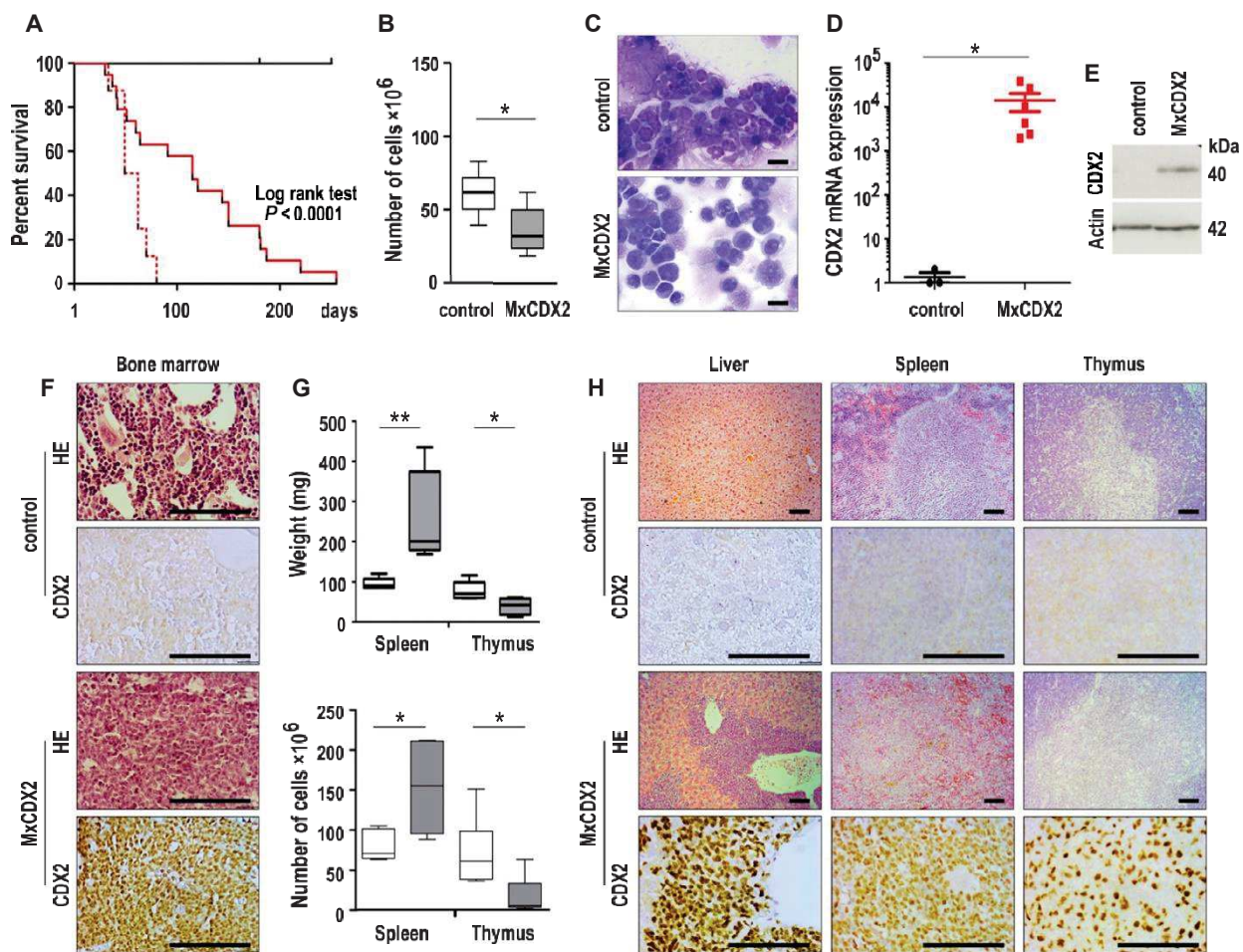


Fig. 2. Ex vivo replating activity of *MxCDX2* bone marrow cells.

(A) Number of colonies grown ex vivo from bone marrow cells of control (white boxes) and *MxCDX2* mice (gray boxes) during five consecutive passages (P1–P5). Data are given means with SEM. $n = 4$. (B) Representative pictures of colonies grown from *MxCDX2* bone marrow cells at the 5th replating step. Bars: 500 μm .

MxCDX2 mice ($n = 8$) developed leukemia within 80 days, in contrast to the recipients transplanted with cells of control mice, which remained healthy ($n = 5$; see Fig. 1A). Leukemia in the transplanted mice was similar to that of the donors, with the expansion of CDX2-positive blasts in the bone marrow, hepatomegaly, and splenomegaly linked to malignant cell invasion (not shown). Altogether, these data provide evidence that aberrant onset of *CDX2* in the hematopoietic lineage generates malignant blasts responsible for acute leukemia, transmissible by transplantation.

3.2. Myeloid vs lymphoid imbalance driven by ectopic CDX2

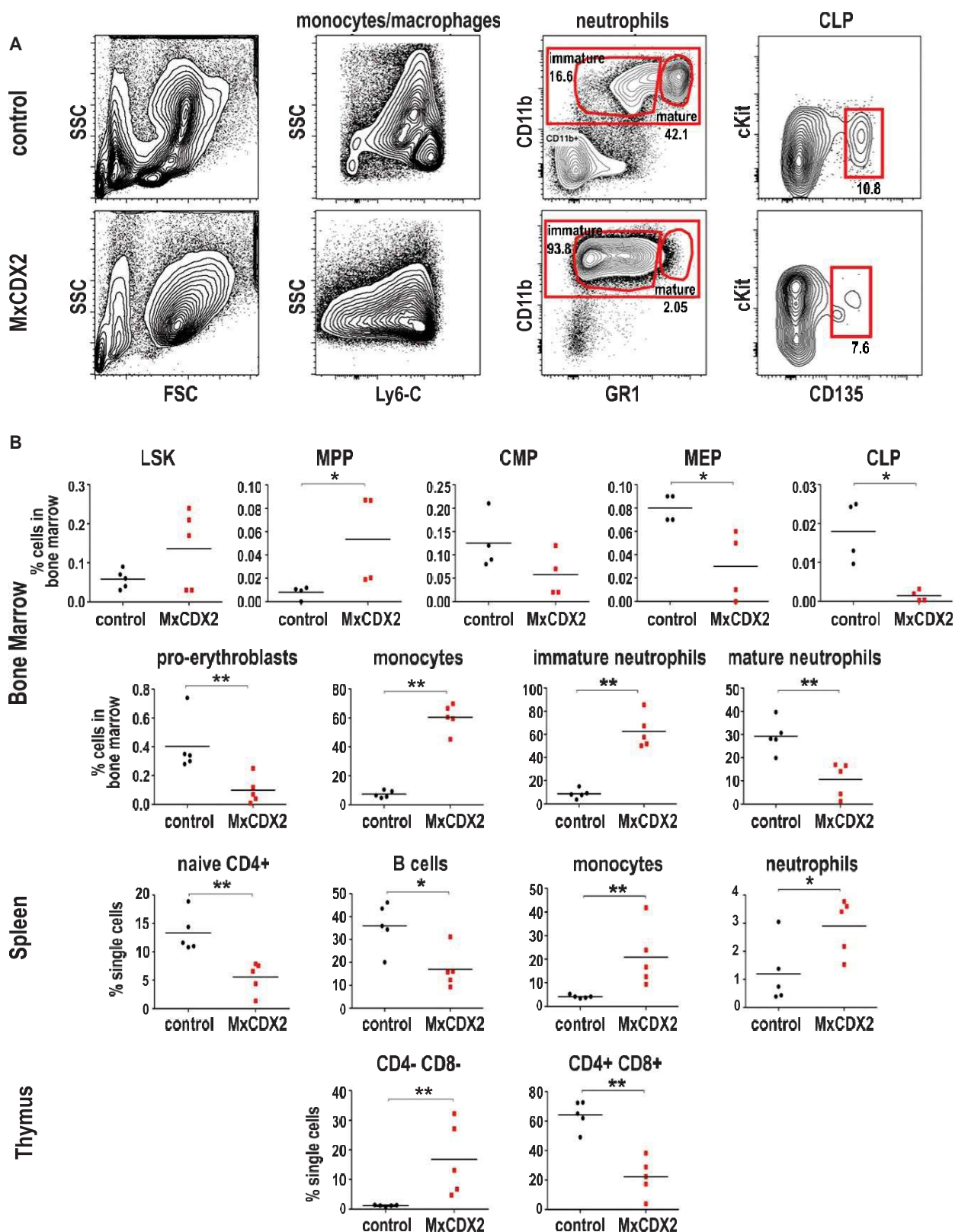
Six weeks after poly(I:C) administration, bone marrow smears from *MxCDX2* mice showed a large proportion of monoblasts, evocative of acute monoblastic leukemia. Flow cytometry and immunophenotyping highlighted the strong imbalance among bone marrow cell populations (Fig. 3A,B), characterized by the expansion of monocytic and granulocytic cells with a nonsignificant increase in Lin Sca⁺cKit⁺ progenitors, a slight increase in multipotent progenitors, and a strong decay of common lymphoid progenitors. In the myeloid lineage, megakaryocyte–erythroid progenitors dropped, leading to a drastic decrease in erythroid cells, reminiscent of the pale color of the bone marrow. On the contrary, monocyte–macrophage and immature neutrophil populations largely expanded. In the thymus, inverse changes in CD4 CD8 and CD4⁺CD8⁺ populations revealed a block in T-cell differentiation. In the spleen, T and B cells decreased, whereas monocyte–macrophage and neutrophil populations increased. These results showed that CDX2 blocks lymphoid commitment and erythroid differentiation to the benefit of monocyte and granulocyte engagement, yet with a defect in complete differentiation.

Transcriptome analyses performed 6 weeks after poly(I:C) administration revealed 2937 differentially expressed genes in the bone marrow of *MxCDX2* compared with control mice ($|\log_2(\text{fold change})| > 1$, $\text{adj}P\text{-value} < 0.05$; Fig. 4A; Table S3). Functional

annotation clustering showed enrichment in genes of the porphyrin biosynthetic pathway corroborating the drop of erythroid cells, as well as in genes involved in cell adhesion and signaling activities suggesting changes in the interactions of hematopoietic cells with their microenvironment (Fig. 4B). Differentially expressed genes also included 221 genes related to DNA binding and regulation of transcription, of which genes encoding important transcription factors for hematopoiesis and leukemogenesis (Table S4). Among them, transcription factor genes involved in monocyte emergence were upregulated (*Irf8*, *Klf4*, *Cebpa*, *Egr2*, *Jun*), whereas those of the erythroid (*E2f4*, *Gata1*, *Gata2*, *Gfi1b*, *Hoxa7*, *Hoxa9*, *Hoxa10*, *Klf1*, *Meis1*, *Pbx1*, *Sox6*, *Tal1*, *Zfpml1*) and lymphoid pathways (*Ets1*, *Etv5*, *Pax5*, *Pbx1*) were downregulated, consistent with the cellular phenotype. 1010 of the 2937 deregulated genes exhibited one or several consensus CDX DNA-binding site(s) in their 2000 to +50 bp promoters, including *Irf8* and *Jun* involved in monopoiesis [31,32]. These binding sites for CDX2 integrate into a complex pattern of deregulated genes encoding transcription factor involved in hematopoiesis and the presence of putative DNA-binding sites in their gene promoters, suggesting a cascade of deregulations initiated by CDX2 (Fig. 4C, Table S5). Interestingly, deregulated genes also included several genes related to DNA damage and repair (*Btg2*, *Ccnd1*, *Ebfl*, *Foxo3*, *Hmga2*, *Mgmt*, *Rad23a*, *Rb1*, *Xrcc5*), suggesting genome instability. This correlated with the occurrence of a number of nucleotide changes in the transcribed sequences of *MxCDX2* mice, of which 1001 recurrent mutations present in all three mutant mice but in none of the three control littermates (Table S6). 64.6% of these changes were transitions and 26.7% transversions, the others being microdeletions or microinsertions involving < 10 nucleotides. The predicted molecular effects are given in Table S7 and summarized in Fig. S1. Mutational signature analysis best fitted with the models A (35%), B and C (20%), and D (15%) of myeloid leukemia [29]. Substitution profiling according to Ref. [30] also suggested possible contribution of defects in PMS2-, Exo1-, PMS1-, and UNG-related DNA repair mechanisms (respectively, 50%, 17%, 12%, and 12% correlation).

Fig. 3. Cellular consequences of the ectopic expression of *CDX2*.

(A) Example of cytometry pattern in one control and one *MxCDX2* mouse. Monocytes–macrophages are Cd11b⁺ Gr1^{low} Ly6C^{and} SSC^{low}. Immature neutrophils are Cd11b⁺ Gr1^{low}; mature neutrophils are Cd11b⁺ Gr1^{high}. Percentages for neutrophils are given compared with single cells, and for CLP compared with Lin II7R Sca1^{low} cKit^{low} cells. In the illustrated examples, CLP represented 0.0243% and 0.000339% of the bone marrow single cells of control and *MxCDX2* mice, respectively. (B) Percent of the indicated cell types in the bone marrow, spleen, and thymus of control and *MxCDX2* mice. CLPs: common lymphoid progenitors; CMPs, common myeloid progenitors; LSK, Lin Sca⁺ cKit⁺ progenitors; MEPs, megakaryocyte–erythroid progenitors; MPPs, multipotent progenitors. Bars represent means, * $P < 0.05$; ** $P < 0.01$.



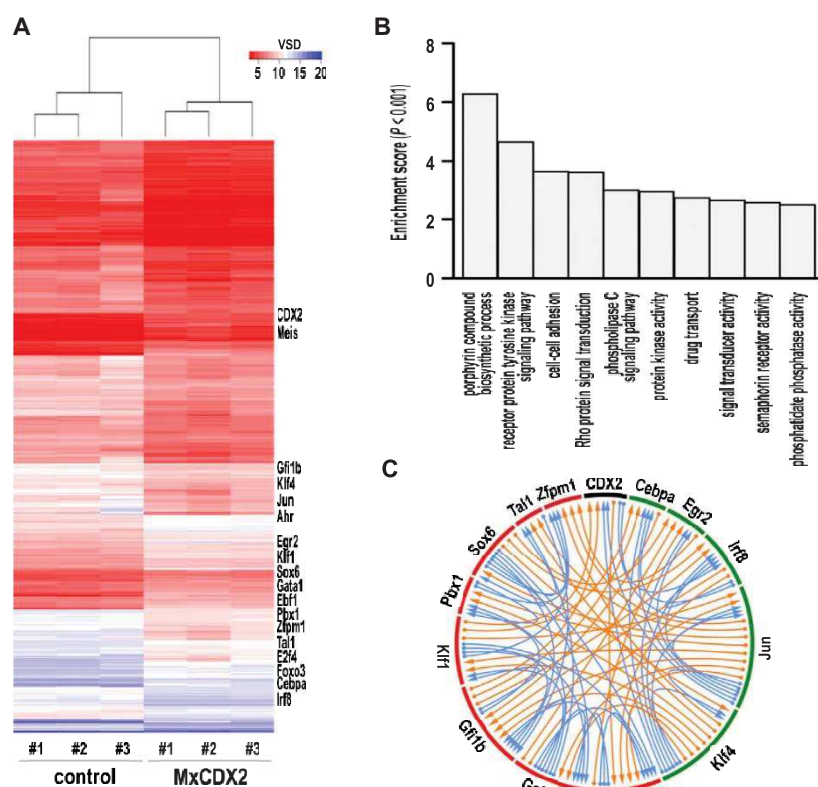
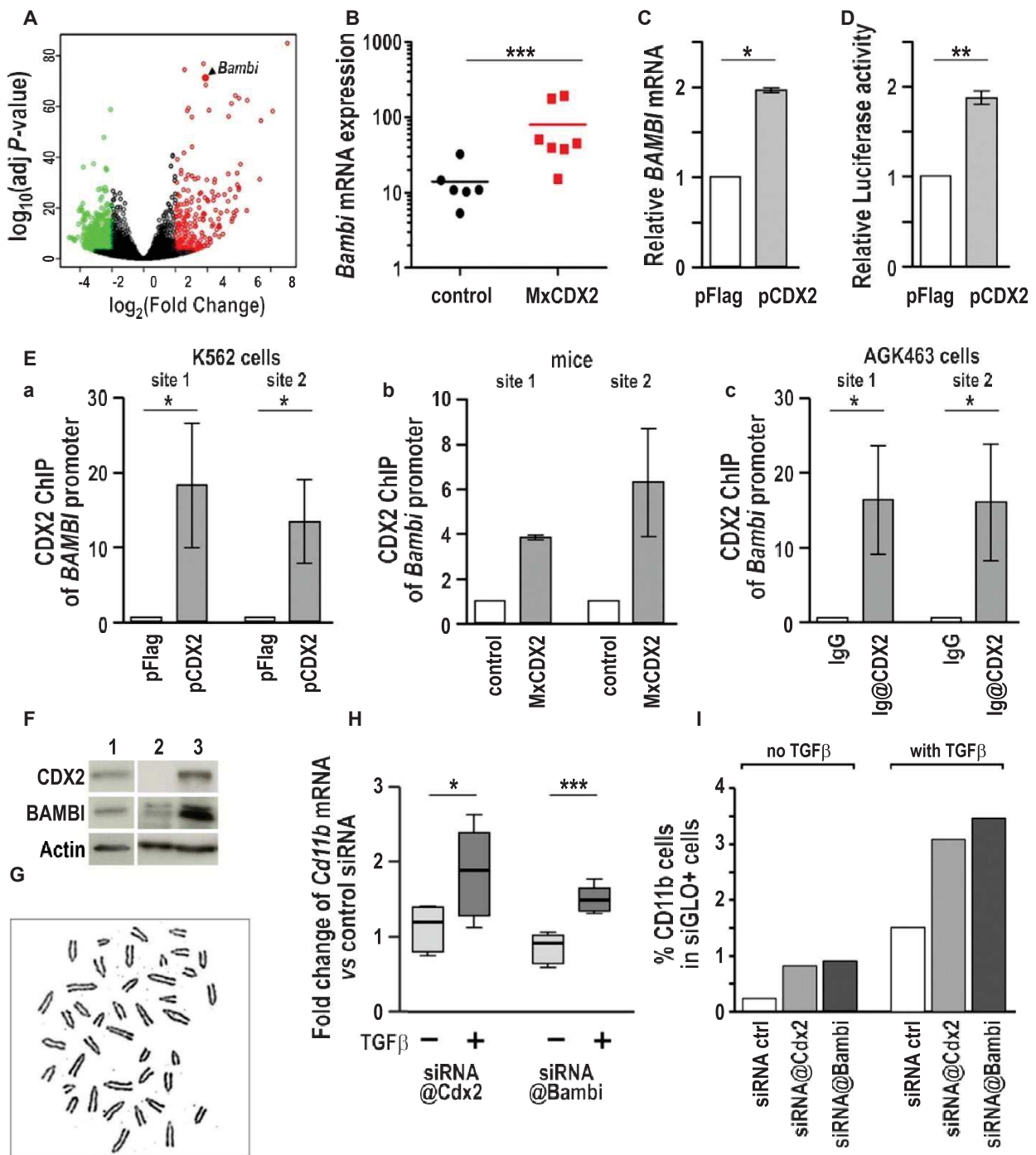


Fig. 4. Molecular consequences of the ectopic expression of *CDX2*.

(A) Heat map of the differentially expressed genes in the bone marrow of three control and three *MxCDX2* mice. The position of the *CDX2* gene and of major genes involved in hematopoiesis is indicated. (B) Gene Ontology term enrichment of the differentially expressed genes. (C) Circos plot of the relationship between the *CDX2* protein or several transcription factors involved in hematopoiesis, and the presence of consensus DNA-binding sites for these factors in their +200 to +50 bp gene promoter. Green boxes and red boxes represent transcription factors whose genes are, respectively, upregulated and downregulated by ectopic expression of *CDX2*. Arrows connect a given transcription factor to the presence of consensus DNA-binding site(s) for this factor in the promoters of transcription factor genes involved in hematopoiesis. Arrows are, respectively, in blue or orange when the factor and its target change in the same way or in opposite ways by ectopic expression of *CDX2*. A table version of this graph is given in Table S5.

Fig. 5. Involvement of *BAMBI* in the oncogenic effect of *CDX2*.

(A) Volcano plot of the differentially expressed genes in the bone marrow of *MxCDX2* compared with control littermates. Red: upregulated genes; green: downregulated genes. The position of *Bambi* is indicated. (B) Comparative expression of the *Bambi* gene by RT-qPCR in *MxCDX2* ($n = 7$) and control mice ($n = 6$). Bars represent means, $***P < 0.001$. (C) Expression of the *BAMBI* gene in human K562 cells transfected with the pFlag-*CDX2* plasmid compared to cells transfected with the control pFlag vector. $n = 3$; data are given means with SEM. $*P < 0.02$. (D) Luciferase activity produced by pBAMBI-Luc in human K562 cells transfected with pFlag-*CDX2* or with pFlag. $n = 9$; data are given means with SEM. $**P < 0.01$. (E) *CDX2* chromatin immunoprecipitation of *BAMBI* promoter fragments overlapping the *CDX*-type DNA-binding sites 1 and 2 (a) in human K562 cells transfected with pFlag-*CDX2* vs pFlag ($n = 5$), (b) in bone marrow cells of *MxCDX2* vs control mice ($n = 2$), and (c) in mouse AGK463 nuclear extracts immunoprecipitated with anti-*CDX2* antibodies or with control IgG ($n = 5$). Data are given means with SEM. $*P < 0.05$. (F) Expression of the *CDX2* and *BAMBI* proteins by western blot in AGK463 cells (lane 1) in comparison with the bone marrow of control and *MxCDX2* mice (respectively, lanes 2 and 3). (G) Karyotype analysis revealing supernumerary chromosomes ($n = 41$) in the AGK463 cell line. (H) Relative *Cd11b* mRNA expression in AGK463 cells transfected with siRNA@*CDX2* or siRNA@*Bambi* compared with control siRNA, in the absence or presence of TGF- β ($n = 9$). Boxes extend from 25th to 75th percentile, and whiskers represent mean to max. $*P < 0.05$; $***P < 0.001$. (I) Proportion of *CD11b*⁺ cells among siGLO⁺ AGK463 cells transfected with siGLO and the indicated siRNA, and treated or not with TGF- β .



The nucleotide changes fell into ~ 400 genes among which 20 are linked to the GO terms hematopoiesis or leukemia: *Chuk*, *Dleu2*, *Gab3*, *Hdac7*, *Hdac9*, *Hectd1*, *Kcnab2*, *Lrrk1*, *Picalm*, *Pkn1*, *Ppargc1b*, *Prdm16*, *Prdx3*, *Ptpnc*, *Rpl22*, *Stap1*, *Trpm2*, *Vps33b*, *Wdr1*, and *Zfp950*.

3.3. CDX2 interferes with TGF- β signaling to perturb the maturation of monoblasts

Although ectopic CDX2 is expected to alter numerous cellular and molecular functions, RNAseq analysis identified BMP and activin membrane-bound

inhibitor (*Bambi*) among the genes with the strongest stimulation in the leukemia cells of *MxCDX2* mice (Fig. 5A). It was confirmed by RT-qPCR in an independent series of control ($n = 6$) and *MxCDX2* ($n = 7$) animals (Fig. 5B). The BAMBI protein is a nonsignaling pseudoreceptor of the TGFBR family that represses the TGF- β pathway [33]. To investigate whether the overexpression of BAMBI may contribute to the malignant phenotype in *MxCDX2* mice, we first investigated the direct effect of CDX2 on the expression of the *Bambi* gene. In human acute myeloid leukemia cells K562, CDX2 overexpression by transfection with the plasmid pFlag-CDX2 increased the level of endogenous *BAMBI* mRNA (Fig. 5C) and also the activity of the [3384/+82] human *BAMBI* gene promoter [16], as shown in luciferase reporter assays (Fig. 5D). *In silico* analysis revealed two consensus CDX-type binding sites within a 300 bp segment conserved between human and mice (80% sequence identity) and located, respectively, 560 and 805 bp upstream of the human and mouse *BAMBI* transcription start site. Chromatin immunoprecipitation (ChIP) with anti-CDX2 antibody demonstrated the occupancy of both CDX-type sites by the CDX2 protein in pFlag-CDX2-transfected K562 cells unlike cells transfected with the control plasmid pFlag (Fig. 5Ea). Moreover, the corresponding sites in the mouse *Bambi* gene promoter were also occupied by CDX2 *in vivo* in bone marrow cells of *MxCDX2* compared with control littermates (Fig. 5Eb).

Next, advantage was taken from the replating potential of *MxCDX2* bone marrow cells to derive a cell line, named AGK463. This cell line expressed the CDX2 and BAMBI proteins (Fig. 5F), exhibited a monoblastic phenotype, and showed aneuploidy ($n = 41$ chromosomes, Fig. 5G), in line with the genome instability mentioned above in bone marrow cells of *MxCDX2* mice. Anti-CDX2 ChIP also demonstrated the binding of the CDX2 protein on both CDX-type binding sites of the *Bambi* promoter in AGK463 cells (see Fig. 5Ec). These cells were used to investigate the effect of CDX2 and BAMBI on their differentiation potential in response to TGF- β . For this purpose, cells were transfected with control siRNA or with siRNA@CDX2 or siRNA@Bambi, which decreased by twofold the corresponding mRNA levels. As shown in Fig. 5H, in the absence of TGF- β , the level of *Cd11b* mRNA was poorly affected in cells transfected with either siRNA@CDX2 or siRNA@Bambi compared with control siRNA. On the contrary, when TGF- β was added in the culture medium, decreasing the level of *CDX2* transcript with

siRNA@CDX2 stimulated *Cd11b* expression. Moreover, *Cd11b* expression was also stimulated upon TGF- β signaling activation in the CDX2-expressing AGK463 cells transfected with siRNA@Bambi. By cytometry, the proportion of AGK463 cells expressing CD11b at the membrane increased by transfection with siRNA@CDX2 or siRNA@Bambi, and the increase was further strengthened after TGF- β treatment (Fig. 5I). Thus, one of the effects of CDX2 is to jeopardize the TGF- β -dependent maturation of monocytes through its stimulatory effect of the expression of the TGF- β signaling inhibitor BAMBI. BAMBI represents a mediator of the oncogenic potential of CDX2 in hematopoietic cells by compromising the impact of TGF- β on monocyte differentiation.

4. Conclusion

In conclusion, this study shows that the ectopic expression of *CDX2* in the hematopoietic lineage triggers acute leukemia associated with genome instability and profound changes in gene expression patterns. The malignant phenotype observed here is more drastic and penetrant than in a recent model leading predominantly to myelodysplasia by targeting *CDX2* selectively in hematopoietic stem cells [10], likely because the induction of *CDX2* driven by the *Mx1Cre* allele affects more cells of the hematopoietic lineage than only stem cells. Previous studies have shown that the proliferation of acute myeloid leukemia cells is compromised when *CDX2* is knocked down, and, conversely, that inducing *Cdx2* *ex vivo* in hematopoietic progenitors increases self-renewal [8]. The present data are in line with these observations, and they further show that *CDX2* also introduces a bias in hematopoiesis in favor of the myeloid lineage, while compromising the maturation of monocytes through its stimulatory effect on a negative regulator of the TGF- β signaling pathway, BAMBI. Thus, unlike its tumor suppressor activity at its physiological site of expression, the gut, the *CDX2* homeobox gene is oncogenic in the ectopic setting of the hematopoietic lineage. Although these puzzling effects are far from being completely understood, they highlight the context-dependent function of this gene. A context-dependent activity of *CDX2* has already been described regarding its nontranscriptional activity on DNA repair between colon cancer and leukemia cells [18]. Recent studies highlighting the role of *CDX2* in the maintenance of open chromatin domains also allow anticipating a context-dependent effect of this homeoprotein at the transcriptional level [34,35], since the ultimate effect of keeping open chromatin regions depends on the specific repertoire of DNA-binding factors able

to interact with these regions in each particular cell type. Besides, we have previously reported that the level of CDX2 has an impact on the cellular and molecular microenvironment in the intestinal epithelium [4], and the present study shows that it also has an effect on the way leukemia cells respond to an extracellular signaling molecule, namely TGF- β . Because of the complexity of TGF- β signaling with regard to the number of members of this pathway and to its connection with other pathways, further studies are needed to thoroughly explore how CDX2 interferes with them. Yet, these data underscore the role played by CDX2 as a major actor of the crosstalk with the extracellular microenvironment in both physiological and pathological situations.

Acknowledgements

This work was supported by Inserm, the Ligue contre le Cancer du Haut-Rhin 2017–18, and the Fondation ARC (PJA 20181208021). We thank Anne Schreck (Karlsruhe Institute of Technology) for technical support in RNAseq.

Data accessibility

RNAseq data have been deposited in the GEO database under the accession number GSE120487.

Author contributions

CD-D, ID, and J-NF conceived the study, analyzed the results, and wrote the manuscript. CD-D, KH, YG, and AG performed animal, cellular, and molecular studies. VG performed sequencing and transcriptomic data analysis. EM, CE, and MT helped in immunohistochemistry, cytometry, and serial replating experiments, respectively. LM, LV, and M-PC performed pathological evaluation of the bone marrow smears and organs at the Hematology Laboratory and at the Pathology Department of the University Hospital of Strasbourg.

Conflict of interest

The authors declare no conflict of interest.

References

- Chawengsaksophak K (2019) Cdx2 animal models reveal developmental origins of cancers. *Genes (Basel)* 10: 928.
- Aoki K, Kakizaki F, Sakashita H, Manabe T, Aoki M & Taketo MM (2011) Suppression of colonic polyposis by homeoprotein CDX2 through its nontranscriptional function that stabilizes p27Kip1. *Cancer Res* 71, 593–602.
- Bonhomme C, Duluc I, Martin E, Chawengsaksophak K, Chenard MP, Kedinger M, Beck F, Freund JN & Domon-Dell C (2003) The Cdx2 homeobox gene has a tumour suppressor function in the distal colon in addition to a homeotic role during gut development. *Gut* 52, 1465–1471.
- Balbinot C, Armant O, Elarouci N, Marisa L, Martin E, De Clara E, Onea A, Deschamps J, Beck F, Freund J-N *et al.* (2018) The Cdx2 homeobox gene suppresses intestinal tumorigenesis through non-cell-autonomous mechanisms. *J Exp Med* 215, 911–926.
- Sakamoto N, Feng Y, Stolfi C, Kurosu Y, Green M, Lin J, Green ME, Sentani K, Yasui W, McMahon M *et al.* (2017) BRAFV600E cooperates with CDX2 inactivation to promote serrated colorectal tumorigenesis. *Elife* 6, e20331.
- Hryniuk A, Grainger S, Savory JGA & Lohnes D (2014) Cdx1 and Cdx2 function as tumor suppressors. *JBiol Chem* 289, 33343–33354.
- Darvishi M, Mashati P & Khosravi A (2018) The clinical significance of CDX2 in leukemia: a new perspective for leukemia research. *Leuk Res* 72, 45–51.
- Scholl C, Bansal D, Dohner K, Eiwen K, Huntly BJ, Lee BH, Rucker FG, Schlenk RF, Bullinger L, Dohner H *et al.* (2007) The homeobox gene CDX2 is aberrantly expressed in most cases of acute myeloid leukemia and promotes leukemogenesis. *J Clin Invest* 117, 1037–1048.
- Thoene S, Rawat VP, Heilmeier B, Hoster E, Metzeler KH, Herold T, Hiddemann W, Gokbuget N, Hoelzer D, Bohlander SK *et al.* (2009) The homeobox gene CDX2 is aberrantly expressed and associated with an inferior prognosis in patients with acute lymphoblastic leukemia. *Leukemia* 23, 649–655.
- Vu T, Straube J, Porter A, Bywater M, Song A, Ling V, Cooper L, Pali G, Brueedigam C, Jacquelin S *et al.* (2020) Hematopoietic stem and progenitor cell-restricted Cdx2 expression induces transformation to myelodysplasia and acute leukemia. *Nat Commun* 15, 3021.
- Bataller A, Montalban-Bravo G, Soltysiak K & Garcia-Manero G (2019) The role of TGF β in hematopoiesis and myeloid disorders. *Leukemia* 3, 1076–1089.
- Blank U & Karlsson S (2015) TGF- β signaling in the control of hematopoietic stem cells. *Blood* 125, 3542–3550.
- Grall E, Gourain V, Naïr A, Martin E, Birling M-C, Freund J-N & Duluc I (2019) Severe head dysgenesis resulting from imbalance between anterior and posterior ontogenetic programs. *Cell Death Dis* 10, 812.
- Kuhn R, Schwenk F, Aguet M & Rajewsky K (1995) Inducible gene targeting in mice. *Science* 269, 1427–1429.

15. Balbinot C, Vanier M, Armant O, Nair A, Penichon J, Soret C, Martin E, Saandi T, Reimund J-M, Deschamps J *et al.* (2017) Fine-tuning and autoregulation of the intestinal determinant and tumor suppressor homeobox gene CDX2 by alternative splicing. *Cell Death Differ* 24, 2173–2186.
16. Sekiya T, Adachi S, Kohu K, Yamada T, Higuchi O, Furukawa Y, Nakamura Y, Nakamura T, Tashiro K, Kuhara S *et al.* (2004) Identification of BMP and activin membrane-bound inhibitor (BAMBI), an inhibitor of transforming growth factor-beta signaling, as a target of the beta-catenin pathway in colorectal tumor cells. *J Biol Chem* 279, 6840–6846.
17. Luzzio CB & Luzzio BB (1975) Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood* 45, 321–334.
18. Renouf B, Soret C, Saandi T, Delalande F, Martin E, Vanier M, Duluc I, Gross I, Freund J-N & Domon-Dell C (2012) Cdx2 homeoprotein inhibits non-homologous end joining in colon cancer but not in leukemia cells. *Nucleic Acids Res* 40, 3456–3469.
19. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M & Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
20. Anders S, Pyl PT & Huber W (2015) HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169.
21. Love MI, Huber W & Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550.
22. Huang DW, Sherman BT & Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4, 44–57.
23. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H & Glass CK (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38, 576–589.
24. Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, Bessy A, Cheneby J, Kulkarni SR, Tan G *et al.* (2018) JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res* 46, D260–D266.
25. Bailey T, Johnson J, Grant C & Noble W (2015) The MEME suite. *Nucleic Acids Res* 43, W39–W49.
26. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J *et al.* (2013) From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 43, 11.10.1–11.10.33.
27. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D *et al.* (2017) Scaling accurate genetic variant discovery to tens of thousands of samples. *bioRxiv* [PREPRINT].
28. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P & Cunningham F (2016) The ensembl variant effect predictor. *Genome Biol* 17, 122.
29. Degasperis A, Amarante TD, Czarnecki J, Shooter S, Zou X, Glodzik D, Morganella S, Nanda AS, Badja C, Koh G *et al.* (2020) A practical framework and online tool for mutational signature analyses show inter-tissue variation and driver dependencies. *Nat Cancer* 1, 249–263.
30. Zou X, Koh GCC, Nanda AS, Degasperis A, Urgo K, Roumeliotis TI, Agu CA, Side L, Brice G, Perez-Alonso V *et al.* (2020) Dissecting mutational mechanisms underpinning signatures caused by replication errors and endogenous DNA damage. *bioRxiv* [PREPRINT].
31. Friedman AD (2007) C/EBP α Induces PU.1 and interacts with AP-1 and NF- κ B to regulate myeloid development. *Blood Cells Mol Dis* 39, 340–343.
32. Kurotaki D, Osato N, Nishiyama A, Yamamoto M, Ban T, Sato H, Nakabayashi J, Umehara M, Miyake N, Matsumoto N *et al.* (2013) Essential role of the IRF8-KLF4 transcription factor cascade in murine monocyte differentiation. *Blood* 121, 1839–1849.
33. Onichtchouk D, Chen Y, Dosch R, Gawantka V, Delius H, Massagué J & Niehrs C (1999) Silencing of TGF-beta signalling by the pseudoreceptor BAMBI. *Nature* 401, 480–485.
34. Saxena M, Roman AKS, O'Neill NK, Sulahian R, Jadhav U & Shivdasani RA (2017) Transcription factor-dependent “anti-repressive” mammalian enhancers exclude H3K27me3 from extended genomic domains. *Genes Dev* 31, 2391–2404.
35. Verzi MP, Shin H, San Roman AK, Liu XS & Shivdasani RA (2013) Intestinal master transcription factor CDX2 controls chromatin access for partner transcription factor binding. *Mol Cell Biol* 33, 281–292.

Supporting information

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

Fig. S1. Proportion of the predicted consequences of the recurrent mutational changes observed in MxCDX2 mice.

Table S1. Primers, probes and siRNA.

Table S2. List of antibodies used for flow cytometry.

Table S3. Differentially-expressed genes in the bone marrow of MxCDX2 mice compared to control littermates.

Table S4. Differentially-expressed genes for DNA binding/Transcription Regulatory Factors.

Table S5. Combined analysis of DNA-binding sites in promoter of hematopoietic transcription factors and their respective deregulated expression.

Table S6. Recurrent nucleotide changes in the transcribed sequences of *MxCDX2* mice compared to control littermates.

Table S7. Prediction of the consequences of the nucleotide changes in *MxCDX2* mice.

Partie 2

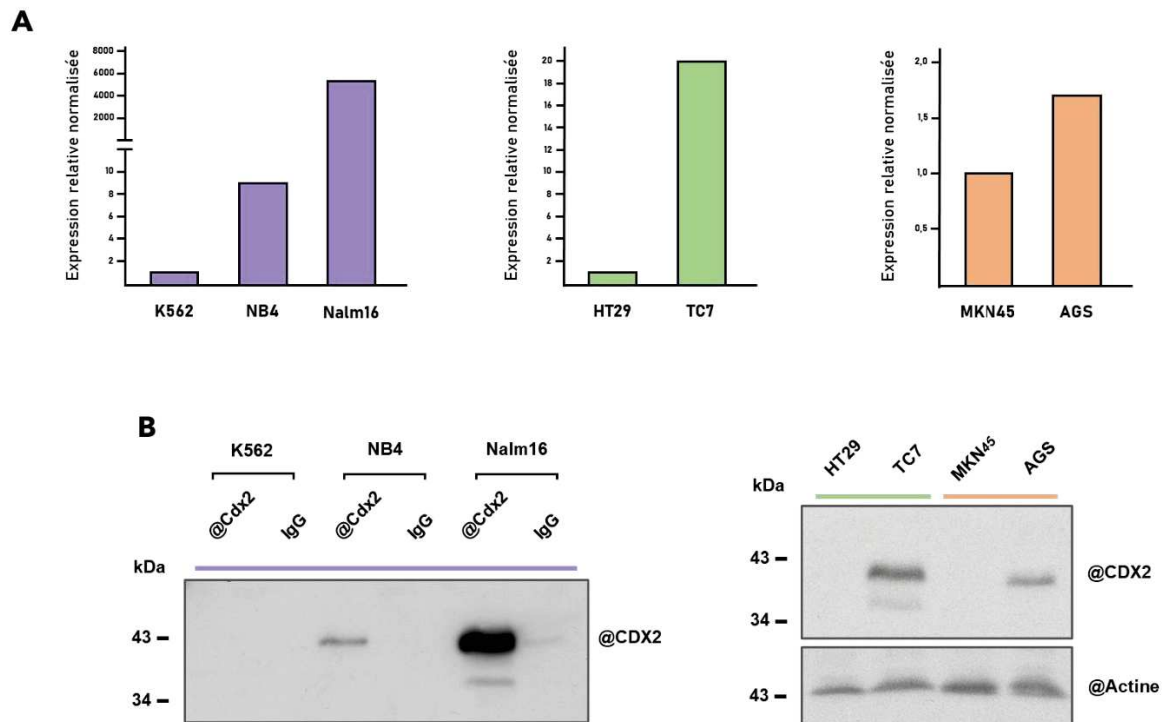


Fig. 6. Expression du gène CDX2

(A) Résultats de RT-qPCR ciblant l'expression du gène CDX2 dans les 3 lignées leucémiques (K562, NB4, Nalm-16), 2 lignées intestinales (HT29, TC7) et 2 lignées gastriques (MKN45, AGS) humaines. Normalisation avec le gène de ménage *TATA-Binding Protein* (TBP). (B) Résultats de Western Blot dans les mêmes lignées cellulaires humaines. Pour les lignées leucémiques, une immunoprécipitation à l'aide d'un anticorps @CDX2 à partir de 2mg de protéines totales a été réalisé. En contrôle, des IgG ont été utilisés pour l'immunoprécipitation.

L'expression du gène *CDX2* est dérégulée dans 3 contextes pathologiques éclectiques que sont les cancers colorectaux (CCR), les carcinomes gastriques et les leucémies aiguës^{291,329,361}. Pourtant, aucun changement de méthylation de son promoteur, d'amplification de son locus ou de mutations l'affectant n'ont été répertoriées de manière systématique chez les patients. Nous faisons donc l'hypothèse que des mécanismes transcriptionnels anormaux se mettent en place.

Pour comprendre les mécanismes de dérégulation mis en jeu, une étude globale de la régulation épigénétique du gène *CDX2* a été amorcée. Elle s'effectue dans 3 contextes en parallèle : des lignées cellulaires cancéreuses intestinales, compartiment dans lequel son expression est physiologique et diminuée dans les CCRs (Caco-2/TC-7, *CDX2*⁺ vs. HT-29, *CDX2*⁻), ainsi que des lignées gastriques (AGS, *CDX2*⁺ vs. MKN-45, *CDX2*⁻) et des lignées leucémiques (NB-4, *CDX2*⁺ vs. K-562, *CDX2*⁻), où son expression est à l'inverse aberrante. L'étude a consisté dans chacun de ces contextes en un travail comparatif entre deux lignées, exprimant ou non *CDX2*. Pour distinguer entre contextes myéloïde et lymphoïde, nous avons, en cours de route, intégré à cette étude une lignée leucémique humaine de LAL-B (Nalm-16, *CDX2*⁺).

La stratégie est en deux étapes : identifier des régions de chromatine active par ATAC-seq et ChIP-seq, puis parmi elles, celles qui se rapprochent physiquement de son promoteur proximal par capture de la conformation du chromosome (en l'occurrence le 4C). Nous avons pour idée de recouper ces 2 sets d'informations afin d'identifier les *enhancers* et *silencers* potentiels de *CDX2*. Peut-être en existe-t-il des spécifiques selon les contextes cellulaires ?

A. Validation de l'expression de *CDX2*

L'expression du gène *CDX2* a été vérifiée dans les 7 lignées cellulaires, au niveau transcriptionnel et protéique. De par son niveau d'expression faible dans la lignée leucémique NB4, une immunoprécipitation ciblant la protéine *CDX2* a été effectuée pour valider son expression dans les lignées leucémiques (**Fig. 6A&B**).

Nous avons ainsi pu confirmer la présence de l'ARNm et de la protéine CDX2 dans les lignées Caco2/TC7, AGS, NB-4 et Nalm-16, et l'absence d'expression ou une faible expression du gène CDX2 dans les lignées HT-29, MKN-45 et K-562. Ces résultats sont en accord avec la littérature.

B. Régions de chromatine active

1. Identification des régions de chromatine actives par ChIP-seq

La première étape de l'investigation a consisté à rechercher des régions de chromatine active, accessible aux facteurs de transcription. Des immunoprécipitations de chromatine, suivies d'un séquençage (ChIP-seq), ici par la technologie Illumina, ont permis d'identifier les régions chromosomiques portant les modifications d'histones activatrices : la triméthylation de la lysine 4 de l'histone H3 (H3K4me3), et l'acétylation de la lysine 27 de l'histone H3 (H3K27ac) (**Fig. 7A**). Lors de la mise au point du ChIP, différentes conditions de sonication ont été testées. Les profils de sonication obtenus nous ont permis d'opter pour un temps de sonication de 10min, pour l'ensemble des lignées (**Fig. 7B**).

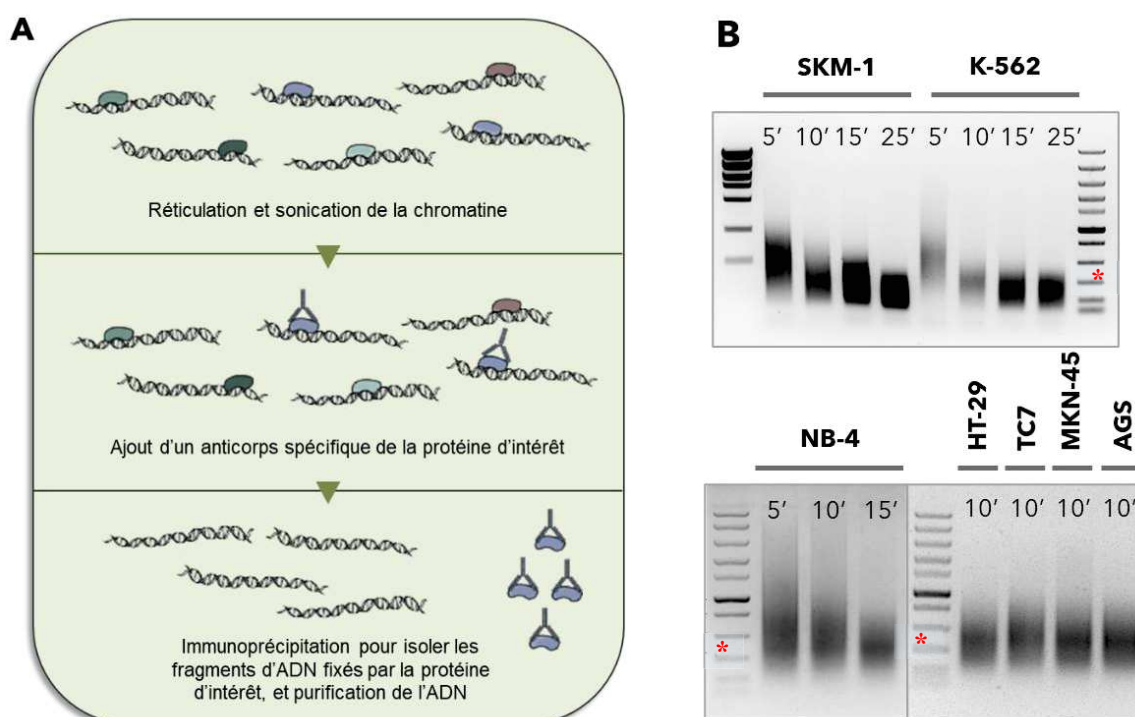


Fig. 7. ChIP-seq : principe de la préparation de la chromatine.

(A) Principe du ChIP-seq : La technique permet d'immunoprécipiter, à l'aide d'anticorps, une protéine donnée et les portions de chromatine soniquée avec lesquelles elle interagit. Les fragments récupérés sont ensuite séquencés par Illumina. (B) Profils obtenus après les temps de sonication indiqués pour les différentes lignées cellulaires. L'astérisque rouge indique la taille de 500 pb.

Préalablement au séquençage, un contrôle qualité des fragments immunoprécipités a été effectué, et a permis de confirmer que les fragments immunoprécipités correspondaient bien majoritairement à de la chromatine active. Nous observons par qPCR un enrichissement de 2 régions connues pour être transcrites (donc porteuses des modifications activatrices étudiées), que sont les gènes ACTB2 et GAPDH. Deux régions intergéniques non transcrites (Neg1/2) n'ont pas été retrouvées parmi les fragments immunoprécipités. (**Fig. 8**).

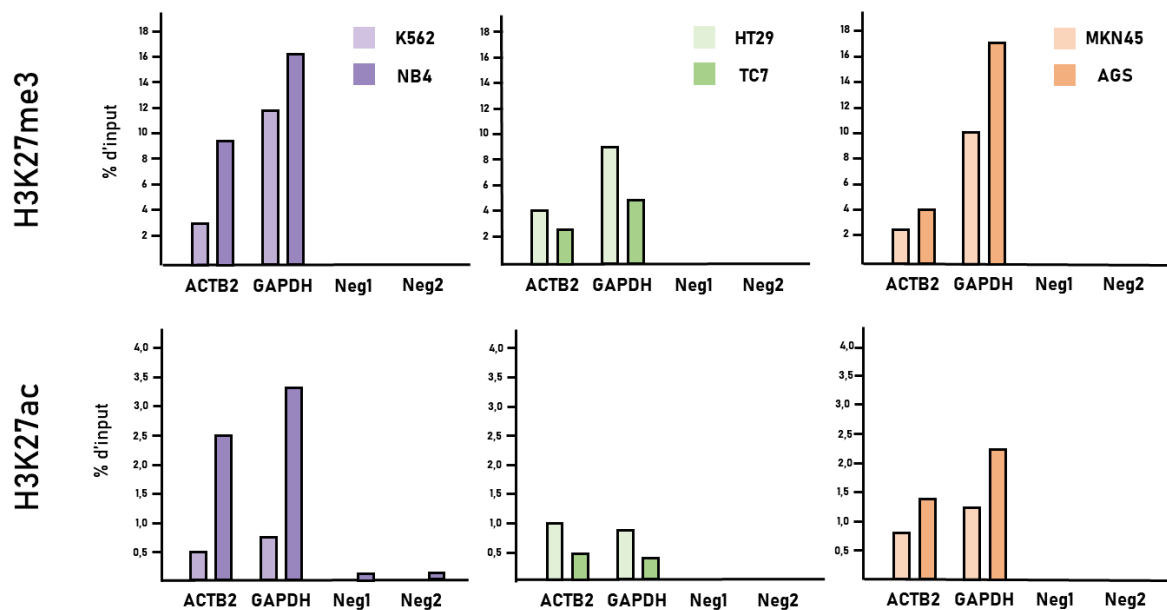


Fig. 8. ChIP-seq : PCR de contrôle des ChIP.

Contrôle par qPCR : vérification de l'enrichissement supérieur en chromatine active (gènes ACTB2 et GAPDH) vs chromatine inactive (Neg1/2 : régions intergéniques non transcrites). Les graphes représentent en ordonnée le % de fragments immunoprécipités par rapport à la quantité initiale pour les 4 régions chromosomiques choisies, les 6 lignées cellulaires et les 2 marques actives.

Les résultats du séquençage se présentent sous la forme de pics obtenus à partir des fichiers .wig de séquençage, grâce au logiciel Integrated Genome Browser (IGB), et correspondent à un enrichissement de la présence d'une région donnée dans l'ensemble des fragments immunoprécipités. A l'échelle du génome, ces résultats seuls ne sont pas très parlants puisqu'il s'agit de modifications d'histones très répandues. Des différences sont néanmoins observables entre les 6 conditions, au niveau du locus de *CDX2* (situé en 13q12.2). La lignée Caco2/TC7 présente un net enrichissement pour les 2 modifications d'histones activatrices, en comparaison avec la lignée HT-29. La lignée gastrique AGS présente davantage de marques H3K27ac que la lignée MKN-45. Ces 2 lignées gastriques présentent toutes deux un enrichissement en marques H3K4me3 (**Fig. 9**).

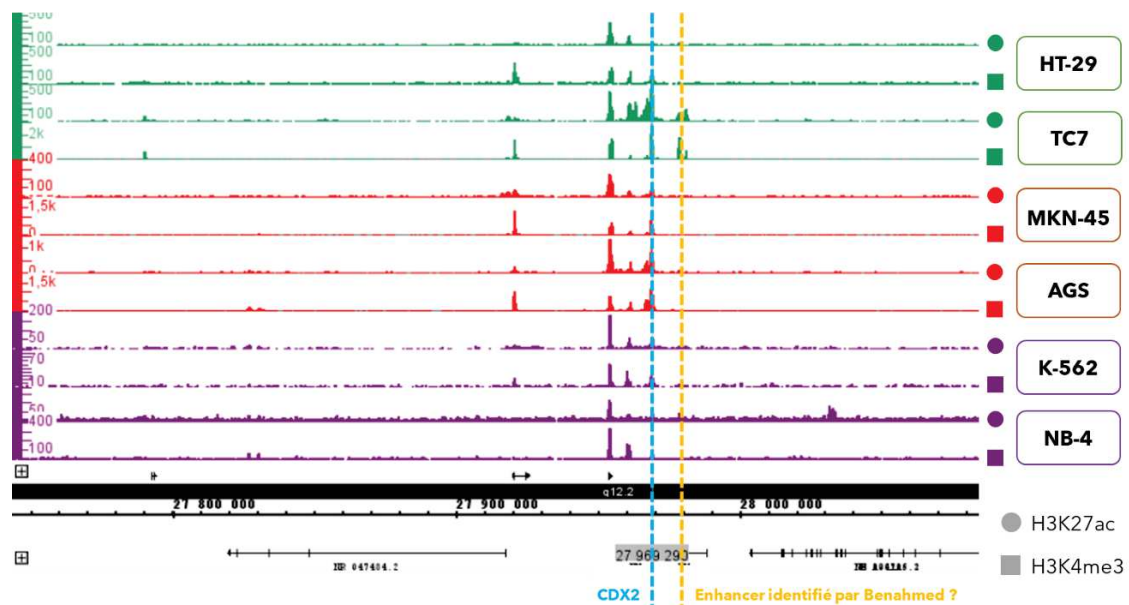


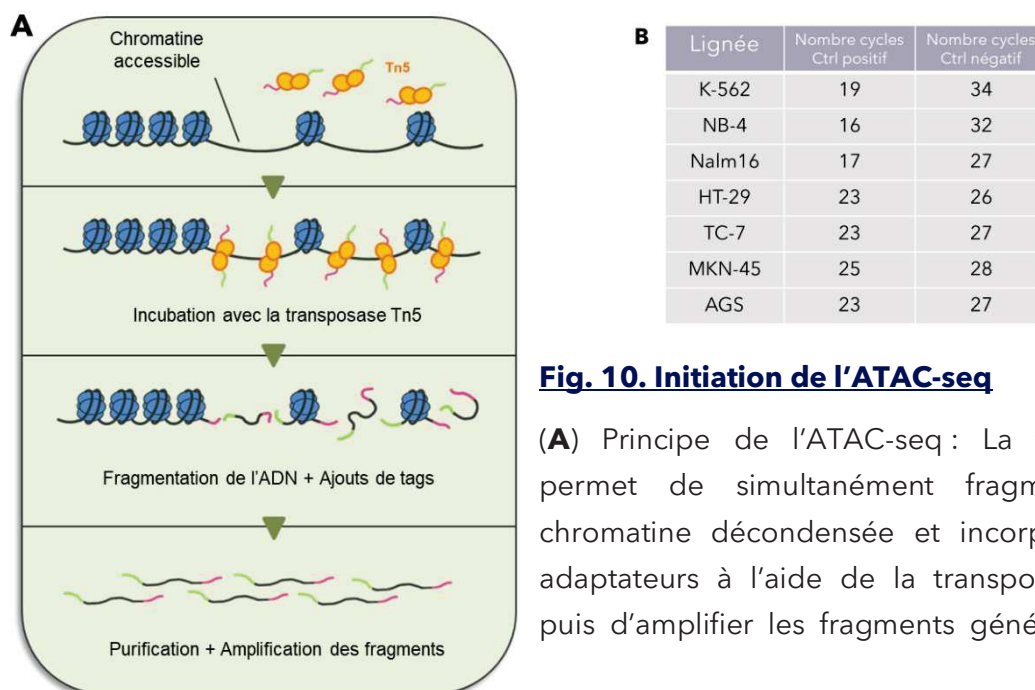
Fig. 9. Profils du ChIP-seq

Les résultats des ChIP-seq se représentent sous forme de pics, correspondant à un enrichissement en régions de chromatine portant une modification post-traductionnelle d'histone donnée. L'analyse est ici centrée sur le locus du gène *CDX2* (indiqué en bleu). Est indiqué en jaune la localisation du enhancer de *CDX2* identifié dans l'épithélium intestinal (Benahmed 2008). Les marques H3K27ac sont représentées par des ronds, et les marques H3K4me4 par des carrés. Leur couleur dépend du compartiment étudié : vert pour l'intestin, rouge pour l'estomac, violet pour le compartiment hématopoïétique.

Nous pouvons remarquer la présence d'un pic en amont du gène *CDX2*, au niveau d'une région située à environ 10000 pb du +1 de transcription. Ce enhancer pourrait correspondre à celui décrit par Benahmed et al., (2008), et n'est pas retrouvé dans les lignées gastriques ni dans les lignées leucémiques suggérant qu'un mécanisme de régulation impliquant cet enhancer serait spécifique de l'intestin. Au niveau des lignées leucémiques qui expriment *CDX2* (NB-4 et Nalm-16), nous observons également une décondensation de la chromatine au niveau du gène *FLT3* situé environ 35 000 (pour l'extrémité 3') à 130 000 (pour l'extrémité 5') pb en amont du gène *CDX2*. Une analyse bioinformatique plus détaillée est nécessaire pour déterminer de manière plus sensible les spécificités des profils.

2. Identification de la chromatine active par ATAC-seq

L'ATAC-seq (*Assay for Transposase-Accessible Chromatin + sequencing*) est une technique complémentaire du ChIP-seq pour étudier la chromatine active : elle consiste à incuber la chromatine avec la transposase Tn5. Cette enzyme accède uniquement aux régions décompactées de la chromatine et va la fragmenter. Des adaptateurs sont incorporés simultanément aux extrémités des fragments, permettant ainsi de les purifier, amplifier et séquencer (**Fig. 10A**). Avant le séquençage, nous avons contrôlé la qualité de la digestion par la Tn5, en testant la présence dans le pool de chromatine purifiée d'une région active (promoteur du gène Actine) et inactive (région intergénique) (**Fig. 10B**).



les séquencer. **(B)** La qPCR a été effectuée en recherchant la présence d'une région de chromatine active (Actine) et inactive (région intergénique) parmi l'ADN fragmenté et purifié. On estime qu'il faut au moins 2Ct de différence entre la région active et inactive pour considérer comme bonne la qualité des échantillons.

A l'issue du séquençage, les pics obtenus correspondent aux régions où la transposase a pu couper, et donc ici aussi aux régions où la chromatine est ouverte. (**Fig. 11**).

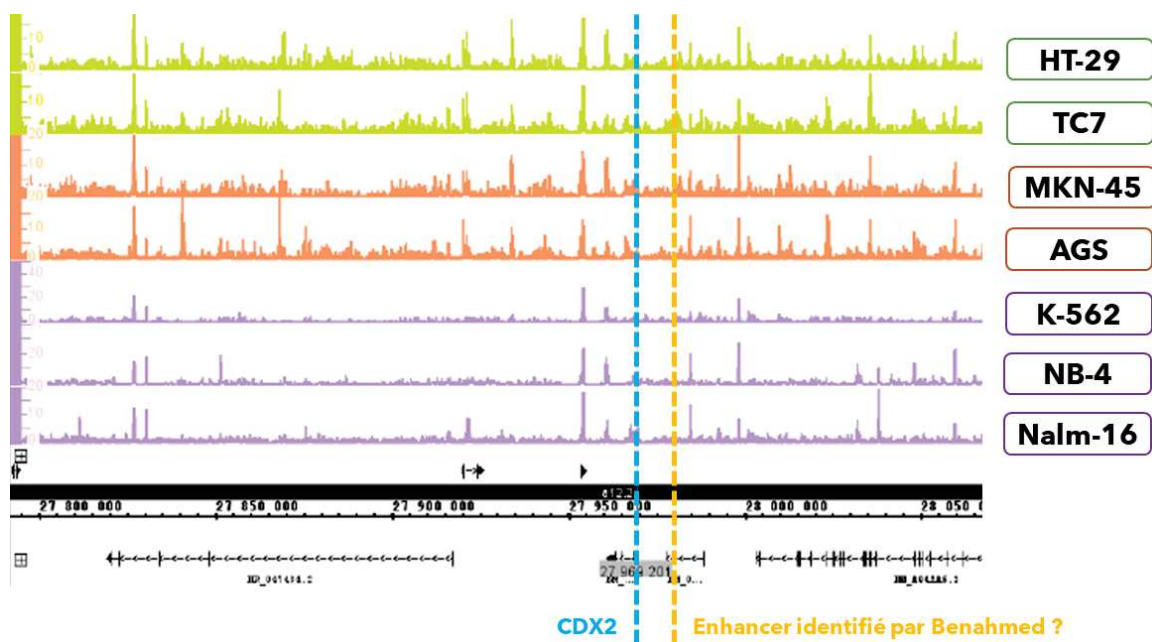


Fig. 11. Profils de l'ATAC-seq.

Les résultats de l'ATAC-seq se représentent sous forme de pics, correspondants aux régions enrichies où la transposase Tn5 a pu couper, et donc aux régions de chromatine accessibles aux facteurs de transcription.

Les profils ainsi obtenus à l'issue du ChIP-seq et de l'ATAC-seq ont été comparés. Pour chaque lignée, nous observons une bonne corrélation entre les deux profils (**Fig. 12**), indiquant que ces deux approches sont robustes pour mettre en évidence les régions actives de la chromatine.

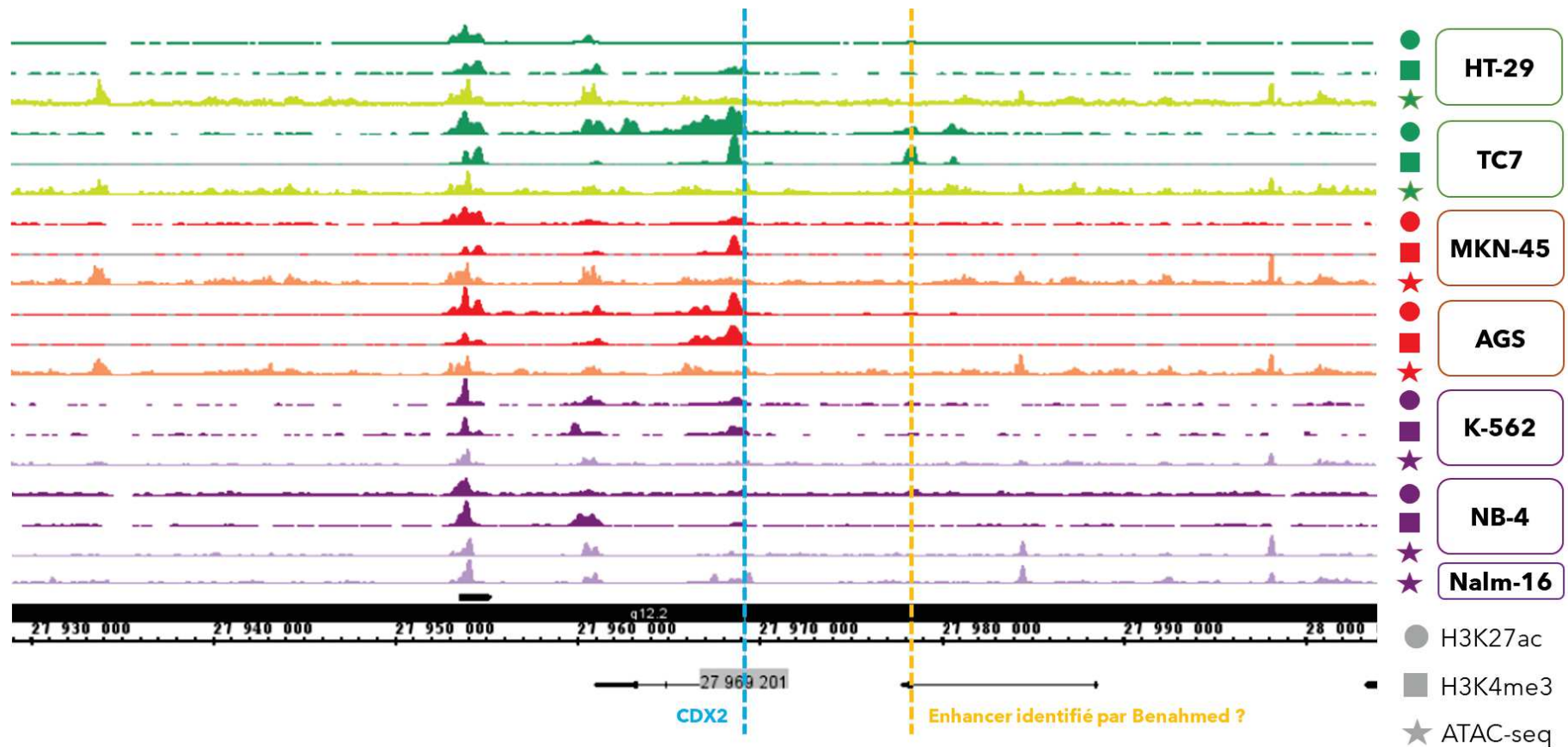


Fig. 12. Superposition des profils de ChIP-seq et ATAC-seq

Pour chaque lignée étudiée indiquée à droite, les profils obtenus par ChIP-seq détectant les régions enrichies en marques actives H3K27ac (rond) ou H3K4me3 (carré) et par ATAC-seq (étoile) autour du locus du gène *CDX2* ont été alignés. Pour les cellules Nalm-16, nous ne disposons que du profil d'ATAC-seq. La position du gène *CDX2* est indiquée par les pointillés bleu.

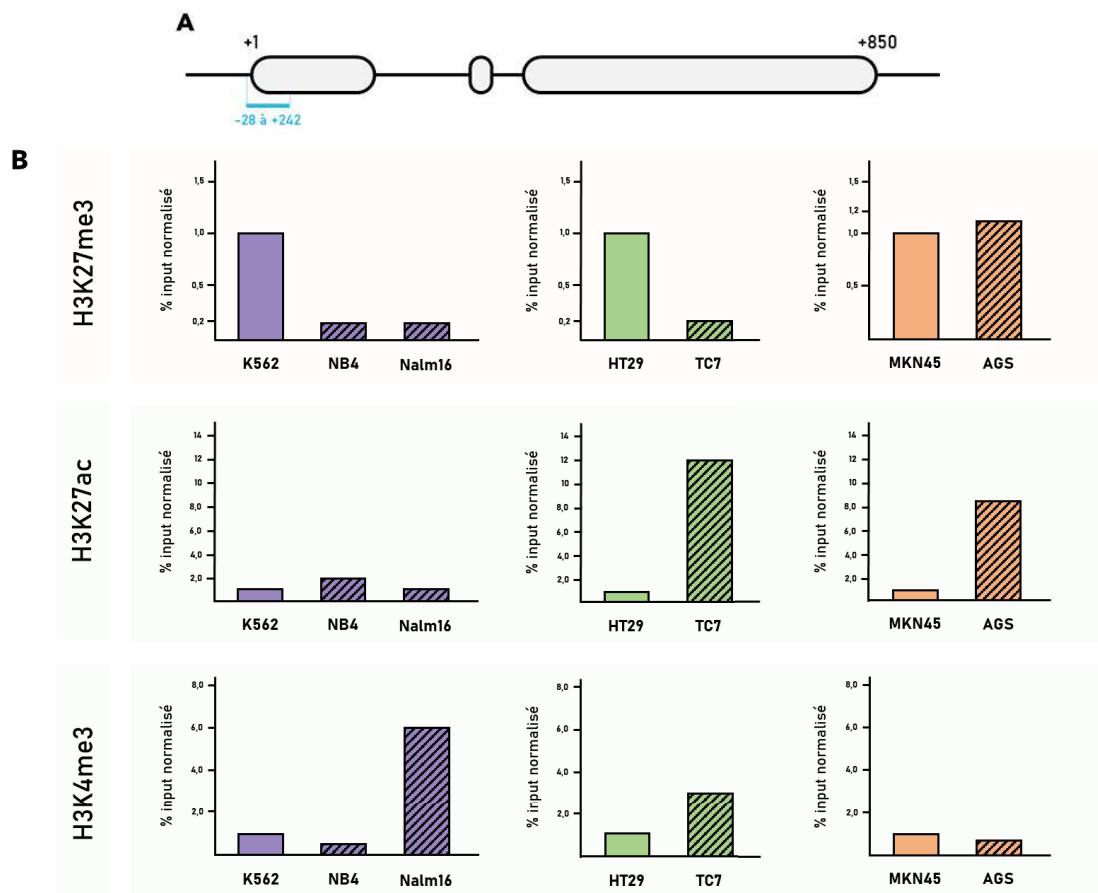
C. Bivalence du promoteur proximal du gène *CDX2*

CDX2 fait partie des gènes développementaux dotés de promoteurs bivalents, caractérisés par la présence de modifications d'histones activatrices et inactivatrices simultanément^{138,326}. Cette cohabitation permet un contrôle rapide et efficace de leur activation, leur permettant une réponse transcriptionnelle efficace. Pour ces gènes, l'expression est conditionnée par la perte de modifications répressives.

Notre étude s'est ainsi poursuivie par une analyse comparative quantitative de modifications d'histones, localisée au niveau du promoteur proximal de *CDX2*, dans les différentes lignées. Nous nous sommes intéressés aux 2 modifications activatrices précédentes (H3K27ac et H3K4me3), mais aussi à la modification inactivatrice très répandue, H3K27me3. L'approche a consisté en des ChIP ciblant ces 3 modifications d'histones, suivis de qPCR amplifiant spécifiquement une portion de la région 5' du gène *CDX2* (**Fig. 13A**). Au niveau du promoteur proximal du gène *CDX2*, une diminution des modifications inhibitrices de type H3K27me3 est effectivement observée dans la lignée intestinale exprimant *CDX2*, en même temps qu'une nette augmentation des marques activatrices (H3K27ac et H3K4me3). Les modifications H3K27me3 et H3K4me3 restent en moyenne inchangées dans les lignées gastriques. Le promoteur proximal de *CDX2* y est en revanche enrichi en modifications H3K27ac. Enfin, dans les lignées leucémiques, l'expression de *CDX2* corrèle avec l'amoindrissement de marques H3K27me3. La lignée NB4 ne présente pas d'enrichissement en modifications activatrices, tandis que la lignée Nalm-16, exprimant beaucoup plus fortement *CDX2*, démontre un fort taux de modifications activatrices H3K4me3 (**Fig. 13B**).

Fig. 13. Étude du promoteur proximal du gène *CDX2*

(**A**) Région du promoteur proximal du gène *CDX2* ciblée par la qPCR, d'après Sun et al., 2017. (**B**) Les résultats de ChIP-qPCR se présentent sous la forme du pourcentage de la portion de chromatine étudiée immunoprécipitée par rapport à la chromatine totale initiale (input) et est normalisée par rapport à la lignée considérée comme négative pour *CDX2*. Le ChIP-qPCR a été réalisé dans les 7 lignées cellulaires humaines K-562, NB-4 et Nalm-16 (leucémiques, HT29 et TC7 (intestinales) et MKN45 et AGS (gastriques). Les modifications H3K27me3 d'une part et H3K27ac et H3K4me3 d'autre part sont considérées respectivement comme inhibitrices et activatrices de la transcription.



D. Régions en interaction avec le promoteur proximal de *CDX2*

Comme mentionné en introduction, les enhancers sont des régions régulatrices situées à distance du site d'initiation de la transcription et qui interagissent par un repliement de la chromatine avec le promoteur proximal. La dernière étape à laquelle j'ai pu contribuer a consisté en la réalisation d'un 4C-seq (*Circularisation Chromosome Conformation Capture-on-Chip + sequencing*) sur les 7 lignées, afin d'identifier les régions de chromatine qui interagissent transitoirement avec le promoteur de *Cdx2*. Cette technique permet une détection non biaisée des régions chromosomiques qui interagissent avec une région donnée spécifique. Elle consiste à maintenir proche les 2 régions temporairement en interaction par cross-linking, puis à les ligaturer bout à bout pour les amplifier et les séquencer. L'alignement sur un génome de référence révèle les régions qui se rapprochent du promoteur proximal de *CDX2* (**Fig. 14A**).

Le prérequis à cette technique est du coup de connaître la séquence du promoteur proximal de Cdx2 pour mettre au point les amorces adéquates pour l'amplification PCR. Son site d'initiation de la transcription (TSS) était connu dans l'intestin²³⁴, mais rien n'assurait qu'il soit maintenu en contexte ectopique. Après avoir vérifié par RACE (*Rapid Amplification of cDNA Ends*) que le site d'initiation de la transcription de *CDX2* était similaire dans les différents contextes cellulaires, l'approche de 4C a été mise au point. (**Fig. 14B**).

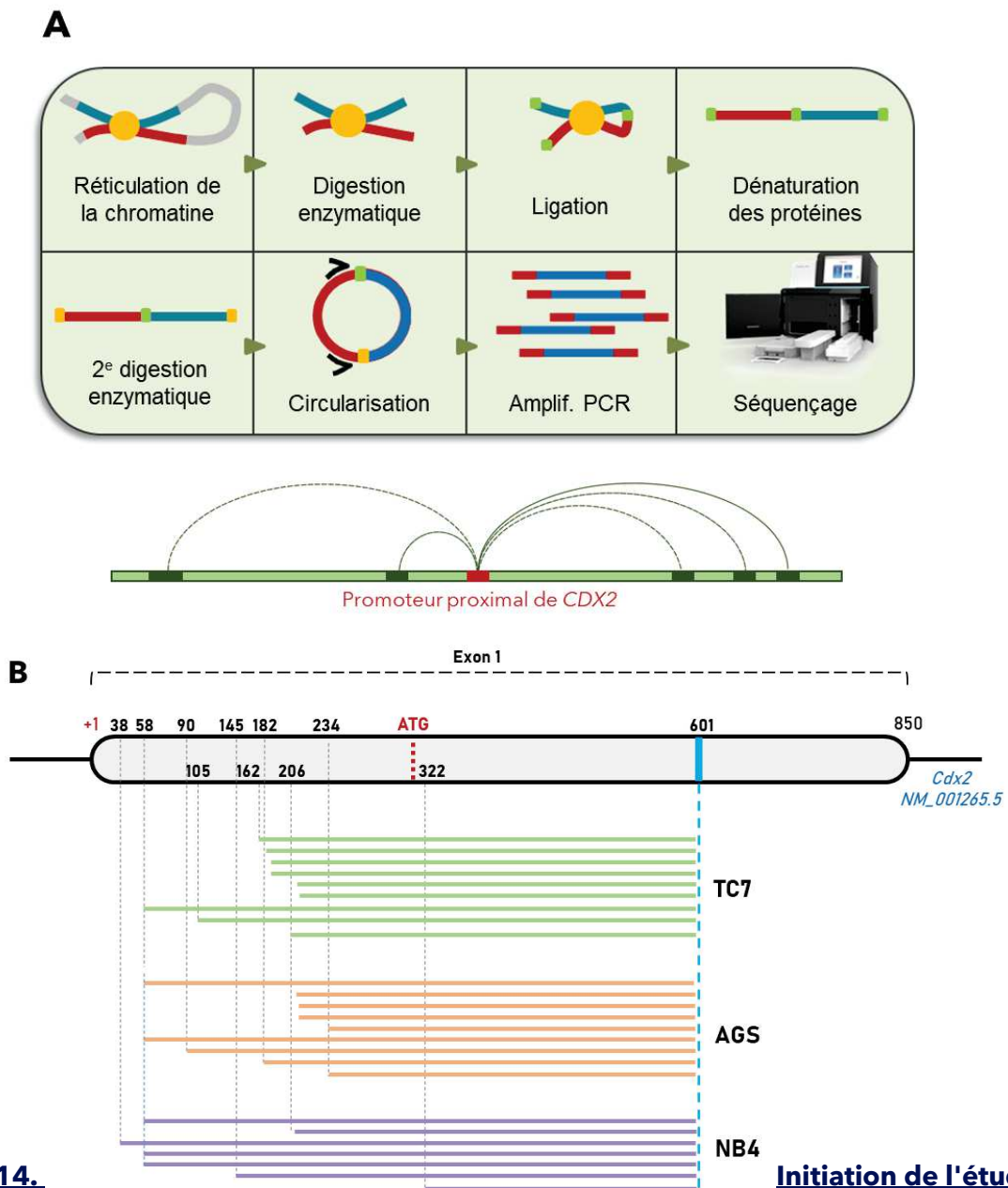


Fig. 14.
par 4C-seq (partie 1).

Initiation de l'étude

(A) Principe du 4C-seq : La technique consiste à maintenir proche les 2 régions temporairement en interaction par cross-linking, puis à les ligaturer bout à bout pour les circulariser, les amplifier et les séquencer. L'alignement sur un génome de référence révèle les régions qui se rapprochent transitoirement du promoteur proximal du gène *CDX2*. (B) Schématisation des résultats du RACE : Le premier exon du gène *CDX2* est représenté en gris. Il porte le codon ATG d'initiation de la traduction. L'amplification des fragments obtenus par RT a été réalisée à l'aide d'une amorce située en position 601 du gène *CDX2*. Pour chaque lignée TC7 (vert), AGS (rouge) et NB4 (bleu), les traits horizontaux représentent la longueur des fragments obtenus par RACE, clonés et séquencés. L'amplification de l'extrémité 5' du gène *CDX2* ne révèle pas de modification du site d'initiation de la transcription entre les 3 contextes cellulaires.

Pour le 4C, l'étape d'amplification par PCR nécessite la mise au point au préalable d'amorces spécifiques de la région d'intérêt, ici le promoteur proximal de *CDX2*, cela en tenant compte de la position des sites de restrictions utilisés pour la digestion enzymatique, ici par *DpnII* et *Csp6I* (**Fig. 15**).

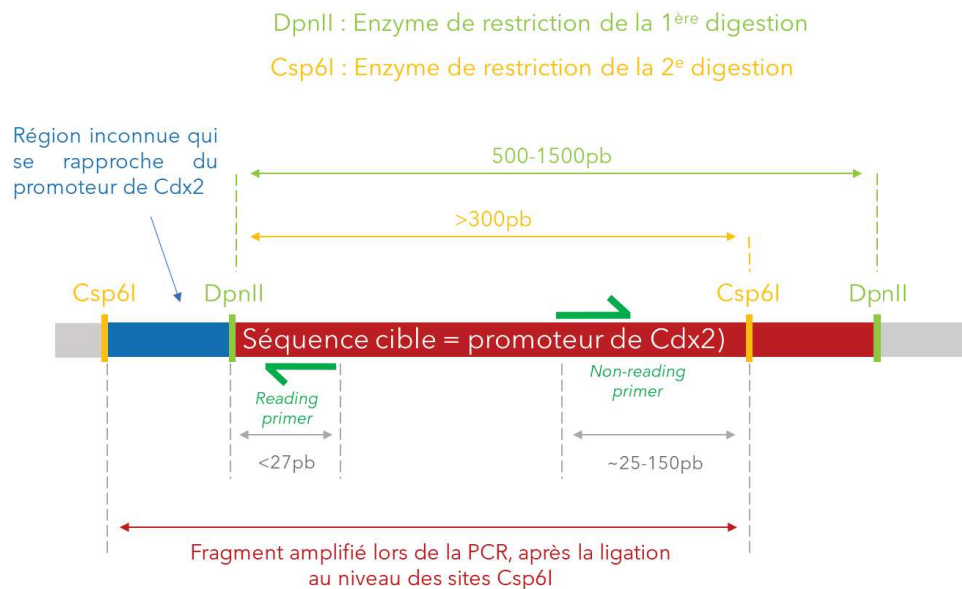


Fig. 15. Choix des amorces pour la PCR

Chaque étape du 4C est ensuite contrôlée par dépôt d'un aliquot des échantillons sur gel d'agarose, et migration électrophorétique (**Fig. 16**).

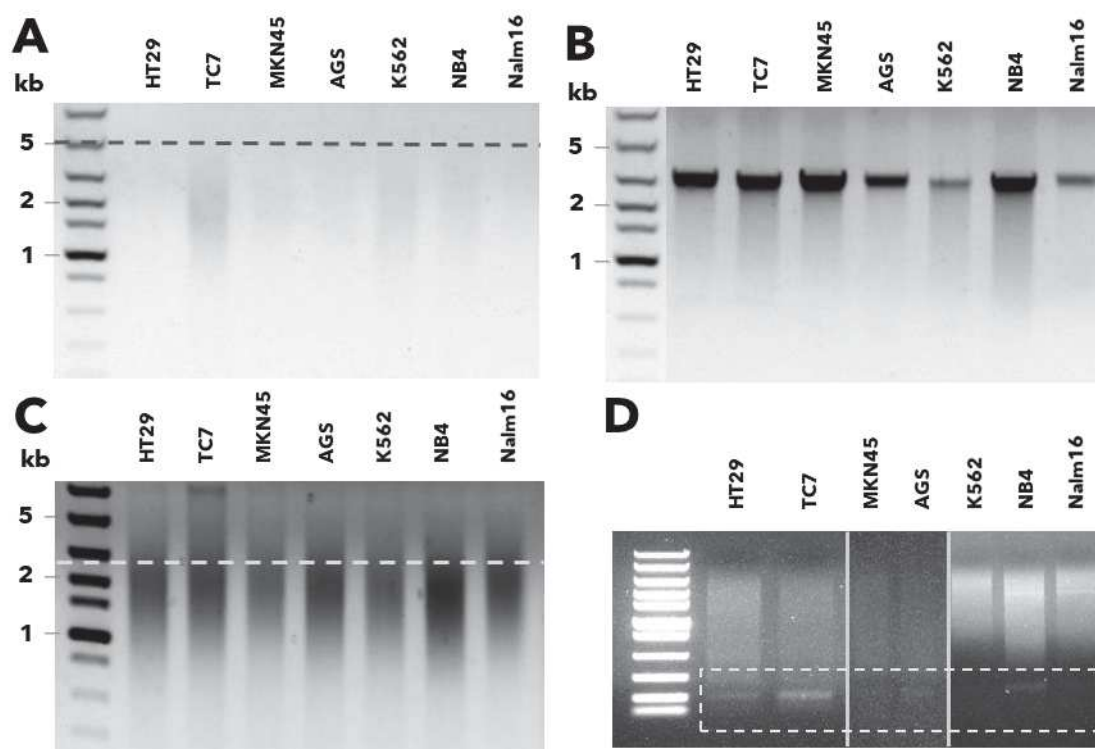


Fig. 16. Contrôles de digestion et ligation.

(A) La majorité des fragments d'ADN doivent être inférieurs à 5kb après la première digestion enzymatique (DpnII). (B) Un shift de poids moléculaire doit être observé à l'issue de la première ligation. (C) La majorité des fragments d'ADN doivent être inférieurs à 2,5kb après la seconde digestion enzymatique. (D) A l'issue de la seconde ligation et de l'amplification PCR, des bandes de faible poids moléculaires doivent être visibles.

Les ADN ont été amplifiés par PCR à l'aide d'oligonucléotides portant un index spécifique de chaque lignée et les ADN amplifiés ont été regroupés et livrés à la plateforme de séquençage. Le séquençage récent a permis de générer 309,547,362 « reads » de qualité validée. L'analyse bioinformatique des résultats de séquençage est en cours.

DISCUSSION

Partie 1

La modélisation des leucémies s'effectuait initialement par exposition des souris à des radiations, agents viraux et composés chimiques. Elle fournissait un support de travail pour la mise au point de chimiothérapies. C'est par intérêt pour les anomalies génétiques sous-jacentes à la pathologie qu'ont commencé à se développer des approches transgéniques, permettant de faire s'exprimer des gènes de fusions (PML-RARA, AML1-ETO) ou des gènes mutés (NMP1c, FLT3-ITD). Les modèles permettent également de moduler l'expression de gènes ayant déjà un rôle hématopoïétique (MLL, IDH1/2), ou bien d'induire une expression ectopique de gènes normalement réprimés dans ce compartiment³⁶².

I. Description de la pathologie se développant dans les souris MxCdx2

Une expression ectopique de *CDX2* a été observée dans 80% des LA, myéloïdes ou lymphoblastiques. Dans le but de comprendre l'implication de *CDX2* dans le développement de ces maladies, nous avons développé au laboratoire un nouveau modèle original de surexpression de *CDX2* dans les cellules sanguines immatures de la moelle osseuse. Après induction de l'expression de *CDX2*, les souris MxCdx2 présentent une moyenne de survie globale de 90 jours. Une expression de *CDX2* est retrouvée dans leurs cellules de moelle osseuse (RT-qPCR et Western Blot), et leur immunophénotypage révèle un blocage de la différenciation lymphoïde et la réduction du lignage mégacaryocyte-érythroïde, associé à un envahissement par une population de cellules monocytaires immatures. L'analyse anatomopathologique de frottis médullaires et sanguins confirme l'envahissement médullaire par ces blastes et leur fuite dans le sang périphérique : avec plus de 80% de monocytes observés dans la moelle osseuse, il s'agit d'une leucémie aiguë myélomonocytaire⁷². L'analyse histologique de rates, foies et thymus atteste d'une atteinte de ces organes, traduite par une infiltration cellulaire et/ou perte d'architecture. Il est admis que les LAM de type myélomonocytaire (M4 et M5 de la classification FAB) présentent un envahissement extramédullaire important³⁶³. Rate et foie sont des sites d'invasion privilégiée, l'hypothèse étant que leur importante vascularisation, associée à la perméabilité des capillaires seraient déterminantes dans le processus d'invasion. L'involution thymique, observable dans une minorité de cas chez notre modèle, est un phénomène normal avec l'âge, et déjà observée dans la modélisation de LAM chez la souris^{364,365}.

	Rawat 2004 / Scholl 2007	Vu 2020		Mx1-Cre 2021
Mode d'induction	Transduction pCdx2 + Greffe	Scl-CreERT2	LysM-CreERT2	Mx1-Cre
Compartiment cellulaire	M.O. totale	CSH	GMP	CSH – MPP ?
Pathologie	LAM	SMD SMP Quelques LAM	SMP	LAM
Lymphopoïèse	 T et B	 T et B	-	 CLP
Erythropoïèse	 Érythrocytes	 MEP	 Érythrocytes	 MEP
Granulo-monocytopoïèse	 Monocytes immatures	 GMP	 Granulocytes et monocytes	 Monocytes immatures

Figure 30. Comparaison de modèles murins d'expression de Cdx2 dans le compartiment hématopoïétique

Au travers des différents modèles, certaines tendances sont observables : perte des populations lymphoïdes et érythroïdes, et expansion de populations myéloïdes : granulocytaires et/ou monocytaires. Mettre en parallèle ces données met en avant que le timing d'induction semble déterminant dans l'agressivité de la pathologie. L'article de Vu met aussi en évidence que même en étant exprimé dans le lignage des granulocytes et monocytes, l'expression de Cdx2 parvient à perturber l'érythropoïèse. Rawat et. al, 2004, Scholl et. al, 2007, Vu et. al, 2020, Galland et. al, 2021.

La modélisation de l'expression de *CDX2* dans le compartiment hématopoïétique avait déjà été entreprise par le passé. Rawat et al., se sont aperçus en 2004 que l'expression seule de *CDX2*, et non pas celle du gène de fusion *ETV6-CDX2* initialement étudiée, induisait une LAM transplantable chez la souris^{355,356}. Ce premier modèle, ainsi que celui développé par Scholl et al., en 2007 impliquaient une transduction de la séquence du gène *CDX2* dans des cellules murines *in vitro*, suivie de leur transplantation dans des souris irradiées. Dans ces deux modèles, des LAM à cellules très peu différenciées (M0-M1) se développent²⁹¹. Le modèle murin se rapprochant le plus des Mx-Cdx2, du point de vue mise au point, est celui développé récemment par Vu et al., permettant une expression inductible *in vivo* de *CDX2*, dans les CSH. L'induction s'effectue par injection de tamoxifène, activant une Cre-ER^{T2} sous le contrôle du promoteur du gène *Tal1/Scl*. Les souris présentent une médiane de survie largement supérieure (301 jours vs 91 jours pour le modèle MxCDX2), et développent des SMD et SMP dont une fraction seulement (23%) évolue en leucémie aiguë de phénotypes divers³⁶⁶, la plupart en LAM myéloérythroïdes, avec parfois des LAL.

Les différentes modélisations employées par ces 2 groupes partagent cette même vocation d'élucider le rôle oncogénique de *CDX2* dans le compartiment hématopoïétique en y induisant son expression. Le contexte cellulaire dans lequel cette induction intervient est cependant différent d'un modèle à un autre. Le promoteur du gène *Tal1/Scl* permet en effet une induction de l'expression du gène *CDX2* ciblée dans les CSH, tandis que, sous la dépendance du promoteur *Mx1*, celle-ci déborde jusque dans les progéniteurs multipotents. Dans le modèle de Scholl, la surexpression est effectuée de manière ciblée dans des CSH *in vitro*. D'ailleurs, l'intensité de cette induction est questionnable, puisque que les souris *Scl-Cdx2* induites par ingestion de tamoxifène (croquettes), présentent une expression de *CDX2* 900 fois inférieure à celles du modèle de Scholl... Les effets observés varient ainsi entre ces 3 contextes, permettant d'établir un parallèle intéressant entre les mécanismes conséquents mis en évidence.

II. Cdx2 transforme les cellules hématopoïétiques

Les tests de réensemencement en série en milieu de méthylcellulose ont montré que seules les cellules de MO de souris Mx-Cdx2 leucémiques sont capables de former des colonies au-

delà de 3 réensemencements. Ce maintien reflète très probablement la capacité de prolifération et/ou une résistance à l'apoptose accrues des monocytes immatures observés dans la MO, et démontre *in vitro* le pouvoir transformant de *CDX2* dans ces cellules. Ces observations sont confirmées *in vivo* par la transplantation d'une fraction (100 000) de cellules de MO totale de souris Mx-Cdx2 induite à des souris Nod Scid Gamma (NSG), conduisant à la même pathologie chez ces dernières et à leur décès encore plus rapide (médiane de survie de 50 jours).

Cette étude est complémentaire à celle de Scholl et al., qui avait montré que la surexpression de Cdx2 dans des cellules de moelle osseuse de souris saine (*wild type*) leur apportait un potentiel de réensemencement en série²⁹¹. Aussi, un abstract publié en 2017, avance que le KD de *CDX2* dans les lignées de LAM SKM1 et EOL-1 réduit également leur clonogénicité, sans pour autant affecter leur prolifération, apoptose et progression dans le cycle cellulaire³⁶⁷. Lors de notre étude, un suivi de la prolifération des cellules leucémiques murines, en diminuant ou non l'expression de *CDX2* (électroporation de siRNA), a été effectué, mais aucune variation n'avait été observée (résultats non publiés). Enfin, les cellules de MO des souris Scl-Cdx2 n'ont à l'inverse pas démontré de hausse de potentiel de réensemencement³⁶⁶. Les auteurs en concluent à une réduction du potentiel d'autorenouvellement des CSH, qu'ils ont confirmée *in vivo*, par la mise en compétition de cellules WT et Scl:Cdx2, transplantées chez des souris irradiées : après injection de tamoxifène, le chimérisme se perd progressivement en faveur des cellules WT³⁶⁶.

III. Cdx2 interfère avec l'hématopoïèse

A. Cdx2 favorise la différenciation des progéniteurs myéloïdes

1. Pas de modification de l'autorenouvellement et de la différenciation

Les immunophénotypages et tri cellulaires effectués à partir de MO totale des souris Mx-Cdx2 induites, n'ont pas révélé de blocage significatif de la maturation des CSH en présence de *CDX2*. Cela, contrairement à des observations antérieures *in vitro*, où son expression diminuait

l'engagement de cellules souches embryonnaires envers le lignage hématopoïétique, et diminuait la différenciation de progéniteurs hématopoïétiques³⁶⁸.

Pour le modèle des souris Scl-Cdx2, l'analyse transcriptomique a permis de montrer que les cellules progénitrices Cdx2⁺ des souris portent en revanche moins de marqueurs d'autorenouvellement (analyse globale par *Gene set enrichment analysis* (GSEA) des gènes impliqués dans l'autorenouvellement), et semblent davantage s'engager dans la différenciation. Y est assimilée la réduction d'expression des gènes *HOX*, allant classiquement de pair avec la baisse d'autorenouvellement et la différenciation cellulaire³⁶⁶.

2. Expansion préférentielle de cellules myéloïdes immatures

Comme mentionné précédemment, les cellules malignes envahissant la cavité médullaire sont monoblastiques et bloquées dans leur différenciation. Les données de l'analyse transcriptomique corrélaient raisonnablement avec cette observation, en révélant une augmentation de l'expression de 5 gènes clés de la différenciation monocyttaire que sont *Irf8*, *Klf4*, *CebpA*, *Egr2* et *Jun*. Ceux-ci contribuent à aiguiller la différenciation myéloïde vers le lignage monocyttaire. L'implication de chacun d'entre eux reste néanmoins paradoxale, du fait de leurs rôles multiples dans les processus hématopoïétiques, mais aussi de leur potentiel antagonisme. Il existe par exemple une balance entre l'expression de *JUN* et *CEBPA* favorisant la prolifération cellulaire dans les LAM³⁶⁹. Il faut d'ailleurs aussi garder à l'esprit que, quel que soit le niveau d'expression d'un gène, l'action de la protéine en découlant peut ensuite être endiguée de diverses façons. A l'exemple de *CEBPA*, dont la contribution à la granulopoïèse est compromise par des interactions protéine-protéine avec *C-JUN*³⁷⁰. La mise en évidence de la dérégulation des protéines codées par ces gènes reste à réaliser. La hausse d'expression de ces gènes est donc associée à un engagement myéloïde préférentiel, mais n'explique pas le blocage de maturation des monocytes.

De la même manière, la perte des progéniteurs mégacaryocytes-érythroïdes corréla avec la diminution significative de l'expression de gènes supportant leur maturation comme *Gfi1b*³⁷¹, *E2f4*³⁷², *Gata-1*³⁷³, *Meis-1*^{374,375}, *Sox-6*³⁷⁶ ou encore *Zfpm1*³⁷⁷.

B. Cdx2 inhibe l'engagement vers le lignage lymphoïde

L'immunophénotypage a révélé une diminution drastique des progéniteurs lymphoïdes (CLP) entraînant une perte quasi-totale du lignage lymphoïde dans les souris Mx-Cdx2, suggérant que la présence de CDX2 interfère avec des mécanismes clés permettant l'engagement lymphoïde.

Cette piste est supportée par la surexpression chez le modèle Mx-Cdx2 de gènes connus pour modifier l'accessibilité de la chromatine en faveur du lignage myéloïde. *In vitro*, il a par exemple été montré que CEBPA et KLF4 participent à la reprogrammation des cellules pre-B en macrophages, notamment via le recrutement de la déméthylase TET2³⁷⁸⁻³⁸⁰. Il a récemment été attribué une action similaire de EGR2³⁸¹.

Au niveau de la chromatine, une analyse en ATAC-seq comparative dans le modèle Scl-Cdx2 vs contrôle, a révélé que la présence de CDX2 augmente l'accessibilité d'enhancers associés à la différenciation myéloïde, les régions rendues davantage accessibles présentant par exemple des sites de fixation de CEBPA³⁶⁶. CDX2 avait déjà démontré sa capacité à impacter l'accessibilité de la chromatine dans le contexte intestinal²⁴⁶. Elle était également observée dans les blastocystes, où il réquisitionne BRG1, sous-unité catalytique de complexes de remodelage de la chromatine, pour réprimer l'expression du gène de pluripotence *OCT4*³⁸². CDX2 et BRG1 cooccupent également plusieurs loci dans la lignée HEK293, suggérant que BRG1 pourrait faciliter l'accessibilité des cibles de CDX2²⁴⁸. BRG1 semble par ailleurs participer au développement des lymphocytes B, en facilitant l'accès de combinaisons de facteurs de transcription spécifiques comme E2A, EBF1 et Ikaros³⁸³ à des enhancers. Il participe également au développement des lymphocytes T³⁸⁴, des granulocytes³⁸⁵ et des érythrocytes³⁸⁶, autant de lignages amoindris en présence de CDX2... !! Peut-être CDX2 le détourne-t-il de ses fonctions hématopoïétiques en s'exprimant dans ce compartiment ? Il serait intéressant de tester si ces 2 protéines interagissent entre elles.

Il a d'ailleurs été montré que la survie et le potentiel d'auto-renouvellement des cellules leucémiques pouvait être supporté par le complexe SWI/SNF : certaines cellules se reposeraient sur BRG1 pour le maintien de leurs programmes transcriptionnels oncogéniques. BRG1 favorise

leur prolifération et viabilité, par exemple en maintenant l'occupation de FTs au niveau d'un enhancer de *MYC*¹⁸¹.

Les analyses transcriptomiques supportent le blocage d'engagement lymphoïde. C'est le cas par exemple de la baisse d'expression d' *E2F4*, dont la délétion dans les CLP diminue l'expression du FT d'engagement lymphoïde *Ikaros*³⁸⁷. L'expression du FT *PBX1* est également diminuée chez notre modèle murin, lequel restreint la maturation myéloïde et préserve le potentiel lymphoïde des progéniteurs hématopoïétiques³⁸⁸.

Curieusement, dans les souris *Scl-Cdx2*, aucune baisse significative des CLP n'est observée. Cependant, alors que la population T n'est pas affectée, une perte des progéniteurs B et des cellules matures est observée chez ces souris, suggérant que *CDX2* bloque la lymphopoïèse B.

Un haut taux de méthylation est requis pour assurer l'engagement lymphoïde. La perte partielle de *DNMT1* est par exemple associée à la hausse d'expression de gènes myéloérythroïdes³⁸⁹. Peut-être l'expression ectopique de *CDX2* entraîne-t-elle une déméthylation de l'ADN, engendrant une restriction myéloérythroïde.

Enfin, la hausse ciblée de la population de MPP pourrait peut-être provenir de l'incapacité de certaines cellules à s'engager dans le lignage lymphoïde, et qui s'accumuleraient donc à ce stade normalement transitoire de différenciation...

C. Qu'en est-il des gènes *HOX*?

Comme mentionné précédemment, *CDX2* est un régulateur des gènes *HOX* lors de l'hématopoïèse primitive. Leur expression se maintient tout au long de la vie adulte, et contribue à l'établissement et la différenciation des différents lignages. De manière générale, leur expression est davantage retrouvée dans les progéniteurs en prolifération, puis diminue au cours de la différenciation. Dans les LAM, leur expression peut être perturbée par des réarrangements les impliquant eux-mêmes ou bien par des mutations et réarrangement affectant leurs régulateurs. De par le rôle des gènes *HOX* dans l'autorenouvellement des

progéniteurs hématopoïétiques, on suppose que la dérégulation de leur expression soutient la maintenance des CSL. Leurs contributions à la leucémogenèse sont progressivement investiguées par des études de perte et de gain de fonction. Leurs profils d'expression dans les cellules leucémiques ne semblent pas uniquement influencés par leur stade de différenciation^{230,390,391}.

Le point commun notable entre les modèles Mx-Cdx2 et Scl-Cdx2 est la diminution globale d'expression des gènes HoxA. HoxA9 est celui dont la diminution est la plus importante dans notre modèle, tandis qu'il présente plutôt une tendance à la surexpression dans les leucémies en général chez l'Homme³⁹². On peut néanmoins associer ici sa diminution à la disparition du lignage lymphoïde, puisqu'il semble, avec FLT3, être essentiel à la lymphopoïèse B³⁹³⁻³⁹⁵. Les souris déficientes en *Hoxa9* présentent une diminution de CLP, CMP et GMP, et de leurs précurseurs lymphoïdes B et T, tout en conservant un pool normal de CSH³⁹⁶. La diminution d'expression du gène *Hoxa7* avait déjà été associée à la réduction de la population de MEP³⁹⁶. Aussi, de manière intéressante, *Hoxc6* est l'unique gène de son cluster à être augmenté en présence de CDX2. Une étude antérieure avait permis d'observer que ce gène privilégie l'engagement myéloïde à celui lymphoïde, son expression diminuant ensuite dans les cellules matures. Sa surexpression bloque la différenciation myéloïde³⁹⁷.

La dérégulation des gènes *HOX* reste cependant un événement rare dans les LAL, en comparaison avec les LAM, suggérant que l'action oncogénique de *CDX2* peut emprunter d'autres réseaux géniques³⁹⁸.

IV. Perturbation de la voie de signalisation du TGF- β

Parmi les nombreuses voies de signalisation dérégulées dans nos souris Mx-Cdx2, la voie du TGF- β a retenu mon attention. Cette voie est essentielle au cours de l'hématopoïèse. Brièvement, elle possède une trentaine de ligands activateurs, variant selon le type cellulaire. Ceux-ci favorisent la dimérisation des récepteurs au TGF- β , transmembranaires et à activité sérine-thréonine kinase. Une fois activés, ils phosphorylent les protéines R-SMAD intracellulaires et permettent leur importation dans le noyau où elles régulent la transcription de leurs gènes cibles.

Dans la phase précoce et la progression des LAM, on note souvent l'apparition d'une résistance à l'action du TGF- β , via un appauvrissement en récepteurs de surface, en Smads, ou via une mutation les rendant dysfonctionnels. Plusieurs travaux attribuent ainsi un rôle suppresseur de tumeur à cette voie de signalisation dans le compartiment hématopoïétique³⁹⁹. Dans le modèle Mx-Cdx2, 2 gènes activateurs de la voie du TGF- β (*Inhibine A* et *Bmp4*) sont sous-exprimés, et 2 gènes inhibiteurs (*Smad7* et *Bambi*) sont sur-exprimés suite à l'induction de l'expression de Cdx2 dans les CSH, suggérant une inhibition de cette voie de signalisation.

A. La surexpression du gène Bambi réduit l'action du TGF- β

La protéine BAMBI (*BMP and Activin Membrane Bound Inhibitor*), par ailleurs connue pour stimuler la voie oncogénique de la β -caténine⁴⁰⁰ dans l'intestin, exerce son rôle inhibiteur de la voie du TGF- β par ancrage dans la membrane et blocage de la dimérisation des sous-unités du récepteur^{401,402}. Ceci diminue l'activation des récepteurs de surface et donc la signalisation en découlant, quels que soient les ligands activateurs. Surexprimé près de 15 fois en présence de CDX2 dans la moelle osseuse des souris Mx-Cdx2, par rapport aux souris contrôles, le gène *Bambi* a été étudié comme acteur potentiel de la leucémogénèse.

1. Bambi est une cible directe de Cdx2

L'interaction de CDX2 avec le promoteur de *Bambi*, au niveau de 2 sites consensus de fixation (A/CTTTATATA/G), ainsi que la corrélation d'expression entre les 2 gènes démontre que *Bambi* est une cible transcriptionnelle directe de CDX2 dans le modèle murin Mx-Cdx2. Chez l'Homme, ces 2 sites consensus de fixation sont conservés, et dans des lignées humaines de LAM (SKM1 et NB4) l'expression de *CDX2* stimule l'activité du promoteur de *Bambi*.

Cependant, aucune augmentation de l'expression du gène *Bambi* n'est répertoriée chez les patients atteints de LAM. Aussi, au même titre que la modulation d'expression de *CDX2*, celle de *Bambi* n'a pas montré d'incidence sur la prolifération des cellules leucémiques murines et humaines (résultats non publiés). Il ne représente donc ici qu'un levier actionné chez la souris pour freiner la signalisation par la voie du TGF- β . Nous avons néanmoins saisi cette perche

tendue pour nous intéresser à l'impact de l'inhibition de cette voie dans les cellules leucémiques murines. Aussi, BAMBI semble, à l'image de *CDX2*, endosser des fonctions duelles selon son compartiment d'expression. Dans l'épithélium intestinal, il se situe par exemple à la croisée des voies de signalisation WNT et TGF- β , en étant respectivement cible transcriptionnelle et activateur d'un côté, et inhibiteur de l'autre⁴⁰³. Investiguer ses modes d'action dans un compartiment reste une source non négligeable d'informations translationnelles...

2. BAMBI contribue au blocage de la différenciation monocyttaire

Les ligands activateurs de la voie du TGF- β sont connus pour leur pléiotropie, également dans le compartiment hématopoïétique⁴⁰⁴. Le TGF- β 1 est un ligand de la voie favorisant l'engagement des CSH au destin cellulaire myéloïde⁴⁰⁵. Il a notamment été montré, à partir de la lignée promyélocytaire THP1, que le TGF- β 1 promouvait la maturation de cellules monocytaires immatures en macrophages⁴⁰⁶. Nous avons ainsi évalué l'impact de l'inhibition de la voie de signalisation sur la différenciation myéloïde, compromise chez le modèle Mx-Cdx2. Le traitement des cellules leucémiques murines, dérivées à partir de notre modèle Mx-Cdx2, au TGF- β 1 ne réhausse que peu l'expression du marqueur de différenciation membranaire CD11B. L'action du TGF- β 1 est en revanche largement amplifiée lors de l'inhibition concomitante de l'expression du gène *CDX2* ou du gène *Bambi*. Ces observations incitent à penser que l'expression ectopique du gène *CDX2* dans les CSH compromet la maturation des monocytes, en partie en inhibant la signalisation par la voie du TGF- β , cela en promouvant la surexpression du gène *Bambi*.

Aussi, le TGF- β est initialement un facteur de survie pour les monocytes, qui entrent en apoptose en son absence⁴⁰⁷. La baisse de signalisation par cette voie observée chez le modèle MxCdx2 semble indiquer que les monoblastes leucémiques s'en sont affranchis.

B. Impact de la voie sur les populations hématopoïétiques

1. Autorenouvellement, quiescence, prolifération et engagement des CSH

La voie de signalisation du TGF- β contribue notamment à l'autorenouvellement⁴⁰⁸ et à la quiescence des CSH⁴⁰⁴. Ces 2 paramètres ne semblent pas particulièrement altérés dans le modèle Mx-Cdx2, cela malgré la perte des mégacaryocytes, acteurs majeurs de la quiescence⁵³, et malgré la hausse d'expression de Smad7, associée dans une autre étude à une hausse de la capacité d'autorenouvellement⁴⁰⁹. L'expression des gènes *Shp-1*, *Cd63* et *Msi2*, connus pour favoriser la quiescence et l'autorenouvellement par le biais de cette voie, reste relativement inchangée en présence de CDX2⁴¹⁰⁻⁴¹².

L'effet antiprolifératif du TGF- β sur les CSH est dépendant de son abondance, et du type de CSH. *In vitro*, il possède un effet antiprolifératif à forte dose (<20pg/mL) sur les CSH au destin cellulaire aussi bien myéloïde (CSH-My) que lymphoïde (CSH-Ly). A faible dose (10pg/mL) en revanche, il promeut la formation de colonies myéloïdes exclusivement. Le mécanisme lui permettant cette fonction duelle n'est pas connu⁴¹³. On peut supposer que la résistance au TGF- β développée par les progéniteurs exprimant CDX2 ne permet plus que l'expansion de ceux au devenir myéloïde, ceux au devenir lymphoïde étant maintenus en quiescence...

Aussi, l'inhibition de la voie par la surexpression de SMAD7 a déjà été décrite comme induisant un biais de différenciation en faveur du myéloïde. Les souris NSG greffées avec des cellules de sang de cordon surexprimant SMAD7 développent davantage de progéniteurs myéloïdes que lymphoïdes⁴¹⁴.

2. Différenciation inégale du lignage myéloïde

In vitro, la signalisation par la voie du TGF- β stimule la différenciation des progéniteurs sensibles au GM-CSH⁴¹⁵. *In vivo*, l'administration de TGF- β augmente le nombre de CFU-GM et de myéloblastes dans les fémurs et augmente le nombre de granulocytes matures dans la MO et le sang périphérique⁴¹⁶. Ces observations soutiennent que le TGF- β promeut la granulopoïèse, et corrèlent avec la perte des GMP observée chez le modèle Mx-Cdx2.

De la même manière, la baisse d'expression du gène *Bmp4* en présence de CDX2, codant pour un ligand soutenant la différenciation érythroïde^{417,418}, corrèle avec l'importante anémie observée dans les souris Mx-Cdx2. La signalisation par la voie du TGF- β , et notamment par BMP4, contribue d'ailleurs également à la maturation des mégacaryocytes^{419,420}.

V. Cdx2 oncogénique dans le lignage lymphoïde ?

Aussi intrigant soit-il, le biais d'engagement cellulaire envers le lignage myéloïde constitue d'une certaine manière la limite majeure des modèles Mx-Cdx2 et de ceux développés par Scholl et al., et Vu et al. Ne développant que des pathologies myéloïdes, ceux-ci ne permettent pas de s'intéresser aux conséquences d'une expression ectopique de *CDX2* dans le lignage lymphoïde, alors même que 80 % des patients atteints de LAL expriment *CDX2* dans leur MO. Le rôle de *CDX2* dans la leucémogénèse lymphoïde reste inconnu.

Pour aborder cette question, un second modèle murin a été mis au point au laboratoire : les souris Rosa26-LSL-hCdx2:CD19-CreER^{T2} (dites CD19-Cdx2), chez lesquelles l'expression de *CDX2* dans les précurseurs lymphoïdes B (stade pré-B) est inducible au tamoxifène⁴²¹. Les études préliminaires de ce projet auquel j'ai participé ont consisté en l'analyse du phénotype des souris CD19-Cdx2 à différents temps après induction. L'induction de l'expression de *CDX2* a été validée, et augmente transitoirement le nombre de progéniteurs B (à 2 semaines après induction), mais se perd au cours du temps. L'expression de *CDX2* ne confère donc pas d'avantage prolifératif aux cellules pré-B mais semble bloquer la différenciation. Cependant, la perte de l'expression de *CDX2*, le moment d'induction de *CDX2* au cours de la lymphopoïèse B et le faible % de cellules exprimant *CDX2* ne nous ont pas permis de montrer le rôle oncogénique de *CDX2* dans ce modèle.

Peut-être une induction de l'expression de *CDX2* dans des cellules déjà engagées, ici au stade de différenciation pré-B, est-elle trop tardive pour engendrer une hémopathie ? L'induction effectuée par Vu et al., dans les GMP, n'a pas non plus permis l'émergence d'une leucémie³⁶⁶.

Pour y pallier, 2 stratégies sont envisageables. Il est possible d'une part d'introduire un avantage prolifératif concomitant à l'expression de *CDX2* dans les cellules pré-B, par exemple en croisant les souris CD19-Cdx2 avec des souris porteuses d'une mutation du gène *Kras*, connu pour stimuler la prolifération⁴²². Il est possible d'autre part d'induire l'expression de *CDX2* plus précocement dans la lymphopoïèse B. Placer la recombinaise CreER^{T2} sous la dépendance du promoteur du gène *Mb1/Cd79A* permettrait par exemple une induction dès le stade pro B (A)⁴²³.

Cependant, induire l'expression de *CDX2* précocement dans la lymphopoïèse B, et cela même au stade de pré-B, fait redouter une dédifférenciation des progéniteurs B. Ils possèdent en effet une certaine plasticité leur permettant de rebifurquer vers le lignage myéloïde⁴²⁴. Le phénomène de switch de lignée s'observe d'ailleurs en conditions pathologiques, où des leucémies lymphoblastiques sont capables, lors d'une rechute, d'évoluer en leucémie myéloïde et vice-versa⁴²⁵.

Partie 2

Nous avons vu que l'expression de *CDX2* fluctue en fonction du contexte cellulaire, physiologique ou pathologique. L'origine de ces dérégulations a déjà fait l'objet de recherches, qui n'ont pour l'instant pas permis d'identifier de mécanisme causal. Nous avons fait l'hypothèse que peut-être l'expression ectopique du gène *CDX2* dans le compartiment hématopoïétique impliquait des éléments de régulation distaux de son site d'initiation de la transcription. Il est en effet démontré depuis plusieurs années que la perturbation de régions enhancers soutient activement le processus de leucémogénèse⁴²⁶.

Cette partie du projet consiste à décrypter les mécanismes de régulation de l'expression du gène *CDX2* en identifiant et comparant les enhancers impliqués dans des lignées cellulaires cancéreuses intestinales, gastriques et leucémiques qui expriment ou non *CDX2*.

Ce 2^e versant de l'étude n'ayant pu être mené très loin, je discuterai ici des bases bibliographiques utilisées pour son initiation, des résultats préliminaires, mais aussi des pistes pour la suite du projet.

I. Validation de l'expression de *Cdx2*

A. Lignées intestinales

Le gène *CDX2* participe à la spécification du tissu intestinal lors du développement embryonnaire, puis son expression se restreint à l'épithélium intestinal, où il sera exprimé tout au long de la vie. Il s'agit de son compartiment d'expression physiologique²³⁹. Les cellules de la lignée Caco2-TC7 sont issues d'un adénocarcinome colique, ayant conservé leur capacité de différenciation spontanée, et leur expression du gène *CDX2*. Les cellules de la lignée HT-29 sont également issues d'un adénocarcinome colique, mais présentent un phénotype indifférencié, associé à une forte diminution de l'expression de *CDX2*. Ces deux lignées reflètent une partie de la diversité d'expression de *CDX2* dans les CCR. Elles sont classiquement utilisées en parallèle pour étudier les fonctions de *CDX2* dans le compartiment intestinal, et les conséquences de sa perte d'expression, typiquement observée lors de cancers colorectaux^{270,277,427,428}. L'analyse de l'expression de *CDX2* menée lors de notre étude coïncide

avec les données de la littérature, puisque nous confirmons une forte expression de CDX2 dans les Caco2/TC7 et une faible expression dans les HT-29.

B. Lignées leucémiques

Il a été montré que 80% des patients atteints de leucémie aiguë expriment de manière aberrante le gène *CDX2* dans leurs cellules de moelle osseuse²⁹¹. Nous avons choisi la lignée cellulaire K-562, issue d'une leucémie myéloïde chronique, comme contrôle n'exprimant pas *CDX2*. La lignée cellulaire NB-4³¹⁷, issue d'une leucémie aiguë promyélocytaire, et la lignée cellulaire Nalm-16, issue d'une leucémie aiguë lymphoïde de précurseurs B, modélisent respectivement l'expression de *CDX2* en contexte de LAM et LAL. Comme attendu, le niveau d'expression -anormale- de *CDX2* dans le compartiment hématopoïétique reste largement inférieur à celui observé dans le compartiment intestinal, impliquant le recours à des immunoprécipitations des protéines CDX2. Aussi, l'intense expression observée dans la lignée Nalm-16 reflète celle retrouvée dans certains sous-groupes de patients atteints de LAL-B, chez lesquels la très forte expression de *CDX2* est associée à une faible survie (Yasuda et al., 2021, et E. Clappier, résultats non publiés).

C. Lignées gastriques

Les adénocarcinomes gastriques peuvent être de 2 types : intestinal ou diffus⁴²⁹. Une expression aberrante du gène *CDX2* est retrouvée notamment dans ceux de type intestinal⁴³⁰. Il a été montré que l'expression ectopique du gène *CDX2* dans l'épithélium gastrique normal conduisait à l'apparition de lésions pré-cancéreuses, notamment via la mise en place d'un programme transcriptionnel intestinal²⁸⁵. Généralement, les adénocarcinomes gastriques de type diffus, caractérisés par une différenciation limitée des cellules, n'expriment pas Cdx2. Les cellules des lignées AGS et MKN-45 sont issues d'adénocarcinomes gastriques de type respectivement intestinal et diffus. Toutes 2 expriment *CDX2*, son expression étant plus forte dans la première³¹⁰. Malgré cette différence, il serait intéressant d'inclure une 3^e lignée gastrique ne présentant pas du tout d'expression de *CDX2* pour renforcer l'étude comparative.

II. Régions de chromatine accessible

A. ChIP-seq et modifications post-traductionnelles d'histones

Les modifications post-traductionnelles (PTM) des histones reflètent l'accessibilité de la chromatine aux facteurs de transcription. Les marques H3K27ac sont particulièrement localisées au niveau de promoteurs et enhancers de gènes activement transcrits. Les marques H3K4me3 sont particulièrement associées aux régions régulatrices proximales, à proximité des sites d'initiation de la transcription de gènes activement transcrits^{431(p202)}. Nous observons dans l'ensemble une abondance supérieure des 2 types de PTM d'histones activatrices, dans les lignées exprimant *CDX2*, en comparaison avec celles ne l'exprimant pas, au niveau de son locus. Ces différences encouragent la poursuite de l'analyse dans ces lignées.

Une portion de chromatine accessible ne correspond cependant pas forcément à une région régulatrice stimulant l'expression du gène *CDX2*. Les « silencers » sont le pendant bifonctionnel des enhancers : ce sont des régions régulatrices distales, bien moins étudiées et caractérisées, présentant une fonction duelle, activatrice ou répressive de l'expression des gènes, selon le contexte cellulaire⁴³². Malgré leur dénomination d'origine, opposer les silencers aux enhancers serait une simplification à outrance, une hypothèse étant que certains silencers, si ce sont pas tous, seraient des enhancers dans d'autres tissus. Le mode d'action des silencers serait analogue à celui des enhancers, via un repliement de la chromatine, d'après des observations faites chez la *Drosophile*⁴³³, mais n'a pour l'instant pas été décrit chez l'Homme.

Il est difficile de discriminer clairement ces 2 types de régions régulatrices sur la base de PTMs qu'elles arborent. Néanmoins, il semblerait possible de détecter les silencers par la co-occurrence de marques H3K27me3 et d'accessibilité de la chromatine à la DNase I⁴³⁴. Il serait ainsi intéressant de compléter notre étude par ce type d'analyse et par la comparaison plus globale des marques activatrices ou inhibitrices au niveau de la région élargie autour du locus du gène *CDX2*, dans nos 7 lignées cellulaires.

Aussi, il a déjà été montré que l'activation de l'expression d'un oncogène pouvait impliquer la perturbation d'un silencer. Le gène *RUNX1* se retrouve par exemple surexprimé dans certaines

LAM suite à une translocation empêchant son interaction avec un silencer⁴³⁵. Cette dualité des silencers motive d'autant plus l'investigation des régions régulatrices distales du gène *CDX2*, lui-même polyvalent.

B. ATAC-seq et espacement des nucléosomes

Les profils de l'ATAC-seq, basés sur l'accessibilité de la chromatine à la Tn5, se superposent à ceux du ChIP-seq. En revanche, l'intensité des pics n'est pas forcément corrélée. D'autres modifications post-traductionnelles d'histones peuvent en effet favoriser l'accessibilité de la chromatine. Pour s'intéresser aux enhancers actifs, il serait par exemple également envisageable de cibler la modification H3K4me1, qui y est particulièrement associée⁴³⁶.

A l'échelle du locus de *CDX2*, les profils d'accessibilité de la chromatine sont cohérents avec les données de la littérature, et permettent par exemple de retrouver l'enhancer de *CDX2*, localisé 9kb en amont de son TSS, décrit par Benahmed et al²³⁴.

L'ensemble de ces résultats restent extrêmement préliminaires. L'analyse comparative approfondie des profils de ChIP-seq et ATAC-seq avec l'aide d'un bioinformaticien est en cours.

III. Bivalence du promoteur proximal de *Cdx2*

De nombreuses études sur le développement très précoce de l'embryon, au moment de la détermination du trophoctoderme, ont montré que le promoteur de *CDX2* porte des modifications post-traductionnelles (PTM) d'histones activatrices (H3K4me3) et inhibitrices (H3K27me3) simultanément, faisant de lui un gène bivalent^{138,326}. Notre hypothèse était donc que la modulation aberrante de l'expression du gène *CDX2* pourrait à nouveau impliquer cette balance entre PTMs antagonistes d'histones au niveau de son promoteur proximal.

Nos résultats vont globalement en faveur de cette théorie. Il y a effectivement une présence différentielle de marques activatrices et inhibitrices en fonction de l'expression ou non de *CDX2*.

Les résultats de ChIP-qPCR obtenus ne nous donnent qu'un échantillon des profils d'accessibilité de la chromatine au niveau du locus proximal de *Cdx2*, puisque nous avons choisi une seule région du promoteur allant de -28 à +242. Cette région avait été choisie par Sun et al., (2017) sur la base de profil ChIP-seq obtenu sur NCBI. Il est évident que des ChIP-seq réalisés avec un anticorps reconnaissant un/des marques répressives, comme un anti-H3K27me3 par exemple, seront nécessaires pour compléter cette étude.

IV. Régions en interaction avec le promoteur proximal de *Cdx2*

Pour détecter les événements de repliement transitoire de la chromatine nous avons choisi d'avoir recours à la technique du 4C-seq⁴³⁷.

Le prérequis à la technique du 4C était de connaître approximativement la localisation du site d'initiation de la transcription (TSS) de *CDX2*. On constate que les fragments amplifiés lors du RACE n'atteignent jamais le +1 connu. Cela peut s'expliquer par la région promotrice riche en GC, pouvant créer des hyperstructures au niveau de l'ARN qui gênent la progression de l'enzyme de l'étape de rétrotranscription. Aussi, la variabilité observée entre les fragments est aussi potentiellement liée à cette absence de TATA box, puisque les gènes avec un promoteur riche en GC ont souvent un site d'initiation de la transcription variable.

Cdx2 s'exprime également de manière aberrante dans l'épithélium oesophagien lors d'épithéliums de Barrett, qui consistent en des transdifférenciations locales en structures de type intestinale⁴³⁸. Dans ce contexte d'adénocarcinome de l'œsophage, un site de fixation cryptique du facteur de transcription FOXA3 a été mis en évidence. Cette région pourrait constituer un enhancer de *Cdx2*⁴³⁹. Il serait pertinent d'inclure des lignées d'adénocarcinome de l'œsophage à cette étude globale.

Enfin, pour en revenir à la notion de détournement d'enhancer, on pourrait imaginer que l'expression aberrante de *Cdx2* soit associée à son intégration au sein d'un TAD. La délétion retrouvée à proximité du gène *FLT3* chez certains patients LAL-B exprimant fortement *Cdx2* (résultats non publiés) pourrait par exemple l'inclure dans le TAD de *FLT3* et l'exposer à son enhancer.

CONCLUSION

I. Fonction oncogénique de Cdx2 dans les leucémies aiguës

L'expression ectopique de *CDX2* dans le compartiment hématopoïétique déclenche une LAM chez les souris Mx-Cdx2, se caractérisant par la perte du lignage lymphoïde, et l'expansion de monocytes immatures.

Une théorie est que les oncogènes façonneraient le destin cellulaire en restreignant à une seule possibilité le choix de lignage des CSH. Ce qui expliquerait pourquoi les CSLs se restreignent généralement à un lignage... Dans le modèle Scl-Cdx2, la capacité de *CDX2* à favoriser le destin cellulaire myéloïde via un remodelage de la chromatine soutient cette théorie. Aussi, le stade de différenciation cellulaire auquel l'expression de Cdx2 est induite semble radicalement orienter le devenir cellulaire, aussi bien en termes de lignage qu'en terme de pathogenèse. Peut-être la divergence entre Mx-Cdx2 et Scl-Cdx2 s'explique-t-elle par une intensité ou un timing d'expression différent ? Recouper ces notions de timing et d'accessibilité de la chromatine rappelle le développement embryonnaire pendant lequel elles sont essentielles pour déterminer le devenir cellulaire.

Chez les patients, l'identité des cellules de la moelle osseuse exprimant *CDX2* et le rôle de *CDX2* sur les cellules de son environnement ne sont pas définis. Transforme-t-il son microenvironnement comme c'est le cas dans les CCR (Balbinot 2018) ?

Les réseaux de voies de signalisation impliqués dans l'établissement de l'hématopoïèse primitive et définitive pendant le développement embryonnaire pourraient être réactivées lors de leucémies, et générer des CSL aux propriétés d'autorenouvellement et de différenciation altérées. Sont concernées par cette théorie les voies WNT, TGF- β , BMP, NOTCH, HH mais aussi l'axe *CDX-HOX*... Ce dernier axe reste justement à creuser.

II. Régulation du gène CDX2 dans les leucémies aiguës

Dans la continuité de l'étude initiée, le but sera de croiser les 2 sets de données : les informations sur l'activité de la chromatine (ChIP-seq et ATAC-seq), et celles sur sa topologie (4C-seq). Une fois les potentielles régions enhancers et silencers identifiées, il sera possible d'en déduire *in silico* les facteurs de transcriptions et autres régulateurs transcriptionnels qui pourraient interagir avec le promoteur. Cela se fera par analyse bio-informatique en se basant sur les sites consensus de fixation connus (via des outils tels Gibbs, HOMER, TFBIND).

Il s'agira ensuite d'éprouver la pertinence des régulateurs potentiels. A partir de lignées cellulaires dans un premier temps, puis sur des échantillons patients, seront recherchées une corrélation d'expression d'une part, et une interaction physique d'autre part, entre le régulateur et le promoteur proximal de *CDX2*. La validation de ces régulateurs se fera par une étude fonctionnelle visant à étudier leur importance dans l'expression du gène *CDX2* et la leucémogénèse à l'aide de modèles cellulaires et murins et par CRISPR-Cas9. Ces régulateurs devront à terme être replacés dans des voies de signalisation, dans le but de trouver un angle d'attaque pour bloquer l'expression de *CDX2*. Ce mécanisme devra être spécifique du compartiment hématopoïétique, pour ne pas cibler son expression intestinale. Cela justifie l'étude comparative entreprise entre les différentes lignées qui a pour but de mettre à jour des enhancers spécifiques des tissus.

MATÉRIELS & MÉTHODES

I. Matériels

A. Lignées cellulaires

La lignée cellulaire cancéreuse HT-29 est dérivée d'un adénocarcinome colorectal (ATCC® HTB-38TM). La lignée cellulaire cancéreuse Caco2/TC7 est dérivée d'un adénocarcinome colorectal (Sigma-Aldrich® SCC209TM). La lignée cellulaire cancéreuse K-562 est dérivée d'une leucémie chronique myéloïde (ATCC® CCL-243TM). La lignée cellulaire cancéreuse NB-4 est dérivée d'une leucémie aiguë promyélocytaire (DSMZ® ACC207TM). La lignée cellulaire cancéreuse Nalm-16 est dérivée d'une leucémie aiguë lymphoblastique B (DSMZ® ACC680TM). La lignée cellulaire cancéreuse MKN-45 est dérivée d'un adénocarcinome gastrique (DSMZ® ACC 409TM). La lignée cellulaire cancéreuse AGS est dérivée d'un adénocarcinome gastrique (ATCC® CRL-1739TM).

B. Anticorps

Tableau 2. Anticorps utilisés pour les immunoprécipitations de chromatine et Western Blot.

Anticorps	Espèce	Fournisseur	Utilisation
@CDX2	Lapin	Abcam #AB76541	ChIP / WB
@CDX2	Souris	Biogenex #AM392GP	ChIP / WB
@ACTINE	Souris	Millipore #MAB1501	Western Blot
@Mouse	Cheval	Cell Signaling #7076S	Western Blot
@Rabbit	Chèvre	Cell Signaling #7074S	Western Blot
@H3K27ac	Lapin	Active Motif #39133	ChIP
@H3K4me3	Lapin	Sigma Aldrich #04-745	ChIP
@H3K27me3	Lapin	Active Motif #39157	ChIP

C. Sondes nucléotidiques

Tableau 3. Sondes utilisées pour la PCR quantitative (Applied Biosystems).

ADNc ciblé	Technologie	Référence ou Séquence	Utilisation
<i>hCDX2</i>	Taqman	Hs00230919_m1	RT-qPCR
<i>hTBP</i>	Taqman	Hs999999910_m1	RT-qPCR
<i>promCDX2</i>	Sybrgreen	5'-GTGGGAGGAAAGAAGGAAGAAAGG-3' 5'-GTACATGCTCACGTCCTTGCCAG-3'	ChIP-qPCR
<i>ACTINEB2</i>	Sybrgreen	<i>Human positive Control Primers Set ACTB2 Active Motif</i>	ChIP-seq / ChIP-qPCR
<i>GAPDH</i>	Sybrgreen	<i>Human positive Control Primers Set GAPDH Active Motif</i>	ChIP-seq / ChIP-qPCR
<i>Neg1</i>	Sybrgreen	<i>Human negative Control Primers Set 1</i>	ChIP-seq / ChIP-qPCR
<i>Neg2</i>	Sybrgreen	<i>Human negative Control Primers Set 2</i>	ChIP-seq / ChIP-qPCR

Tableau 2. Amorces PCR des bibliothèques d'ATAC-seq

Lignée	Nom	Séquence
-	Ad1_noMX	AATGATACGGCGACCACCGAGATCTCACTCGTCGGCAGCGTCAGATGTG
HT-29	Ad2.1	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGGAGATGT
TC7	Ad2.2	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGGAGATGT
MKN-45	Ad2.3	CAAGCAGAAGACGGCATACGAGATTCTGCCTGTCTCGTGGGCTCGGAGATGT
AGS	Ad2.4	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGGGCTCGGAGATGT
K-562	Ad2.22	CAAGCAGAAGACGGCATACGAGATTGTGACCAGTCTCGTGGGCTCGGAGATGT
NB-4	Ad2.23	CAAGCAGAAGACGGCATACGAGATAGGGTCAAGTCTCGTGGGCTCGGAGATGT
Nalm-16	Ad2.24	CAAGCAGAAGACGGCATACGAGATAGGAGTGGGTCTCGTGGGCTCGGAGATGT

Tableau 3. Amorces PCR des bibliothèques de 4C-seq

Nom	Séquence
N-hCdx2-Csp6I	CAAGCAGAAGACGGCATACGAGCTCTCCGATCTTCCTCGAGTCCCCTCGCC
A-hCdx2-DpnII	AATGATACGGCGACCACCGAGATCTCACTCTTCCCTACACGAGGCTCTCCGATCTXXXXXXCCTCCGCACAAGACAAATA

Lignée	Référence	N° d'index	Séquence de l'index Illumina
HT-29	JLFD81	6	GCCAAT
TC7	JLFD80	12	GCCAAT
MKN-45	JLFD83	5	ACAGTG
AGS	JFLD82	19	GTGAAA
K-562	JLFD84	10	TAGCTT
NB-4	JLFD85	2	CGATGT
Nalm-16	JLFD86	16b	CCGTTC

En rouge : séquence complémentaire à la séquence de Cdx2. En bleu : index permettant de discriminer les différentes lignées lors du démultiplexage. En jaune : adaptateur Illumina.

II. Méthodes

A. Extraction des ARN totaux

Les cellules sont rincées dans du PBS puis lysées 5min à température ambiante dans 1mL de réactif *TRI Reagent* (Sigma-Aldrich) à base de phénol (Ambion). Le protocole suivi consiste en une extraction phénol-chloroforme en présence de thiocyanate de guanidine, suivie d'une précipitation des acides nucléiques à l'isopropanol.

Le culot d'ARN extraits est remis en suspension dans de l'eau ultra pure, puis leur concentration et pureté sont simultanément estimées par le rapport des densités optiques mesurées à 230, 260 et 280 nm (Spectrophotomètre NanoDrop 1000 Thermo Fisher Scientific).

B. Transcription réverse (RT)

Les ARN extraits sont rétrotranscrits en ADNc à l'aide du kit *High-Capacity cDNA Reverse Transcription* en suivant les indications du fournisseur (ThermoFisher). Les milieux réactionnels, d'un volume final de 20 μ L, à partir de 2 μ g d'ARN, puis placés dans le Bio Rad iCycler (10 min 25°C, 120 min 37°C, 5 min 85°C). Les échantillons d'ADNc se conservent ensuite à -20°C.

C. PCR quantitative en temps réel (qPCR)

Les milieux réactionnels, d'un volume final de 20 μ L, se composent de 16 μ L de milieu réactionnel de la technologie choisie (Fast Start Universal Probe Master Mix pour les sondes Taqman ou Fast Start Universal SybrGreen Master Mix, Roche) et 4 μ L d'ADN (ADNc issu de la RT, ou ADN purifié). Après activation de la Taq polymérase 10 min à 95°C, l'ADN est amplifié par 40 cycles : 15 sec 95°C (dénaturation de l'ADN) ; 1 min 60°C (hybridation des amorces (**Tableau 3**) et polymérisation). Les résultats de la qPCR sont analysés selon la méthode de quantification $\Delta \Delta C_t$ par le logiciel *QuantStudio™ Design & Analysis* associé à l'appareil, puis traités sur Excel.

D. Western Blot

Les cellules des lignées K-562, NB-4 et Nalm-16 sont au préalable immunoprécipitées. Elles sont lysées avec un buffer contenant 50 mM Tris buffer pH 7.5, 150 mM NaCl, 2% NP40 et des inhibiteurs de protéases 1x (Roche Diagnostic). 2mg de protéines sont ensuite incubés overnight à 4°C, sous agitation, dans 1mL de tampon de lyse contenant 2µg d'anticorps anti-CDX2. 25µL de billes d'agarose fixées par des protéines G, préalablement lavées dans le tampon de lyse, sont ensuite ajoutées et incubées dans la suspension pendant 2H à 4°C. Le matériel immunoprécipité est ensuite élué.

Pour un extrait cellulaire total, les cellules d'intérêt sont lysées dans du tampon dénaturant Laemmli 1X (10mM Tris-HCl pH 6,5, 100 mM DTT, 1% SDS). Quarante µg de protéines totales sont utilisés pour le Western Blot. Les Western Blots sont ensuite effectués comme décrit précédemment (Gross et al., 2005), sur gels à 12% de polyacrylamide contenant du SDS.

E. Immunoprécipitation de chromatine (ChIP)

Les cellules (10^6) sont reprises dans 1mL de PBS 1X et fixées par ajout de formaldéhyde (1% final), pendant 10 min à température ambiante. Une solution de glycine (0,125 M final) est ajoutée pour une incubation de 5 min, pour arrêter le processus de réticulation. La chromatine est ensuite soniquée 5min. Les ChIPs sont effectués à l'aide du kit *EZ-Magna ChIP™ G Chromatin Immunoprecipitation* (Millipore). La chromatine réticulée est incubée overnight à 4°C avec 2 µg d'IgG de souris (Millipore) ou 2 µg d'anticorps approprié. Une fraction du lysat de chromatine réticulée et fragmentée est prélevée avant l'ajout de ces anticorps, pour servir d'« input ». Les inputs permettent de normaliser la quantité de fragments d'intérêts immunoprécipités par rapport à leur quantité totale dans le lysat cellulaire. A l'issue du ChIP, l'ADN est purifié à l'aide du kit *QIAquick PCR Purification* (Qiagen), et est élué dans un volume de 50 µL de tampon d'élution.

F. Immunoprécipitation de chromatine + séquençage (ChIP-seq)

Le protocole suivi est celui développé par Martianov et al., Scientific Report, 2016. La méthode comprend la préparation de la chromatine et le ChIP-seq à partir d'un volume de 400 -500 µL de culot cellulaire.

Les cellules HT-29, TC7, MKN-45, AGS, K-562 et NB-4 sont reprises dans un volume de 10 mL de PBS 1X, leur chromatine est fixée par ajout de formaldéhyde (0,4% final) pendant 10 min. La réticulation est arrêtée par ajout d'une solution de glycine (0.4 M final). Le culot cellulaire est collecté par centrifugation pendant 5 minutes à 500 g, lysé avec 5 volumes de tampon de lyse (50 mM Tris pH8, 10 mM EDTA, 1% SDS, inhibiteurs de protéases 1x) et soniqué pendant 10 min à l'aide du sonicateur Covaris E220. Un aliquot est prélevée, l'ADN purifié et quantifié. Cet aliquot constituera l'ADN « input ». La taille des fragments d'ADN fragmenté est évaluée sur gel d'agarose 1% (TBE). Une moyenne de taille autour de 400 pb est attendue.

La chromatine soniquée est pré-clarifiée à l'aide de billes d'agarose couplés à des protéines G (préalablement saturées en présence de BSA et de tRNA de levure). La chromatine (l'équivalent de 50 µg d'ADN) est incubée 12H à 4°C, sous agitation, en présence d'anticorps de lapin anti-H3K27ac (Active Motif) ou de lapin anti-H3K4me3 (clone MC315, Millipore). Les fragments de chromatine sont précipités à l'aide de nouvelles billes d'agarose couplées à des protéines G saturées, lavés et élués. L'ADN est purifié par une extraction phénol/chloroforme volume à volume, suivi d'une précipitation à l'isopropanol. L'ADN est élué dans 20 µL d'eau dépourvue de nucléases.

Les banques pour le séquençage ont été préparées et séquencées sur la plateforme de séquençage de l'IGBMC (GenomeEast, Illkirch) sur un Illumina Hi-seq2500 en « simple extrémité » et avec des « reads » de 50 bases. Après le séquençage, la détection des pics se fait à partir des fichiers .wig à l'aide du logiciel *Integrated Genome Browser* (IGB).

G. Assay for Transposase-Accessible Chromatin (ATAC)

Le protocole est adapté de celui développé par Buenrostro et al., 2015. Les cellules ($50 \cdot 10^3$ par point) sont lavées dans 50 μ L de PBS 1X puis re-suspendues dans 50 μ L de tampon de lyse (10 mM Tris-HCl pH 7.5 ; 10 mM NaCl ; 3mM MgCl₂ ; 0,1% NP40) froid. Le matériel est immédiatement centrifugé (10 min 500g 4°C). Chaque échantillon est repris individuellement dans une solution de transposition (kit *Nextera DNA Library 15028212*) : 22,5 μ L d'eau dépourvue de nucléases ; 25 μ L de tampon TD ; 2,5 μ L de transposase Tn5 (TDE1). Le matériel y est incubé pendant 1H à 37°C, sous agitation douce (balancier). L'ADN est ensuite purifié à l'aide du kit *Qiagen MinElute PCR Purification* (Qiagen), et est élué dans un volume de 10 μ L de tampon d'élution (10 mM Tris pH 8.0).

Les fragments d'ADN sont amplifiés par PCR, selon le mélange réactionnel suivant : 10 μ L d'ADN transposé ; 10 μ L d'eau dépourvue de nucléases ; 2,5 μ L d'amorces AD1.1 (25 μ M) ; 2,5 μ L d'amorces AD2.2 (25 μ M) (**Tableau 4**) ; 25 μ L *NEBNext High-Fidelity 2x* (PCR Master mix). Le programme PCR est le suivant : (5 min à 72°C ; 30 sec à 98°C) x1 (10 sec à 98°C ; 30 sec à 63°C ; 1 min à 72°C) x14. L'ADN est ensuite purifié à l'aide du kit *Qiagen MinElute PCR Purification*, et est élué dans 20 μ L de tampon d'élution (10 mM Tris-HCl pH 8.0). La qualité des fragments collectés est évaluée via une qPCR sur un aliquot de 2 μ L, à l'aide de sondes ciblant une région contrôle positive, et une région contrôle négative (**Tableau 2**).

Les banques pour le séquençage ont été séquencées sur la plateforme de séquençage de l'IGBMC (GenomeEast, Illkirch) sur un Illumina Hi-seq2500 en « simple extrémité » et avec des « reads » de 50 bases. Après le séquençage, la détection des pics se fait à partir des fichiers .wig à l'aide du logiciel *Integrated Genome Browser* (IGB).

H. Rapid Amplification of cDNA Ends (RACE)

La transcription réverse RT de l'ARN (1µg) en ADNc, puis l'amplification PCR des extrémités 5' des fragments d'ADNc s'effectuent à l'aide du kit *SMARTer® RACE 5'/3'* (Takara Bio USA), en suivant les indications du fournisseur. Les conditions de la RT sont les suivantes : (90min à 42°C) x 1 ; (10 min à 70°C) x 1. Les conditions de la PCR sont les suivantes : (30 sec à 94°C ; 30 sec à 68°C ; 3min à 72°C) x 25. Les fragments d'ADNc amplifiés sont ensuite séparés par migration électrophorétique sur gel à 1% d'agarose, en présence de Sybrgreen (cx10000). Les fragments d'ADNc de poids moléculaire d'intérêt sont extraits du gel, puis purifiés sur colonnes à l'aide du kit *QIAQuick Gel Extraction* (Qiagen). A 7 µL de l'élution de fragments d'ADNc sont ensuite ajoutés 1 µL Vecteurs pRACE et 2 µL de Master Mix In-Fusion HD fournis par le kit *SMARTer® RACE 5'/3'* (Takara Bio USA), de sorte à les intégrer dans des plasmides. Ces plasmides sont ensuite transformés dans des bactéries E. coli compétentes. Les bactéries sont étalées sur du milieu LB + antibiotiques, et incubées 24 H à 37°C. Les colonies s'étant développées sont amplifiées individuellement (~10 clones par lignée) 24h à 37°C sous agitation (240rpm). Les plasmides amplifiés sont extraits à l'aide du kit Qiagen Plasmid Midi Prep (Qiagen). 2 µL de plasmides resuspendus dans de l'eau dépourvue de nucléases sont ensuite ajoutés dans un mélange de 0,2 µL d'enzyme de restriction EcoRI, 0,2 µL d'enzyme HindIII, ainsi que 16 µL de réactifs G2O et 2 µL de Tampon Cutsmart fournis par le kit *SMARTer® RACE 5'/3'* (Takara Bio USA). Après une incubation de 1H à 37°C, l'ADNc est purifié à l'aide du kit Qiagen PCR Purification (Qiagen), et repris dans de l'eau dépourvue de nucléases. La concentration des échantillons est déterminée au spectrophotomètre comme décrit précédemment. Un aliquot de 20 µL de chacun d'entre eux, à la concentration ajustée entre 80-100 ng/µL, est transmise à la plateforme *Eurofins Genomics Germany* (Ebersberg, Allemagne) pour leur séquençage. L'analyse des séquences a été réalisée à l'aide du programme d'alignement de séquences de NCBI Blast2seq.

I. Circularised Chromosome Conformation Capture (4C)

Dix millions de cellules sont reprises dans 875µL de PBS 1X, puis fixées pendant 10min à température ambiante par ajout de formaldéhyde (2% final). De la Glycine (0,125M final) est ajoutée pour une incubation de 5min à température ambiante, pour arrêter la réticulation de la chromatine. Les cellules sont ensuite placées sur glace 15min, puis lavées 2 fois avec du PBS 1X. Les cellules sont lysées dans 1mL de tampon de lyse (10mM Tris-HCl pH8, 10mM NaCl, 0,2% NP40, inhibiteurs de protéases 1X), pendant 1H à 4°C sur portoir rotatif.

Les noyaux récupérés sont lavés dans le tampon de l'enzyme *DpnII*, puis perméabilisés 1H à 37°C, sous agitation (1000rpm), avec 1% de SDS final dans la solution. La suspension de noyaux est scindée en 2 aliquots, puis ils sont à nouveau incubés 1H à 37°C, sous agitation (1000rpm), dans le même tampon contenant cette fois 3,3% de Triton. Une fois l'ADN remis en suspension dans 500 µL de tampon, 1500U d'enzyme *DpnII* sont ajoutés pour procéder à la première digestion enzymatique, qui s'opère overnight, à 37°C, sous agitations (800rpm).

L'ADN est lavé dans le tampon de la T4 DNA Ligase, puis incubé dans ce même tampon avec 10 000 U de l'enzyme une nuit à 16°, sous agitation (400rpm), pour effectuer la première ligation. La T4 DNA ligase est inactivée par ajout de 10µg de protéinase K et incubation une nuit à 65°C, sous agitation (750rpm), puis par ajout de 200µg de RNase A pour une incubation 1H à 37°C, sous agitation (750rpm). L'ADN est ensuite purifié par une extraction phénol/chloroforme volume à volume, et une précipitation à l'isopropanol en présence de 300mM d'acétate d'Ammonium. Les culots sont resuspendus dans 50uL d'eau dépourvue de nucléase (Ambion). La concentration et la pureté de l'ADN sont estimées au spectrophotomètre comme décrit précédemment.

L'ADN est alors dilué à 100ng/µL à l'aide du tampon de l'enzyme *Csp6I*, puis 5U d'enzyme par µg d'ADN sont ajoutés. Les échantillons sont incubés une nuit à 37°C, sous agitation (800rpm), il s'agit de la deuxième digestion enzymatique. L'ADN est à nouveau purifié et quantifié comme précédemment. Le culot est repris dans 50µL d'une solution à 10mM Tris-HCl pH 8 et 0,1mM EDTA. La concentration en ADN est à nouveau quantifiée au spectrophotomètre.

L'ADN est enfin dilué à 5ng/μL dans le tampon de la T4 DNA Ligase, auquel est ajouté 200 U de l'enzyme, pour une incubation overnight à 16°C, sous agitation (400 rpm). A l'issue de cette deuxième ligation, l'ADN est purifié et quantifié comme précédemment.

La préparation des banques de 4C s'effectue sur 600ng de matériel. L'ADN est amplifié par 1,75U de la polymérase du kit *Expand™ Long Template PCR System* (Roche). Les milieux réactionnels, d'un volume final de 150μL, contiennent 15 μL de tampon PCR, 6 μL de chaque dNTP (100 mM), et 1,5 μL d'amorces sens et antisens attribuées à chaque lignée (**Tableau 5**). Le mélange est scindé en 6 réactions, qui seront regroupées à l'issue de la PCR. Le programme PCR est le suivant : (2 min à 94°C) x1 ; (15 sec à 94°C ; 1 min à 55°C ; 3 min à 68°C) x27 ; (7 min à 68°C) x1. Les échantillons sont ensuite quantifiés par spectrophotométrie, puis, dans notre cas ont été transmis à la plateforme de séquençage de l'IGBMC. Chaque étape de digestion et ligation est contrôlée par dépôt sur gel d'agarose (1%) d'un aliquot (~1/10^e de la suspension).

BIBLIOGRAPHIE

1. Moore MA, Metcalf D. Ontogeny of the haemopoietic system: yolk sac origin of in vivo and in vitro colony forming cells in the developing mouse embryo. *Br J Haematol.* 1970;18(3):279-296. doi:10.1111/j.1365-2141.1970.tb01443.x
2. Palis J, Robertson S, Kennedy M, Wall C, Keller G. Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development.* 1999;126(22):5073-5084. doi:10.1242/dev.126.22.5073
3. Tober J, Koniski A, McGrath KE, et al. The megakaryocyte lineage originates from hemangioblast precursors and is an integral component both of primitive and of definitive hematopoiesis. *Blood.* 2007;109(4):1433-1441. doi:10.1182/blood-2006-06-031898
4. Canu G, Ruhrberg C. First blood: the endothelial origins of hematopoietic progenitors. *Angiogenesis.* 2021;24(2):199-211. doi:10.1007/s10456-021-09783-9
5. McGrath KE, Frame JM, Fegan KH, et al. Distinct sources of hematopoietic progenitors emerge before HSCs and provide functional blood cells in the mammalian embryo. *Cell Rep.* 2015;11(12):1892-1904. doi:10.1016/j.celrep.2015.05.036
6. Kasaai B, Caolo V, Peacock HM, et al. Erythro-myeloid progenitors can differentiate from endothelial cells and modulate embryonic vascular remodeling. *Sci Rep.* 2017;7(1):43817. doi:10.1038/srep43817
7. Oberlin E, Hafny BE, Petit-Cocault L, Souyri M. Definitive human and mouse hematopoiesis originates from the embryonic endothelium: a new class of HSCs based on VE-cadherin expression. *Int J Dev Biol.* 2010;54(6-7):1165-1173. doi:10.1387/ijdb.103121eo
8. Boisset JC, van Cappellen W, Andrieu-Soler C, Galjart N, Dzierzak E, Robin C. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature.* 2010;464(7285):116-120. doi:10.1038/nature08764
9. Medvinsky A, Rybtsov S, Taoudi S. Embryonic origin of the adult hematopoietic system: advances and questions. *Development.* 2011;138(6):1017-1031. doi:10.1242/dev.040998
10. Julien E, El Omar R, Tavian M. Origin of the hematopoietic system in the human embryo. *FEBS Lett.* 2016;590(22):3987-4001. doi:10.1002/1873-3468.12389
11. Bertrand JY, Chi NC, Santoso B, Teng S, Stainier DYR, Traver D. Hematopoietic stem cells derive directly from aortic endothelium during development. *Nature.* 2010;464(7285):108-111. doi:10.1038/nature08738
12. Ema H, Nakauchi H. Expansion of hematopoietic stem cells in the developing liver of a mouse embryo. *Blood.* 2000;95(7):2284-2288. doi:10.1182/blood.V95.7.2284
13. Morita Y, Iseki A, Okamura S, Suzuki S, Nakauchi H, Ema H. Functional characterization of hematopoietic stem cells in the spleen. *Exp Hematol.* 2011;39(3):351-359.e3. doi:10.1016/j.exphem.2010.12.008
14. Christensen JL, Wright DE, Wagers AJ, Weissman IL. Circulation and Chemotaxis of Fetal Hematopoietic Stem Cells. *PLoS Biol.* 2004;2(3):e75. doi:10.1371/journal.pbio.0020075
15. Laurenti E, Göttgens B. From haematopoietic stem cells to complex differentiation landscapes. *Nature.* 2018;553(7689):418-426. doi:10.1038/nature25022
16. Liu C, Saito F, Liu Z, et al. Coordination between CCR7- and CCR9-mediated chemokine signals in prevascular fetal thymus colonization. *Blood.* 2006;108(8):2531-2539. doi:10.1182/blood-2006-05-024190
17. Douagi I, André I, Ferraz JC, Cumano A. Characterization of T cell precursor activity in the murine fetal thymus: evidence for an input of T cell precursors between days 12 and 14 of gestation. *European Journal of Immunology.* 2000;30(8):2201-2210. doi:10.1002/1521-4141(2000)30:8<2201::AID-IMMU2201>3.0.CO;2-2

18. Seita J, Weissman IL. Hematopoietic stem cell: self-renewal versus differentiation: Hematopoietic stem cell. *WIREs Syst Biol Med*. 2010;2(6):640-653. doi:10.1002/wsbm.86
19. Notta F, Doulatov S, Lauretti E, Poeppl A, Jurisica I, Dick JE. Isolation of single human hematopoietic stem cells capable of long-term multilineage engraftment. *Science*. 2011;333(6039):218-221. doi:10.1126/science.1201219
20. Morrison SJ, Wandycz AM, Hemmati HD, Wright DE, Weissman IL. Identification of a lineage of multipotent hematopoietic progenitors. *Development*. 1997;124(10):1929-1939.
21. Morrison SJ, Uchida N, Weissman IL. The biology of hematopoietic stem cells. *Annu Rev Cell Dev Biol*. 1995;11:35-71. doi:10.1146/annurev.cb.11.110195.000343
22. Lee J, Zhou YJ, Ma W, et al. Lineage specification of human dendritic cell is marked by expression of the transcription factor IRF8 in HSCs and MPPs. *Nat Immunol*. 2017;18(8):877-888. doi:10.1038/ni.3789
23. Notta F, Zandi S, Takayama N, et al. DISTINCT ROUTES OF LINEAGE DEVELOPMENT RESHAPE THE HUMAN BLOOD HIERARCHY ACROSS ONTOGENY. *Science*. 2016;351(6269):aab2116. doi:10.1126/science.aab2116
24. Velten L, Haas SF, Raffel S, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol*. 2017;19(4):271-281. doi:10.1038/ncb3493
25. Jacobsen SEW, Nerlov C. Haematopoiesis in the era of advanced single-cell technologies. *Nat Cell Biol*. 2019;21(1):2-8. doi:10.1038/s41556-018-0227-8
26. Gekas C, Graf T. CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age. *Blood*. 2013;121(22):4463-4472. doi:10.1182/blood-2012-09-457929
27. Shimazu T, Iida R, Zhang Q, et al. CD86 is expressed on murine hematopoietic stem cells and denotes lymphopoietic potential. *Blood*. 2012;119(21):4889-4897. doi:10.1182/blood-2011-10-388736
28. Pietras EM, Reynaud D, Kang YA, et al. Functionally distinct subsets of lineage-biased multipotent progenitors control blood production in normal and regenerative conditions. *Cell Stem Cell*. 2015;17(1):35-46. doi:10.1016/j.stem.2015.05.003
29. Wilson NK, Foster SD, Wang X, et al. Combinatorial Transcriptional Control In Blood Stem/Progenitor Cells: Genome-wide Analysis of Ten Major Transcriptional Regulators. *Cell Stem Cell*. 2010;7(4):532-544. doi:10.1016/j.stem.2010.07.016
30. Ding Y, Liu Z, Liu F. Transcriptional and epigenetic control of hematopoietic stem cell fate decisions in vertebrates. *Dev Biol*. 2021;475:156-164. doi:10.1016/j.ydbio.2021.03.003
31. Stanley ER. Lineage Commitment: Cytokines Instruct, At Last! *Cell Stem Cell*. 2009;5(3):234-236. doi:10.1016/j.stem.2009.08.015
32. Sanjuan-Pla A, Macaulay IC, Jensen CT, et al. Platelet-biased stem cells reside at the apex of the haematopoietic stem-cell hierarchy. *Nature*. 2013;502(7470):232-236. doi:10.1038/nature12495
33. Mossadegh-Keller N, Sarrazin S, Kandalla PK, et al. M-CSF instructs myeloid lineage fate in single haematopoietic stem cells. *Nature*. 2013;497(7448):239-243. doi:10.1038/nature12026
34. Kondo M, Scherer DC, Miyamoto T, et al. Cell-fate conversion of lymphoid-committed progenitors by instructive actions of cytokines. *Nature*. 2000;407(6802):383-386. doi:10.1038/35030112
35. Brown G, Tsapogas P, Ceredig R. The changing face of hematopoiesis: a spectrum of options is available to stem cells. *Immunology & Cell Biology*. 2018;96(9):898-911. doi:10.1111/imcb.12055

36. Travlos GS. Normal Structure, Function, and Histology of the Bone Marrow. *Toxicol Pathol.* 2006;34(5):548-565. doi:10.1080/01926230600939856
37. Lucas D. Structural organization of the bone marrow and its role in hematopoiesis. *Curr Opin Hematol.* 2021;28(1):36-42. doi:10.1097/MOH.0000000000000621
38. Chen JY, Miyanishi M, Wang SK, et al. Hoxb5 marks long-term haematopoietic stem cells and reveals a homogenous perivascular niche. *Nature.* 2016;530(7589):223-227. doi:10.1038/nature16943
39. Pinho S, Frenette PS. Haematopoietic stem cell activity and interactions with the niche. *Nat Rev Mol Cell Biol.* 2019;20(5):303-320. doi:10.1038/s41580-019-0103-9
40. Yamazaki S, Ema H, Karlsson G, et al. Nonmyelinating Schwann Cells Maintain Hematopoietic Stem Cell Hibernation in the Bone Marrow Niche. *Cell.* 2011;147(5):1146-1158. doi:10.1016/j.cell.2011.09.053
41. Brown G. Hematopoietic Stem Cells: Nature and Niche Nurture. *Bioengineering (Basel).* 2021;8(5):67. doi:10.3390/bioengineering8050067
42. Comazzetto S, Shen B, Morrison SJ. Niches that regulate stem cells and hematopoiesis in adult bone marrow. *Dev Cell.* 2021;56(13):1848-1860. doi:10.1016/j.devcel.2021.05.018
43. Méndez-Ferrer S, Lucas D, Battista M, Frenette PS. Haematopoietic stem cell release is regulated by circadian oscillations. *Nature.* 2008;452(7186):442-447. doi:10.1038/nature06685
44. Gao X, Zhang D, Xu C, Li H, Caron KM, Frenette PS. Nociceptive nerves regulate haematopoietic stem cell mobilization. *Nature.* 2021;589(7843):591-596. doi:10.1038/s41586-020-03057-y
45. Upadhaya S, Krichevsky O, Akhmetzyanova I, Sawai CM, Fooksman DR, Reizis B. Intravital imaging reveals motility of adult hematopoietic stem cells in the bone marrow niche. *Cell Stem Cell.* 2020;27(2):336-345.e4. doi:10.1016/j.stem.2020.06.003
46. Ding L, Morrison SJ. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature.* 2013;495(7440):231-235. doi:10.1038/nature11885
47. Shen B, Tasdogan A, Ubellacker JM, et al. A mechanosensitive peri-arteriolar niche for osteogenesis and lymphopoiesis. *Nature.* 2021;591(7850):438-444. doi:10.1038/s41586-021-03298-5
48. Zhang J, Wu Q, Johnson CB, et al. In situ mapping identifies distinct vascular niches for myelopoiesis. *Nature.* 2021;590(7846):457-462. doi:10.1038/s41586-021-03201-2
49. Comazzetto S, Murphy MM, Berto S, Jeffery E, Zhao Z, Morrison SJ. Restricted Hematopoietic Progenitors and Erythropoiesis Require SCF from Leptin Receptor+ Niche Cells in the Bone Marrow. *Cell Stem Cell.* 2019;24(3):477-486.e6. doi:10.1016/j.stem.2018.11.022
50. Hur J, Choi JI, Lee H, et al. CD82/KAI1 Maintains the Dormancy of Long-Term Hematopoietic Stem Cells through Interaction with DARC-Expressing Macrophages. *Cell Stem Cell.* 2016;18(4):508-521. doi:10.1016/j.stem.2016.01.013
51. Chow A, Lucas D, Hidalgo A, et al. Bone marrow CD169+ macrophages promote the retention of hematopoietic stem and progenitor cells in the mesenchymal stem cell niche. *J Exp Med.* 2011;208(2):261-271. doi:10.1084/jem.20101688
52. Bruns I, Lucas D, Pinho S, et al. Megakaryocytes regulate hematopoietic stem cell quiescence through CXCL4 secretion. *Nat Med.* 2014;20(11):1315-1320. doi:10.1038/nm.3707
53. Zhao M, Perry JM, Marshall H, et al. Megakaryocytes maintain homeostatic quiescence and promote post-injury regeneration of hematopoietic stem cells. *Nat Med.* 2014;20(11):1321-1326. doi:10.1038/nm.3706

54. Schepers K, Pietras EM, Reynaud D, et al. Myeloproliferative Neoplasia Remodels the Endosteal Bone Marrow Niche into a Self-Reinforcing Leukemic Niche. *Cell Stem Cell*. 2013;13(3):285-299. doi:10.1016/j.stem.2013.06.009
55. Kumar B, Garcia M, Weng L, et al. Acute myeloid leukemia transforms the bone marrow niche into a leukemia-permissive microenvironment through exosome secretion. *Leukemia*. 2018;32(3):575-587. doi:10.1038/leu.2017.259
56. Blau O, Hofmann WK, Baldus CD, et al. Chromosomal aberrations in bone marrow mesenchymal stroma cells from patients with myelodysplastic syndrome and acute myeloblastic leukemia. *Experimental Hematology*. 2007;35(2):221-229. doi:10.1016/j.exphem.2006.10.012
57. Blau O, Baldus CD, Hofmann WK, et al. Mesenchymal stromal cells of myelodysplastic syndrome and acute myeloid leukemia patients have distinct genetic abnormalities compared with leukemic blasts. *Blood*. 2011;118(20):5583-5592. doi:10.1182/blood-2011-03-343467
58. Kim JA, Shim JS, Lee GY, et al. Microenvironmental remodeling as a parameter and prognostic factor of heterogeneous leukemogenesis in acute myelogenous leukemia. *Cancer Res*. 2015;75(11):2222-2231. doi:10.1158/0008-5472.CAN-14-3379
59. Kode A, Manavalan JS, Mosialou I, et al. Leukemogenesis Induced by an Activating β -catenin mutation in Osteoblasts. *Nature*. 2014;506(7487):240-244. doi:10.1038/nature12883
60. Raaijmakers MHGP, Mukherjee S, Guo S, et al. Bone progenitor dysfunction induces myelodysplasia and secondary leukemia. *Nature*. 2010;464(7290):852-857. doi:10.1038/nature08851
61. Walkley CR, Shea JM, Sims NA, Purton LE, Orkin SH. pRb Extrinsically Regulates Hematopoietic Stem Cells via Myeloid Cell -Bone Marrow Microenvironment Interactions. *Cell*. 2007;129(6):1081-1095. doi:10.1016/j.cell.2007.03.055
62. Inra CN, Zhou BO, Acar M, et al. A perisinusoidal niche for extramedullary haematopoiesis in the spleen. *Nature*. 2015;527(7579):466-471. doi:10.1038/nature15530
63. Dutta P, Hoyer FF, Grigoryeva LS, et al. Macrophages retain hematopoietic stem cells in the spleen via VCAM-1. *J Exp Med*. 2015;212(4):497-512. doi:10.1084/jem.20141642
64. Kampen KR. The discovery and early understanding of leukemia. *Leukemia Research*. 2012;36(1):6-13. doi:10.1016/j.leukres.2011.09.028
65. Bennett JH. On Leucocythemia, or Blood Containing an Unusual Number of Colourless Corpuscles. *Mon J Med Sci*. 1851;3(13):17-38.
66. Kay AB. Paul Ehrlich and the Early History of Granulocytes. *Microbiology Spectrum*. 2016;4(4):4.4.62. doi:10.1128/microbiolspec.MCHD-0032-2016
67. Piller GJ. Leukaemia – a brief historical review from ancient times to 1950. *British Journal of Haematology*. 2001;112(2):282-292. doi:10.1046/j.1365-2141.2001.02411.x
68. Bennett JM, Catovsky D, Daniel MT, et al. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br J Haematol*. 1976;33(4):451-458. doi:10.1111/j.1365-2141.1976.tb03563.x
69. Vardiman JW, Harris NL, Brunning RD. The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood*. 2002;100(7):2292-2302. doi:10.1182/blood-2002-04-1199

70. Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009;114(5):937-951. doi:10.1182/blood-2009-03-209262
71. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391-2405. doi:10.1182/blood-2016-03-643544
72. Estey E. Acute myeloid leukemia: 2016 Update on risk-stratification and management. *Am J Hematol*. 2016;91(8):824-846. doi:10.1002/ajh.24439
73. Kishtagari A, Levine RL, Viny AD. Driver mutations in acute myeloid leukemia. *Curr Opin Hematol*. 2020;27(2):49-57. doi:10.1097/MOH.0000000000000567
74. Kantarjian H, Kadia T, DiNardo C, et al. Acute myeloid leukemia: current progress and future directions. *Blood Cancer J*. 2021;11(2):1-25. doi:10.1038/s41408-021-00425-3
75. DiNardo CD, Ravandi F, Agresta S, et al. Characteristics, clinical outcome, and prognostic significance of IDH mutations in AML. *Am J Hematol*. 2015;90(8):732-736. doi:10.1002/ajh.24072
76. Takeda A, Goolsby C, Yaseen NR. NUP98-HOXA9 induces long-term proliferation and blocks differentiation of primary human CD34+ hematopoietic cells. *Cancer Res*. 2006;66(13):6628-6637. doi:10.1158/0008-5472.CAN-06-0458
77. Li B, Huang Q, Wei GH. The Role of HOX Transcription Factors in Cancer Predisposition and Progression. *Cancers (Basel)*. 2019;11(4):E528. doi:10.3390/cancers11040528
78. Chopra M, Bohlander SK. The cell of origin and the leukemia stem cell in acute myeloid leukemia. *Genes Chromosomes Cancer*. 2019;58(12):850-858. doi:10.1002/gcc.22805
79. Ding L, Ley TJ, Larson DE, et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*. 2012;481(7382):506-510. doi:10.1038/nature10738
80. Morita K, Wang F, Jahn K, et al. Clonal evolution of acute myeloid leukemia revealed by high-throughput single-cell genomics. *Nat Commun*. 2020;11(1):5327. doi:10.1038/s41467-020-19119-8
81. Stieglitz E, Loh ML. Genetic predispositions to childhood leukemia. *Ther Adv Hematol*. 2013;4(4):270-290. doi:10.1177/2040620713498161
82. Porter CC. Germ line mutations associated with leukemias. *Hematology Am Soc Hematol Educ Program*. 2016;2016(1):302-308.
83. Maloney KW, Taub JW, Ravindranath Y, Roberts I, Vyas P. Down syndrome preleukemia and leukemia. *Pediatr Clin North Am*. 2015;62(1):121-137. doi:10.1016/j.pcl.2014.09.009
84. Alter BP. Fanconi anemia and the development of leukemia. *Best Pract Res Clin Haematol*. 2014;27(0):214-221. doi:10.1016/j.beha.2014.10.002
85. Smith MA, McCaffrey RP, Karp JE. The Secondary Leukemias: Challenges and Research Directions. *JNCI: Journal of the National Cancer Institute*. 1996;88(7):407-418. doi:10.1093/jnci/88.7.407
86. Kuykendall A, Duployez N, Boissel N, Lancet JE, Welch JS. Acute Myeloid Leukemia: The Good, the Bad, and the Ugly. *Am Soc Clin Oncol Educ Book*. 2018;38:555-573. doi:10.1200/EDBK_199519
87. Goldstein BD, Smith MT. Benzene and haematological cancers. In: Baan RA, Stewart BW, Straif K, eds. *Tumour Site Concordance and Mechanisms of Carcinogenesis*. IARC Scientific Publications. International Agency for Research on Cancer; 2019. Accessed August 26, 2021. <http://www.ncbi.nlm.nih.gov/books/NBK570338/>

88. Kang DS, Kim HS, Jung JH, Lee CM, Ahn YS, Seo YR. Formaldehyde exposure and leukemia risk: a comprehensive review and network-based toxicogenomic approach. *Genes and Environment*. 2021;43(1):13. doi:10.1186/s41021-021-00183-5
89. Hatano Y, Ideta T, Hirata A, et al. Virus-Driven Carcinogenesis. *Cancers (Basel)*. 2021;13(11):2625. doi:10.3390/cancers13112625
90. Manivannan AC, Devaraju V, Velmurugan P, et al. Tumorigenesis and diagnostic practice applied in two oncogenic viruses: Epstein Barr virus and T-cell lymphotropic virus-1—Mini review. *Biomedicine & Pharmacotherapy*. 2021;142:111974. doi:10.1016/j.biopha.2021.111974
91. Tebbi CK. Etiology of Acute Leukemia: A Review. *Cancers (Basel)*. 2021;13(9):2256. doi:10.3390/cancers13092256
92. Silva P, Neumann M, Schroeder MP, et al. Acute myeloid leukemia in the elderly is characterized by a distinct genetic and epigenetic landscape. *Leukemia*. 2017;31(7):1640-1644. doi:10.1038/leu.2017.109
93. Hoffmann D, Hoffmann I. The changing cigarette, 1950-1995. *J Toxicol Environ Health*. 1997;50(4):307-364. doi:10.1080/009841097160393
94. Fircanis S, Merriam P, Khan N, Castillo JJ. The relation between cigarette smoking and risk of acute myeloid leukemia: An updated meta-analysis of epidemiological studies. *American Journal of Hematology*. 2014;89(8):E125-E132. doi:10.1002/ajh.23744
95. Poynter JN, Richardson M, Blair CK, et al. Obesity over the Life Course and Risk of Acute Myeloid Leukemia and Myelodysplastic Syndromes. *Cancer Epidemiol*. 2016;40:134-140. doi:10.1016/j.canep.2015.12.005
96. Dhakal P, Lyden E, Lee A, et al. Effects of Obesity on Overall Survival of Adults With Acute Myeloid Leukemia. *Clin Lymphoma Myeloma Leuk*. 2020;20(3):e131-e136. doi:10.1016/j.clml.2019.11.001
97. Tomaszewski EL, Fickley CE, Maddux L, et al. The Patient Perspective on Living with Acute Myeloid Leukemia. *Oncol Ther*. 2016;4(2):225-238. doi:10.1007/s40487-016-0029-8
98. Bullinger L, Döhner K, Döhner H. Genomics of Acute Myeloid Leukemia Diagnosis and Pathways. *J Clin Oncol*. 2017;35(9):934-946. doi:10.1200/JCO.2016.71.2208
99. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424-447. doi:10.1182/blood-2016-08-733196
100. Lin TL, Pagano L. The important role of intensive induction chemotherapy in the treatment of acute myeloid leukemia. *Expert Rev Hematol*. 2021;14(3):303-314. doi:10.1080/17474086.2021.1886920
101. Heuser M, Ofran Y, Boissel N, et al. Acute myeloid leukaemia in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up†. *Annals of Oncology*. 2020;31(6):697-712. doi:10.1016/j.annonc.2020.02.018
102. Mayer LD, Tardi P, Louie AC. CPX-351: a nanoscale liposomal co-formulation of daunorubicin and cytarabine with unique biodistribution and tumor cell uptake properties. *Int J Nanomedicine*. 2019;14:3819-3830. doi:10.2147/IJN.S139450
103. Wei AH, Döhner H, Pocock C, et al. Oral Azacitidine Maintenance Therapy for Acute Myeloid Leukemia in First Remission. *New England Journal of Medicine*. 2020;383(26):2526-2537. doi:10.1056/NEJMoa2004444
104. Loke J, Buka R, Craddock C. Allogeneic Stem Cell Transplantation for Acute Myeloid Leukemia: Who, When, and How? *Front Immunol*. 2021;12:659595. doi:10.3389/fimmu.2021.659595

105. Short NJ, Konopleva M, Kadia TM, et al. Advances in the Treatment of Acute Myeloid Leukemia: New Drugs and New Challenges. *Cancer Discov.* 2020;10(4):506-525. doi:10.1158/2159-8290.CD-19-1011
106. Liu Y, Bewersdorf JP, Stahl M, Zeidan AM. Immunotherapy in acute myeloid leukemia and myelodysplastic syndromes: The dawn of a new era? *Blood Rev.* 2019;34:67-83. doi:10.1016/j.blre.2018.12.001
107. Allen C, Zeidan AM, Bewersdorf JP. BiTEs, DARTS, BiKEs and TriKEs—Are Antibody Based Therapies Changing the Future Treatment of AML? *Life (Basel).* 2021;11(6):465. doi:10.3390/life11060465
108. Köhler N, Ruess DA, Kesselring R, Zeiser R. The Role of Immune Checkpoint Molecules for Relapse After Allogeneic Hematopoietic Cell Transplantation. *Front Immunol.* 2021;12:634435. doi:10.3389/fimmu.2021.634435
109. Sendker S, Reinhardt D, Niktoreh N. Redirecting the Immune Microenvironment in Acute Myeloid Leukemia. *Cancers (Basel).* 2021;13(6):1423. doi:10.3390/cancers13061423
110. Zhou K, Gaullier G, Luger K. Nucleosome structure and dynamics are coming of age. *Nat Struct Mol Biol.* 2019;26(1):3-13. doi:10.1038/s41594-018-0166-x
111. Olins DE, Olins AL. Chromatin history: our view from the bridge. *Nat Rev Mol Cell Biol.* 2003;4(10):809-814. doi:10.1038/nrm1225
112. Li G, Zhu P. Structure and organization of chromatin fiber in the nucleus. *FEBS Letters.* 2015;589(20, Part A):2893-2904. doi:10.1016/j.febslet.2015.04.023
113. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell.* 1999;99(3):247-257. doi:10.1016/s0092-8674(00)81656-6
114. Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. Dnmt3L and the establishment of maternal genomic imprints. *Science.* 2001;294(5551):2536-2539. doi:10.1126/science.1065848
115. Sharif J, Muto M, Takebayashi S ichiro, et al. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature.* 2007;450(7171):908-912. doi:10.1038/nature06397
116. He YF, Li BZ, Li Z, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science.* 2011;333(6047):1303-1307. doi:10.1126/science.1210944
117. Ito S, Shen L, Dai Q, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science.* 2011;333(6047):1300-1303. doi:10.1126/science.1210597
118. Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet.* 2017;18(9):517-534. doi:10.1038/nrg.2017.33
119. Bird AP. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Research.* 1980;8(7):1499-1504. doi:10.1093/nar/8.7.1499
120. Schulz WA, Steinhoff C, Florl AR. Methylation of endogenous human retroelements in health and disease. *Curr Top Microbiol Immunol.* 2006;310:211-250. doi:10.1007/3-540-31181-5_11
121. Neri F, Rapelli S, Krepelova A, et al. Intragenic DNA methylation prevents spurious transcription initiation. *Nature.* 2017;543(7643):72-77. doi:10.1038/nature21373
122. Du Q, Luu PL, Stirzaker C, Clark SJ. Methyl-CpG-binding domain proteins: readers of the epigenome. *Epigenomics.* 2015;7(6):1051-1073. doi:10.2217/epi.15.39

123. Mahmood N, Rabbani SA. DNA Methylation Readers and Cancer: Mechanistic and Therapeutic Applications. *Front Oncol.* 2019;9:489. doi:10.3389/fonc.2019.00489
124. Nan X, Ng HH, Johnson CA, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature.* 1998;393(6683):386-389. doi:10.1038/30764
125. Zhang Y, Sun Z, Jia J, et al. Overview of Histone Modification. *Adv Exp Med Biol.* 2021;1283:1-16. doi:10.1007/978-981-15-8104-5_1
126. Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. *Cell Res.* 2011;21(4):564-578. doi:10.1038/cr.2011.42
127. Marmorstein R, Zhou MM. Writers and Readers of Histone Acetylation: Structure, Mechanism, and Inhibition. *Cold Spring Harb Perspect Biol.* 2014;6(7):a018762. doi:10.1101/cshperspect.a018762
128. Margueron R, Justin N, Ohno K, et al. Role of the polycomb protein Eed in the propagation of repressive histone marks. *Nature.* 2009;461(7265):762-767. doi:10.1038/nature08398
129. Santos-Rosa H, Schneider R, Bannister AJ, et al. Active genes are tri-methylated at K4 of histone H3. *Nature.* 2002;419(6905):407-411. doi:10.1038/nature01080
130. Zhao Q, Zhang J, Chen R, et al. Dissecting the precise role of H3K9 methylation in crosstalk with DNA maintenance methylation in mammals. *Nat Commun.* 2016;7(1):12464. doi:10.1038/ncomms12464
131. Bian C, Xu C, Ruan J, et al. Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *EMBO J.* 2011;30(14):2829-2842. doi:10.1038/emboj.2011.193
132. Zhang T, Cooper S, Brockdorff N. The interplay of histone modifications - writers that read. *EMBO Rep.* 2015;16(11):1467-1481. doi:10.15252/embr.201540945
133. Li Y, Chen X, Lu C. The interplay between DNA and histone methylation: molecular mechanisms and disease implications. *EMBO Rep.* 2021;22(5):e51803. doi:10.15252/embr.202051803
134. Jiao L, Liu X. Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science.* 2015;350(6258):aac4383. doi:10.1126/science.aac4383
135. Reddington JP, Perricone SM, Nestor CE, et al. Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. *Genome Biol.* 2013;14(3):R25. doi:10.1186/gb-2013-14-3-r25
136. Voigt P, Tee WW, Reinberg D. A double take on bivalent promoters. *Genes Dev.* 2013;27(12):1318-1338. doi:10.1101/gad.219626.113
137. Kumar D, Cinghu S, Oldfield AJ, Yang P, Jothi R. Decoding the function of bivalent chromatin in development and cancer. *Genome Res.* Published online October 19, 2021:gr.275736.121. doi:10.1101/gr.275736.121
138. Alder O, Laval F, Helness A, et al. Ring1B and Suv39h1 delineate distinct chromatin states at bivalent genes during early mouse lineage commitment. *Development.* 2010;137(15):2483-2492. doi:10.1242/dev.048363
139. Sun X, Yang Q, Rogers CJ, Du M, Zhu MJ. AMPK improves gut epithelial differentiation and barrier function via regulating Cdx2 expression. *Cell Death Differ.* 2017;24(5):819-831. doi:10.1038/cdd.2017.14
140. Buck MJ, Lieb JD. A chromatin-mediated mechanism for specification of conditional transcription factor targets. *Nat Genet.* 2006;38(12):1446-1451. doi:10.1038/ng1917

141. Boyle AP, Song L, Lee BK, et al. High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells. *Genome Res.* 2011;21(3):456-464. doi:10.1101/gr.112656.110
142. Zaret KS, Mango S. Pioneer Transcription Factors, Chromatin Dynamics, and Cell Fate Control. *Curr Opin Genet Dev.* 2016;37:76-81. doi:10.1016/j.gde.2015.12.003
143. Morrison O, Thakur J. Molecular Complexes at Euchromatin, Heterochromatin and Centromeric Chromatin. *International Journal of Molecular Sciences.* 2021;22(13):6922. doi:10.3390/ijms22136922
144. Schaffner W. Enhancers, enhancers – from their discovery to today's universe of transcription enhancers. *Biological Chemistry.* 2015;396(4):311-327. doi:10.1515/hsz-2014-0303
145. Peng Y, Zhang Y. Enhancer and super-enhancer: Positive regulators in gene transcription. *Animal Model Exp Med.* 2018;1(3):169-179. doi:10.1002/ame2.12032
146. Moorthy SD, Davidson S, Shchuka VM, et al. Enhancers and super-enhancers have an equivalent regulatory role in embryonic stem cells through regulation of single or multiple genes. *Genome Res.* 2017;27(2):246-258. doi:10.1101/gr.210930.116
147. He B, Chen C, Teng L, Tan K. Global view of enhancer-promoter interactome in human cells. *PNAS.* 2014;111(21):E2191-E2199. doi:10.1073/pnas.1320308111
148. Sexton T, Cavalli G. The role of chromosome domains in shaping the functional genome. *Cell.* 2015;160(6):1049-1059. doi:10.1016/j.cell.2015.02.040
149. Lewis MW, Li S, Franco HL. Transcriptional control by enhancers and enhancer RNAs. *Transcription.* 2019;10(4-5):171-186. doi:10.1080/21541264.2019.1695492
150. Arnold PR, Wells AD, Li XC. Diversity and Emerging Roles of Enhancer RNA in Regulation of Gene Expression and Cell Fate. *Frontiers in Cell and Developmental Biology.* 2020;7:377. doi:10.3389/fcell.2019.00377
151. Szabo Q, Bantignies F, Cavalli G. Principles of genome folding into topologically associating domains. *Sci Adv.* 2019;5(4):eaaw1668. doi:10.1126/sciadv.aaw1668
152. Gruber S, Haering CH, Nasmyth K. Chromosomal Cohesin Forms a Ring. *Cell.* 2003;112(6):765-777. doi:10.1016/S0092-8674(03)00162-4
153. Ohlsson R, Renkawitz R, Lobanenko V. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet.* 2001;17(9):520-527. doi:10.1016/s0168-9525(01)02366-6
154. Nora EP, Goloborodko A, Valton AL, et al. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell.* 2017;169(5):930-944.e22. doi:10.1016/j.cell.2017.05.004
155. Rao SSP, Huntley MH, Durand NC, et al. A three-dimensional map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell.* 2014;159(7):1665-1680. doi:10.1016/j.cell.2014.11.021
156. Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* 2016;15(9):2038-2049. doi:10.1016/j.celrep.2016.04.085
157. Fudenberg G, Abdennur N, Imakaev M, Goloborodko A, Mirny LA. Emerging Evidence of Chromosome Folding by Loop Extrusion. *Cold Spring Harb Symp Quant Biol.* 2017;82:45-55. doi:10.1101/sqb.2017.82.034710

158. Le Dily F, Baù D, Pohl A, et al. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* 2014;28(19):2151-2162. doi:10.1101/gad.241422.114
159. Dixon JR, Gorkin DU, Ren B. Chromatin Domains: the Unit of Chromosome Organization. *Mol Cell.* 2016;62(5):668-680. doi:10.1016/j.molcel.2016.05.018
160. Kalverda B, Röling MD, Fornerod M. Chromatin organization in relation to the nuclear periphery. *FEBS Letters.* 2008;582(14):2017-2022. doi:10.1016/j.febslet.2008.04.015
161. Cavalli G, Misteli T. Functional implications of genome topology. *Nat Struct Mol Biol.* 2013;20(3):290-299. doi:10.1038/nsmb.2474
162. Mehdipour P, Santoro F, Minucci S. Epigenetic alterations in acute myeloid leukemias. *The FEBS Journal.* 2015;282(9):1786-1800. doi:10.1111/febs.13142
163. Lara-Astiaso D, Weiner A, Lorenzo-Vivas E, et al. Immunogenetics. Chromatin state dynamics during blood formation. *Science.* 2014;345(6199):943-949. doi:10.1126/science.1256271
164. Rodrigues CP, Shvedunova M, Akhtar A. Epigenetic Regulators as the Gatekeepers of Hematopoiesis. *Trends in Genetics.* 2021;37(2):125-142. doi:10.1016/j.tig.2020.09.015
165. Martin EW, Krietsch J, Reggiardo RE, Sousae R, Kim DH, Forsberg EC. Chromatin accessibility maps provide evidence of multilineage gene priming in hematopoietic stem cells. *Epigenetics & Chromatin.* 2021;14(1):2. doi:10.1186/s13072-020-00377-1
166. Petruk S, Mariani SA, De Dominici M, et al. Structure of Nascent Chromatin Is Essential for Hematopoietic Lineage Specification. *Cell Reports.* 2017;19(2):295-306. doi:10.1016/j.celrep.2017.03.035
167. Goyama S, Kitamura T. Epigenetics in normal and malignant hematopoiesis: An overview and update 2017. *Cancer Sci.* 2017;108(4):553-562. doi:10.1111/cas.13168
168. Howard G, Eiges R, Gaudet F, Jaenisch R, Eden A. Activation and transposition of endogenous retroviral elements in hypomethylation induced tumors in mice. *Oncogene.* 2008;27(3):404-408. doi:10.1038/sj.onc.1210631
169. A. Khrabrova D, G. Loiko A, A. Tolkacheva A, et al. Functional Analysis of DNMT3A DNA Methyltransferase Mutations Reported in Patients with Acute Myeloid Leukemia. *Biomolecules.* 2019;10(1):8. doi:10.3390/biom10010008
170. Jeong M, Park HJ, Celik H, et al. Loss of Dnmt3a Immortalizes Hematopoietic Stem Cells In Vivo. *Cell Rep.* 2018;23(1):1-10. doi:10.1016/j.celrep.2018.03.025
171. Huang HT, Figueroa ME. Epigenetic deregulation in myeloid malignancies. *Blood.* 2021;138(8):613-624. doi:10.1182/blood.2019004262
172. Ilango S, Paital B, Jayachandran P, Padma PR, Nirmaladevi R. Epigenetic alterations in cancer. *Front Biosci (Landmark Ed).* 2020;25:1058-1109. doi:10.2741/4847
173. Moran-Crusio K, Reavie L, Shih A, et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell.* 2011;20(1):11-24. doi:10.1016/j.ccr.2011.06.001
174. Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell.* 2010;18(6):553-567. doi:10.1016/j.ccr.2010.11.015

175. Kroeze LI, Nikoloski G, da Silva-Coelho P, et al. Genetic defects in PRC2 components other than EZH2 are not common in myeloid malignancies. *Blood*. 2012;119(5):1318-1319. doi:10.1182/blood-2011-07-365213
176. Göllner S, Oellerich T, Agrawal-Singh S, et al. Loss of the Histone Methyltransferase EZH2 induces Resistance to Multiple Drugs in Acute Myeloid Leukemia. *Nat Med*. 2017;23(1):69-78. doi:10.1038/nm.4247
177. Birch NW, Shilatifard A. The role of histone modifications in leukemogenesis. *J Biosci*. 2020;45:6.
178. Zeisig BB, Milne T, García-Cuellar MP, et al. Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. *Mol Cell Biol*. 2004;24(2):617-628. doi:10.1128/MCB.24.2.617-628.2004
179. Wang J, Hoshino T, Redner RL, Kajigaya S, Liu JM. ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex. *Proc Natl Acad Sci U S A*. 1998;95(18):10860-10865.
180. Corces MR, Buenrostro JD, Wu B, et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat Genet*. 2016;48(10):1193-1203. doi:10.1038/ng.3646
181. Shi J, Whyte WA, Zepeda-Mendoza CJ, et al. Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. *Genes Dev*. 2013;27(24):2648-2662. doi:10.1101/gad.232710.113
182. Puente XS, Beà S, Valdés-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2015;526(7574):519-524. doi:10.1038/nature14666
183. Rasmussen KD, Jia G, Johansen JV, et al. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. *Genes Dev*. 2015;29(9):910-922. doi:10.1101/gad.260174.115
184. Clermont PL, Parolia A, Liu HH, Helgason CD. DNA methylation at enhancer regions: Novel avenues for epigenetic biomarker development. *Front Biosci (Landmark Ed)*. 2016;21:430-446. doi:10.2741/4399
185. Yang M, Safavi S, Woodward EL, et al. 13q12.2 deletions in acute lymphoblastic leukemia lead to upregulation of FLT3 through enhancer hijacking. *Blood*. 2020;136(8):946-956. doi:10.1182/blood.2019004684
186. Fisher JB, McNulty M, Burke MJ, Crispino JD, Rao S. Cohesin Mutations in Myeloid Malignancies. *Trends Cancer*. 2017;3(4):282-293. doi:10.1016/j.trecan.2017.02.006
187. Smith JS, Lappin KM, Craig SG, et al. Chronic loss of STAG2 leads to altered chromatin structure contributing to de-regulated transcription in AML. *J Transl Med*. 2020;18(1):339. doi:10.1186/s12967-020-02500-y
188. Mujahed H, Miliara S, Neddermeyer A, et al. AML displays increased CTCF occupancy associated with aberrant gene expression and transcription factor binding. *Blood*. 2020;136(3):339-352. doi:10.1182/blood.2019002326
189. Lee E, Koh Y, Hong J, Eom HS, Yoon SS. Recent Clinical Update of Acute Myeloid Leukemia: Focus on Epigenetic Therapies. *J Korean Med Sci*. 2021;36(13):e85. doi:10.3346/jkms.2021.36.e85
190. Zhao X, Liu HQ, Wang LN, Yang L, Liu XL. Current and emerging molecular and epigenetic disease entities in acute myeloid leukemia and a critical assessment of their therapeutic modalities. *Semin Cancer Biol*. Published online November 23, 2020:S1044-579X(20)30246-7. doi:10.1016/j.semcancer.2020.11.010
191. Lewis EB. A gene complex controlling segmentation in *Drosophila*. *Nature*. 1978;276(5688):565-570. doi:10.1038/276565a0
192. Hajirnis N, Mishra RK. Homeotic Genes: Clustering, Modularity, and Diversity. *Front Cell Dev Biol*. 2021;9:718308. doi:10.3389/fcell.2021.718308

193. Garcia-Fernández J. Hox, ParaHox, ProtoHox: facts and guesses. *Heredity (Edinb)*. 2005;94(2):145-152. doi:10.1038/sj.hdy.6800621
194. Gehring WJ, Kloter U, Suga H. Evolution of the Hox gene complex from an evolutionary ground state. *Curr Top Dev Biol*. 2009;88:35-61. doi:10.1016/S0070-2153(09)88002-2
195. Wu LH, Lengyel JA. Role of caudal in hindgut specification and gastrulation suggests homology between *Drosophila* amnioproctodeal invagination and vertebrate blastopore. *Development*. 1998;125(13):2433-2442.
196. Gross I, Lhermitte B, Domon-Dell C, et al. Phosphorylation of the homeotic tumor suppressor Cdx2 mediates its ubiquitin-dependent proteasome degradation. *Oncogene*. 2005;24(54):7955-7963. doi:10.1038/sj.onc.1208945
197. Witek ME, Snook AE, Lin JE, et al. A Novel CDX2 Isoform Regulates Alternative Splicing. *PLoS One*. 2014;9(8):e104293. doi:10.1371/journal.pone.0104293
198. Balbinot C, Vanier M, Armant O, et al. Fine-tuning and autoregulation of the intestinal determinant and tumor suppressor homeobox gene CDX2 by alternative splicing. *Cell Death Differ*. 2017;24(12):2173-2186. doi:10.1038/cdd.2017.140
199. Marlétaz F, Maeso I, Faas L, Isaacs HV, Holland PWH. Cdx ParaHox genes acquired distinct developmental roles after gene duplication in vertebrate evolution. *BMC Biology*. 2015;13(1):56. doi:10.1186/s12915-015-0165-x
200. Subramanian V, Meyer BI, Gruss P. Disruption of the murine homeobox gene Cdx1 affects axial skeletal identities by altering the mesodermal expression domains of Hox genes. *Cell*. 1995;83(4):641-653. doi:10.1016/0092-8674(95)90104-3
201. Deschamps J, van Nes J. Developmental regulation of the Hox genes during axial morphogenesis in the mouse. *Development*. 2005;132(13):2931-2942. doi:10.1242/dev.01897
202. Neijts R, Amin S, van Rooijen C, Deschamps J. Cdx is crucial for the timing mechanism driving colinear Hox activation and defines a trunk segment in the Hox cluster topology. *Developmental Biology*. 2017;422(2):146-154. doi:10.1016/j.ydbio.2016.12.024
203. Yagi R, Kohn MJ, Karavanova I, et al. Transcription factor TEAD4 specifies the trophectoderm lineage at the beginning of mammalian development. *Development*. 2007;134(21):3827-3836. doi:10.1242/dev.010223
204. Rayon T, Menchero S, Nieto A, et al. Notch and Hippo Converge on Cdx2 to Specify the Trophectoderm Lineage in the Mouse Blastocyst. *Developmental Cell*. 2014;30(4):410-422. doi:10.1016/j.devcel.2014.06.019
205. Watanabe Y, Miyasaka KY, Kubo A, et al. Notch and Hippo signaling converge on Strawberry Notch 1 (Sbno1) to synergistically activate Cdx2 during specification of the trophectoderm. *Sci Rep*. 2017;7:46135. doi:10.1038/srep46135
206. Nishioka N, Inoue K, Adachi K, et al. The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. *Dev Cell*. 2009;16(3):398-410. doi:10.1016/j.devcel.2009.02.003
207. Home P, Ray S, Dutta D, Bronshteyn I, Larson M, Paul S. GATA3 Is Selectively Expressed in the Trophectoderm of Peri-implantation Embryo and Directly Regulates Cdx2 Gene Expression. *J Biol Chem*. 2009;284(42):28729-28737. doi:10.1074/jbc.M109.016840
208. Sebastiano V, Dalvai M, Gentile L, et al. Oct1 regulates trophoblast development during early mouse embryogenesis. *Development*. 2010;137(21):3551-3560. doi:10.1242/dev.047027

209. Tolkunova E, Cavaleri F, Eckardt S, et al. The caudal-related protein cdx2 promotes trophoblast differentiation of mouse embryonic stem cells. *Stem Cells*. 2006;24(1):139-144. doi:10.1634/stemcells.2005-0240
210. Strumpf D, Mao CA, Yamanaka Y, et al. Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development*. 2005;132(9):2093-2102. doi:10.1242/dev.01801
211. Ralston A, Rossant J. Cdx2 acts downstream of cell polarization to cell-autonomously promote trophectoderm fate in the early mouse embryo. *Developmental Biology*. 2008;313(2):614-629. doi:10.1016/j.ydbio.2007.10.054
212. Bou G, Liu S, Guo J, et al. Cdx2 represses Oct4 function via inducing its proteasome-dependent degradation in early porcine embryos. *Developmental Biology*. 2016;410(1):36-44. doi:10.1016/j.ydbio.2015.12.014
213. Huang D, Guo G, Yuan P, et al. The role of Cdx2 as a lineage specific transcriptional repressor for pluripotent network during the first developmental cell lineage segregation. *Scientific Reports*. 2017;7(1). doi:10.1038/s41598-017-16009-w
214. Beck F, Erler T, Russell A, James R. Expression of Cdx-2 in the mouse embryo and placenta: Possible role in patterning of the extra-embryonic membranes. *Developmental Dynamics*. 1995;204(3):219-227. doi:10.1002/aja.1002040302
215. Niwa H, Toyooka Y, Shimosato D, et al. Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation. *Cell*. 2005;123(5):917-929. doi:10.1016/j.cell.2005.08.040
216. van den Akker E, Forlani S, Chawengsaksophak K, et al. Cdx1 and Cdx2 have overlapping functions in anteroposterior patterning and posterior axis elongation. *Development*. 2002;129(9):2181-2193.
217. van Nes J, de Graaff W, Lebrin F, Gerhard M, Beck F, Deschamps J. The Cdx4 mutation affects axial development and reveals an essential role of Cdx genes in the ontogenesis of the placental labyrinth in mice. *Development*. 2006;133(3):419-428. doi:10.1242/dev.02216
218. Young T, Deschamps J. Hox, Cdx, and anteroposterior patterning in the mouse embryo. *Curr Top Dev Biol*. 2009;88:235-255. doi:10.1016/S0070-2153(09)88008-3
219. Beck F, Stringer EJ. The role of Cdx genes in the gut and in axial development. *Biochem Soc Trans*. 2010;38(2):353-357. doi:10.1042/BST0380353
220. Amin S, Neijts R, Simmini S, et al. Cdx and T Brachyury Co-activate Growth Signaling in the Embryonic Axial Progenitor Niche. *Cell Reports*. 2016;17(12):3165-3177. doi:10.1016/j.celrep.2016.11.069
221. Chawengsaksophak K, James R, Hammond VE, Köntgen F, Beck F. Homeosis and intestinal tumours in Cdx2 mutant mice. *Nature*. 1997;386(6620):84-87. doi:10.1038/386084a0
222. Chawengsaksophak K, de Graaff W, Rossant J, Deschamps J, Beck F. Cdx2 is essential for axial elongation in mouse development. *Proc Natl Acad Sci U S A*. 2004;101(20):7641-7645. doi:10.1073/pnas.0401654101
223. Savory JGA, Bouchard N, Pierre V, et al. Cdx2 regulation of posterior development through non-Hox targets. *Development*. 2009;136(24):4099-4110. doi:10.1242/dev.041582
224. Zhao T, Gan Q, Stokes A, et al. β -catenin regulates Pax3 and Cdx2 for caudal neural tube closure and elongation. *Development*. 2014;141(1):148-157. doi:10.1242/dev.101550
225. Creamer JP, Dege C, Ren Q, et al. Human definitive hematopoietic specification from pluripotent stem cells is regulated by mesodermal expression of CDX4. *Blood*. 2017;129(22):2988-2992. doi:10.1182/blood-2016-11-749382

226. Lengerke C, Schmitt S, Bowman TV, et al. BMP and Wnt specify hematopoietic fate by activation of the Cdx-Hox pathway. *Cell Stem Cell*. 2008;2(1):72-82. doi:10.1016/j.stem.2007.10.022
227. Wang Y, Yabuuchi A, McKinney-Freeman S, et al. Cdx gene deficiency compromises embryonic hematopoiesis in the mouse. *Proc Natl Acad Sci USA*. 2008;105(22):7756-7761. doi:10.1073/pnas.0708951105
228. Jung HS, Uenishi G, Park MA, et al. SOX17 integrates HOXA and arterial programs in hemogenic endothelium to drive definitive lympho-myeloid hematopoiesis. *Cell Rep*. 2021;34(7):108758. doi:10.1016/j.celrep.2021.108758
229. Braekeleer ED, Douet-Guilbert N, Basinko A, Bris MJL, Morel F, Braekeleer MD. Hox gene dysregulation in acute myeloid leukemia. *Future Oncology*. 2014;10(3):475-495. doi:10.2217/fon.13.195
230. Alharbi RA, Pettengell R, Pandha HS, Morgan R. The role of HOX genes in normal hematopoiesis and acute leukemia. *Leukemia*. 2013;27(5):1000-1008. doi:10.1038/leu.2012.356
231. Brooke-Bisschop T, Savory JGA, Foley T, Ringuette R, Lohnes D. Essential roles for Cdx in murine primitive hematopoiesis. *Developmental Biology*. 2017;422(2):115-124. doi:10.1016/j.ydbio.2017.01.002
232. El Omar R, Julien E, Biasch K, et al. CDX2 regulates ACE expression in blood development and leukemia cells. *Blood Adv*. 2021;5(7):2012-2016. doi:10.1182/bloodadvances.2020003563
233. Sherwood RI, Maehr R, Mazzoni EO, Melton DA. Wnt Signaling Specifies and Patterns Intestinal Endoderm. *Mech Dev*. 2011;128(7-10):387-400. doi:10.1016/j.mod.2011.07.005
234. Benahmed F, Gross I, Gaunt SJ, et al. Multiple regulatory regions control the complex expression pattern of the mouse Cdx2 homeobox gene. *Gastroenterology*. 2008;135(4):1238-1247, 1247.e1-3. doi:10.1053/j.gastro.2008.06.045
235. Gao N, White P, Kaestner KH. Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2. *Dev Cell*. 2009;16(4):588-599. doi:10.1016/j.devcel.2009.02.010
236. Stringer EJ, Duluc I, Saandi T, et al. Cdx2 determines the fate of postnatal intestinal endoderm. *Development*. 2012;139(3):465-474. doi:10.1242/dev.070722
237. Grainger S, Savory JGA, Lohnes D. Cdx2 regulates patterning of the intestinal epithelium. *Dev Biol*. 2010;339(1):155-165. doi:10.1016/j.ydbio.2009.12.025
238. Banerjee KK, Saxena M, Kumar N, et al. Enhancer, transcriptional, and cell fate plasticity precedes intestinal determination during endoderm development. *Genes Dev*. 2018;32(21-22):1430-1442. doi:10.1101/gad.318832.118
239. James R, Erler T, Kazenwadel J. Structure of the murine homeobox gene cdx-2. Expression in embryonic and adult intestinal epithelium. *J Biol Chem*. 1994;269(21):15229-15237.
240. Silberg DG, Swain GP, Suh ER, Traber PG. Cdx1 and cdx2 expression during intestinal development. *Gastroenterology*. 2000;119(4):961-971. doi:10.1053/gast.2000.18142
241. Barros R, da Costa LT, Pinto-de-Sousa J, et al. CDX2 autoregulation in human intestinal metaplasia of the stomach: impact on the stability of the phenotype. *Gut*. 2011;60(3):290-298. doi:10.1136/gut.2010.222323
242. Taylor JK, Levy T, Suh ER, Traber PG. Activation of enhancer elements by the homeobox gene Cdx2 is cell line specific. *Nucleic Acids Res*. 1997;25(12):2293-2300. doi:10.1093/nar/25.12.2293
243. Verzi MP, Shin H, Ho LL, Liu XS, Shivdasani RA. Essential and redundant functions of caudal family proteins in activating adult intestinal genes. *Mol Cell Biol*. 2011;31(10):2026-2039. doi:10.1128/MCB.01250-10

244. Gu W, Wang H, Huang X, et al. SATB2 preserves colon stem cell identity and mediates ileum-colon conversion via enhancer remodeling. *Cell Stem Cell*. Published online September 22, 2021:S1934-5909(21)00380-5. doi:10.1016/j.stem.2021.09.004
245. Verzi MP, Shin H, He HH, et al. Differentiation-specific histone modifications reveal dynamic chromatin interactions and alternative partners for the intestinal transcription factor CDX2. *Dev Cell*. 2010;19(5):713-726. doi:10.1016/j.devcel.2010.10.006
246. Verzi MP, Shin H, San Roman AK, Liu XS, Shivdasani RA. Intestinal Master Transcription Factor CDX2 Controls Chromatin Access for Partner Transcription Factor Binding. *Mol Cell Biol*. 2013;33(2):281-292. doi:10.1128/MCB.01185-12
247. Kumar A, D'Souza SS, Thakur AS. Understanding the Journey of Human Hematopoietic Stem Cell Development. *Stem Cells International*. 2019;2019:1-13. doi:10.1155/2019/2141475
248. Nguyen TT, Savory JGA, Brooke-Bisschop T, et al. Cdx2 Regulates Gene Expression through Recruitment of Brg1-associated Switch-Sucrose Non-fermentable (SWI-SNF) Chromatin Remodeling Activity. *J Biol Chem*. 2017;292(8):3389-3399. doi:10.1074/jbc.M116.752774
249. Katz JP, Perreault N, Goldstein BG, et al. The zinc-finger transcription factor Klf4 is required for terminal differentiation of goblet cells in the colon. *Development*. 2002;129(11):2619-2628.
250. Mutoh H, Sakamoto H, Hayakawa H, et al. The intestine-specific homeobox gene Cdx2 induces expression of the basic helix-loop-helix transcription factor Math1. *Differentiation*. 2006;74(6):313-321. doi:10.1111/j.1432-0436.2006.00074.x
251. Grainger S, Lam J, Savory JGA, Mears AJ, Rijli FM, Lohnes D. Cdx regulates Dll1 in multiple lineages. *Dev Biol*. 2012;361(1):1-11. doi:10.1016/j.ydbio.2011.09.034
252. San Roman AK, Tovaglieri A, Breault DT, Shivdasani RA. Distinct Processes and Transcriptional Targets Underlie CDX2 Requirements in Intestinal Stem Cells and Differentiated Villus Cells. *Stem Cell Reports*. 2015;5(5):673-681. doi:10.1016/j.stemcr.2015.09.006
253. Suh E, Chen L, Taylor J, Traber PG. A homeodomain protein related to caudal regulates intestine-specific gene transcription. *Mol Cell Biol*. 1994;14(11):7340-7351.
254. Troelsen JT, Mitchelmore C, Spodsberg N, Jensen AM, Norén O, Sjöström H. Regulation of lactase-phlorizin hydrolase gene expression by the caudal-related homeodomain protein Cdx-2. *Biochem J*. 1997;322(Pt 3):833-838.
255. Uesaka T, Kageyama N, Watanabe H. Identifying target genes regulated downstream of Cdx2 by microarray analysis. *J Mol Biol*. 2004;337(3):647-660. doi:10.1016/j.jmb.2004.01.061
256. Ma L, Jüttner M, Kullak-Ublick GA, Eloranta JJ. Regulation of the gene encoding the intestinal bile acid transporter ASBT by the caudal-type homeobox proteins CDX1 and CDX2. *American Journal of Physiology-Gastrointestinal and Liver Physiology*. 2012;302(1):G123-G133. doi:10.1152/ajpgi.00102.2011
257. Yamamoto H, Bai YQ, Yuasa Y. Homeodomain protein CDX2 regulates goblet-specific MUC2 gene expression. *Biochem Biophys Res Commun*. 2003;300(4):813-818. doi:10.1016/s0006-291x(02)02935-2
258. Jonckheere N, Vincent A, Perrais M, et al. The human mucin MUC4 is transcriptionally regulated by caudal-related homeobox, hepatocyte nuclear factors, forkhead box A, and GATA endodermal transcription factors in epithelial cancer cells. *J Biol Chem*. 2007;282(31):22638-22650. doi:10.1074/jbc.M700905200
259. Modica S, Cariello M, Morgano A, et al. Transcriptional regulation of the intestinal nuclear bile acid farnesoid X receptor (FXR) by the caudal-related homeobox 2 (CDX2). *J Biol Chem*. 2014;289(41):28421-28432. doi:10.1074/jbc.M114.571513

260. Modica S, Morgano A, Salvatore L, et al. Expression and localisation of insulin receptor substrate 2 in normal intestine and colorectal tumours. Regulation by intestine-specific transcription factor CDX2. *Gut*. 2009;58(9):1250-1259. doi:10.1136/gut.2008.158386
261. Hinoi T, Lucas PC, Kuick R, Hanash S, Cho KR, Fearon ER. CDX2 regulates liver intestine-cadherin expression in normal and malignant colon epithelium and intestinal metaplasia. *Gastroenterology*. 2002;123(5):1565-1577. doi:10.1053/gast.2002.36598
262. Keller MS, Ezaki T, Guo RJ, Lynch JP. Cdx1 or Cdx2 expression activates E-cadherin-mediated cell-cell adhesion and compaction in human COLO 205 cells. *Am J Physiol Gastrointest Liver Physiol*. 2004;287(1):G104-114. doi:10.1152/ajpgi.00484.2003
263. Sakaguchi T, Gu X, Golden HM, Suh E, Rhoads DB, Reinecker HC. Cloning of the human claudin-2 5'-flanking region revealed a TATA-less promoter with conserved binding sites in mouse and human for caudal-related homeodomain proteins and hepatocyte nuclear factor-1alpha. *J Biol Chem*. 2002;277(24):21361-21370. doi:10.1074/jbc.M110261200
264. Verzi MP, Hatzis P, Sulahian R, et al. TCF4 and CDX2, major transcription factors for intestinal function, converge on the same cis-regulatory regions. *PNAS*. 2010;107(34):15157-15162. doi:10.1073/pnas.1003822107
265. Boyd M, Hansen M, Jensen TGK, et al. Genome-wide analysis of CDX2 binding in intestinal epithelial cells (Caco-2). *J Biol Chem*. 2010;285(33):25115-25125. doi:10.1074/jbc.M109.089516
266. Hinkel I, Duluc I, Martin E, Guenot D, Freund JN, Gross I. Cdx2 controls expression of the protocadherin Mucdhl, an inhibitor of growth and β -catenin activity in colon cancer cells. *Gastroenterology*. 2012;142(4):875-885.e3. doi:10.1053/j.gastro.2011.12.037
267. Olsen AK, Coskun M, Bzorek M, et al. Regulation of APC and AXIN2 expression by intestinal tumor suppressor CDX2 in colon cancer cells. *Carcinogenesis*. 2013;34(6):1361-1369. doi:10.1093/carcin/bgt037
268. Bai YQ, Miyake S, Iwai T, Yuasa Y. CDX2, a homeobox transcription factor, upregulates transcription of the p21/WAF1/CIP1 gene. *Oncogene*. 2003;22(39):7942-7949. doi:10.1038/sj.onc.1206634
269. Zhu Y, Hryniuk A, Foley T, Hess B, Lohnes D. Cdx2 Regulates Intestinal EphrinB1 through the Notch Pathway. *Genes (Basel)*. 2021;12(2):188. doi:10.3390/genes12020188
270. Lorentz O, Duluc I, Arcangelis AD, Simon-Assmann P, Kedinger M, Freund JN. Key role of the Cdx2 homeobox gene in extracellular matrix-mediated intestinal cell differentiation. *J Cell Biol*. 1997;139(6):1553-1565. doi:10.1083/jcb.139.6.1553
271. Chewchuk S, Jahan S, Lohnes D. Cdx2 regulates immune cell infiltration in the intestine. *Sci Rep*. 2021;11:15841. doi:10.1038/s41598-021-95412-w
272. Renouf B, Soret C, Saandi T, et al. Cdx2 homeoprotein inhibits non-homologous end joining in colon cancer but not in leukemia cells. *Nucleic Acids Research*. 2012;40(8):3456-3469. doi:10.1093/nar/gkr1242
273. Soret C, Martin E, Duluc I, et al. Distinct mechanisms for opposite functions of homeoproteins Cdx2 and HoxB7 in double-strand break DNA repair in colon cancer cells. *Cancer Lett*. 2016;374(2):208-215. doi:10.1016/j.canlet.2016.02.026
274. Aoki K, Kakizaki F, Sakashita H, Manabe T, Aoki M, Taketo MM. Suppression of colonic polyposis by homeoprotein CDX2 through its nontranscriptional function that stabilizes p27Kip1. *Cancer Res*. 2011;71(2):593-602. doi:10.1158/0008-5472.CAN-10-2842

275. Guo RJ, Funakoshi S, Lee HH, Kong J, Lynch JP. The intestine-specific transcription factor Cdx2 inhibits β -catenin/TCF transcriptional activity by disrupting the β -catenin-TCF protein complex. *Carcinogenesis*. 2010;31(2):159-166. doi:10.1093/carcin/bgp213
276. Beck F, Chawengsaksohak K, Waring P, Playford RJ, Furness JB. Reprogramming of intestinal differentiation and intercalary regeneration in Cdx2 mutant mice. *Proc Natl Acad Sci U S A*. 1999;96(13):7318-7323. doi:10.1073/pnas.96.13.7318
277. Marisa L, de Reyniès A, Duval A, et al. Gene Expression Classification of Colon Cancer into Molecular Subtypes: Characterization, Validation, and Prognostic Value. *PLoS Med*. 2013;10(5):e1001453. doi:10.1371/journal.pmed.1001453
278. Bakaris S, Cetinkaya A, Ezberci F, Ekerbicer H. Expression of homeodomain protein CDX2 in colorectal adenoma and adenocarcinoma. *Histol Histopathol*. 2008;23(9):1043-1047. doi:10.14670/HH-23.1043
279. Olsen J, Eiholm S, Kirkeby LT, et al. CDX2 downregulation is associated with poor differentiation and MMR deficiency in colon cancer. *Exp Mol Pathol*. 2016;100(1):59-66. doi:10.1016/j.yexmp.2015.11.009
280. Dalerba P, Sahoo D, Paik S, et al. CDX2 as a Prognostic Biomarker in Stage II and Stage III Colon Cancer. *N Engl J Med*. 2016;374(3):211-222. doi:10.1056/NEJMoa1506597
281. Tóth C, Sükösd F, Valicsek E, Herpel E, Schirmacher P, Tiszlavicz L. Loss of CDX2 gene expression is associated with DNA repair proteins and is a crucial member of the Wnt signaling pathway in liver metastasis of colorectal cancer. *Oncol Lett*. 2018;15(3):3586-3593. doi:10.3892/ol.2018.7756
282. Neumann J, Heinemann V, Engel J, Kirchner T, Stintzing S. The prognostic impact of CDX2 correlates with the underlying mismatch repair status and BRAF mutational status but not with distant metastasis in colorectal cancer. *Virchows Arch*. 2018;473(2):199-207. doi:10.1007/s00428-018-2360-y
283. Bruun J, Sveen A, Barros R, et al. Prognostic, predictive, and pharmacogenomic assessments of CDX2 refine stratification of colorectal cancer. *Mol Oncol*. 2018;12(9):1639-1655. doi:10.1002/1878-0261.12347
284. Dawson H, Galván JA, Helbling M, et al. Possible role of Cdx2 in the serrated pathway of colorectal cancer characterized by BRAF mutation, high-level CpG Island methylator phenotype and mismatch repair-deficiency. *Int J Cancer*. 2014;134(10):2342-2351. doi:10.1002/ijc.28564
285. Camilo V, Garrido M, Valente P, et al. Differentiation reprogramming in gastric intestinal metaplasia and dysplasia: role of SOX2 and CDX2. *Histopathology*. 2015;66(3):343-350. doi:10.1111/his.12544
286. Masood MA, Loya A, Yusuf MA. CDX2 as a prognostic marker in gastric cancer. *Acta Gastroenterol Belg*. 2016;79(2):197-200.
287. Seno H, Oshima M, Taniguchi MA, et al. CDX2 expression in the stomach with intestinal metaplasia and intestinal-type cancer: Prognostic implications. *Int J Oncol*. 2002;21(4):769-774. doi:10.3892/ijo.21.4.769
288. Liu Q, Teh M, Ito K, Shah N, Ito Y, Yeoh KG. CDX2 expression is progressively decreased in human gastric intestinal metaplasia, dysplasia and cancer. *Mod Pathol*. 2007;20(12):1286-1297. doi:10.1038/modpathol.3800968
289. Saito M, Okayama H, Saito K, et al. CDX2 is involved in microRNA-associated inflammatory carcinogenesis in gastric cancer. *Oncol Lett*. 2017;14(5):6184-6190. doi:10.3892/ol.2017.6956
290. Lopes N, Bergsland C, Bruun J, et al. A panel of intestinal differentiation markers (CDX2, GPA33, and LI-cadherin) identifies gastric cancer patients with favourable prognosis. *Gastric Cancer*. 2020;23(5):811-823. doi:10.1007/s10120-020-01064-6

291. Scholl C, Bansal D, Döhner K, et al. The homeobox gene CDX2 is aberrantly expressed in most cases of acute myeloid leukemia and promotes leukemogenesis. *Journal of Clinical Investigation*. 2007;117(4):1037-1048. doi:10.1172/JCI30182
292. Riedt T, Ebinger M, Salih HR, et al. Aberrant expression of the homeobox gene CDX2 in pediatric acute lymphoblastic leukemia. *Blood*. 2009;113(17):4049-4051. doi:10.1182/blood-2008-12-196634
293. Thoene S, Rawat VPS, Heilmeier B, et al. The homeobox gene CDX2 is aberrantly expressed and associated with an inferior prognosis in patients with acute lymphoblastic leukemia. *Leukemia*. 2009;23(4):649-655. doi:10.1038/leu.2008.355
294. Darvishi M, Mashati P, Khosravi A. The clinical significance of CDX2 in leukemia: A new perspective for leukemia research. *Leuk Res*. 2018;72:45-51. doi:10.1016/j.leukres.2018.07.021
295. Monk M, Holding C. Human embryonic genes re-expressed in cancer cells. *Oncogene*. 2001;20(56):8085-8091. doi:10.1038/sj.onc.1205088
296. Yu X, Xu J. A 'Goldmine' for digging cancer-specific targets: the genes essential for embryo development but non-essential for adult life. *J Mol Cell Biol*. 2020;12(9):669-673. doi:10.1093/jmcb/mjaa024
297. Bonhomme C, Duluc I, Martin E, et al. The Cdx2 homeobox gene has a tumour suppressor function in the distal colon in addition to a homeotic role during gut development. *Gut*. 2003;52(10):1465-1471.
298. Aoki K, Tamai Y, Horiike S, Oshima M, Taketo MM. Colonic polyposis caused by mTOR-mediated chromosomal instability in Apc+/ Δ 716 Cdx2+/- compound mutant mice. *Nat Genet*. 2003;35(4):323-330. doi:10.1038/ng1265
299. Sakamoto N, Feng Y, Stolfi C, et al. BRAFV600E cooperates with CDX2 inactivation to promote serrated colorectal tumorigenesis. *Elife*. 2017;6:e20331. doi:10.7554/eLife.20331
300. Balbinot C, Armant O, Elarouci N, et al. The Cdx2 homeobox gene suppresses intestinal tumorigenesis through non-cell-autonomous mechanisms. *J Exp Med*. 2018;215(3):911-926. doi:10.1084/jem.20170934
301. Yu J, Liu D, Sun X, et al. CDX2 inhibits the proliferation and tumor formation of colon cancer cells by suppressing Wnt/ β -catenin signaling via transactivation of GSK-3 β and Axin2 expression. *Cell Death Dis*. 2019;10(1):26. doi:10.1038/s41419-018-1263-9
302. Gross I, Duluc I, Benameur T, et al. The intestine-specific homeobox gene Cdx2 decreases mobility and antagonizes dissemination of colon cancer cells. *Oncogene*. 2008;27(1):107-115. doi:10.1038/sj.onc.1210601
303. Yu J, Li S, Xu Z, et al. CDX2 inhibits epithelial-mesenchymal transition in colorectal cancer by modulation of Snail expression and β -catenin stabilisation via transactivation of PTEN expression. *Br J Cancer*. 2021;124(1):270-280. doi:10.1038/s41416-020-01148-1
304. Salari K, Spulak ME, Cuff J, et al. CDX2 is an amplified lineage-survival oncogene in colorectal cancer. *Proc Natl Acad Sci U S A*. 2012;109(46):E3196-E3205. doi:10.1073/pnas.1206004109
305. Silberg DG, Sullivan J, Kang E, et al. Cdx2 ectopic expression induces gastric intestinal metaplasia in transgenic mice. *Gastroenterology*. 2002;122(3):689-696.
306. Yamamichi N, Inada K, Furukawa C, et al. Cdx2 and the Brm-type SWI/SNF complex cooperatively regulate villin expression in gastrointestinal cells. *Exp Cell Res*. 2009;315(10):1779-1789. doi:10.1016/j.yexcr.2009.01.006
307. Mutoh H, Hakamata Y, Sato K, et al. Conversion of gastric mucosa to intestinal metaplasia in Cdx2-expressing transgenic mice. *Biochem Biophys Res Commun*. 2002;294(2):470-479. doi:10.1016/S0006-291X(02)00480-1

308. Mutoh H, Sakurai S, Satoh K, et al. Development of Gastric Carcinoma from Intestinal Metaplasia in Cdx2-transgenic Mice. *Cancer Res.* 2004;64(21):7740-7747. doi:10.1158/0008-5472.CAN-04-1617
309. Mutoh H, Hayakawa H, Sashikawa M, Sakamoto H, Sugano K. Direct repression of Sonic Hedgehog expression in the stomach by Cdx2 leads to intestinal transformation. *Biochem J.* 2010;427(3):423-434. doi:10.1042/BJ20091177
310. Chai D, Du H, Li K, et al. CDX2 and Reg IV expression and correlation in gastric cancer. *BMC Gastroenterol.* 2021;21:92. doi:10.1186/s12876-021-01678-9
311. Wang XT, Xie YB, Xiao Q. siRNA targeting of Cdx2 inhibits growth of human gastric cancer MGC-803 cells. *World J Gastroenterol.* 2012;18(16):1903-1914. doi:10.3748/wjg.v18.i16.1903
312. Zhang JF, Qu LS, Qian XF, Xia BL, Mao ZB, Chen WC. Nuclear transcription factor CDX2 inhibits gastric cancer-cell growth and reverses epithelial-to-mesenchymal transition in vitro and in vivo. *Mol Med Rep.* 2015;12(4):5231-5238. doi:10.3892/mmr.2015.4114
313. Rawat VPS, Humphries RK, Buske C. Beyond Hox: the role of ParaHox genes in normal and malignant hematopoiesis. *Blood.* 2012;120(3):519-527. doi:10.1182/blood-2012-02-385898
314. Rawat VPS, Thoene S, Naidu VM, et al. Overexpression of CDX2 perturbs HOX gene expression in murine progenitors depending on its N-terminal domain and is closely correlated with deregulated HOX gene expression in human acute myeloid leukemia. *Blood.* 2008;111(1):309-319. doi:10.1182/blood-2007-04-085407
315. Argiropoulos B, Humphries RK. Hox genes in hematopoiesis and leukemogenesis. *Oncogene.* 2007;26(47):6766-6776. doi:10.1038/sj.onc.1210760
316. Fröhling S, Scholl C, Bansal D, Huntly BJP. HOX gene regulation in acute myeloid leukemia: CDX marks the spot? *Cell Cycle.* 2007;6(18):2241-2245. doi:10.4161/cc.6.18.4656
317. Faber K, Bullinger L, Ragu C, et al. CDX2-driven leukemogenesis involves KLF4 repression and deregulated PPAR γ signaling. *Journal of Clinical Investigation.* 2013;123(1):299-314. doi:10.1172/JCI64745
318. Tagawa T, Haraguchi T, Hiramatsu H, et al. Multiple microRNAs induced by Cdx1 suppress Cdx2 in human colorectal tumour cells. *Biochem J.* 2012;447(3):449-455. doi:10.1042/BJ20120434
319. Buchet-Poyau K, Courchet J, Le Hir H, et al. Identification and characterization of human Mex-3 proteins, a novel family of evolutionarily conserved RNA-binding proteins differentially localized to processing bodies. *Nucleic Acids Res.* 2007;35(4):1289-1300. doi:10.1093/nar/gkm016
320. Pereira B, Sousa S, Barros R, et al. CDX2 regulation by the RNA-binding protein MEX3A: impact on intestinal differentiation and stemness. *Nucleic Acids Res.* 2013;41(7):3986-3999. doi:10.1093/nar/gkt087
321. Rings EH, Boudreau F, Taylor JK, Moffett J, Suh ER, Traber PG. Phosphorylation of the serine 60 residue within the Cdx2 activation domain mediates its transactivation capacity. *Gastroenterology.* 2001;121(6):1437-1450. doi:10.1053/gast.2001.29618
322. Houde M, Laprise P, Jean D, Blais M, Asselin C, Rivard N. Intestinal epithelial cell differentiation involves activation of p38 mitogen-activated protein kinase that regulates the homeobox transcription factor CDX2. *J Biol Chem.* 2001;276(24):21885-21894. doi:10.1074/jbc.M100236200
323. Krueger F, Madeja Z, Hemberger M, McMahon M, Cook SJ, Gaunt SJ. Down-regulation of Cdx2 in colorectal carcinoma cells by the Raf-MEK-ERK 1/2 pathway. *Cell Signal.* 2009;21(12):1846-1856. doi:10.1016/j.cellsig.2009.07.020

324. Boulanger J, Vézina A, Mongrain S, et al. Cdk2-dependent phosphorylation of homeobox transcription factor CDX2 regulates its nuclear translocation and proteasome-mediated degradation in human intestinal epithelial cells. *J Biol Chem*. 2005;280(18):18095-18107. doi:10.1074/jbc.M502184200
325. Wang WCH, Shashikant CS. Evidence for positive and negative regulation of the mouse Cdx2 gene. *J Exp Zool B Mol Dev Evol*. 2007;308(3):308-321. doi:10.1002/jez.b.21154
326. Rayon T, Menchero S, Rollán I, et al. Distinct mechanisms regulate Cdx2 expression in the blastocyst and in trophoblast stem cells. *Sci Rep*. 2016;6:27139. doi:10.1038/srep27139
327. Yeap LS, Hayashi K, Surani MA. ERG-associated protein with SET domain (ESET)-Oct4 interaction regulates pluripotency and represses the trophectoderm lineage. *Epigenetics Chromatin*. 2009;2:12. doi:10.1186/1756-8935-2-12
328. Lohmann F, Loureiro J, Su H, et al. KMT1E Mediated H3K9 Methylation Is Required for the Maintenance of Embryonic Stem Cells by Repressing Trophectoderm Differentiation. *STEM CELLS*. 2010;28(2):201-212. doi:10.1002/stem.278
329. Subtil C, Guérin E, Schneider A, et al. Frequent rearrangements and amplification of the CDX2 homeobox gene in human sporadic colorectal cancers with chromosomal instability. *Cancer Lett*. 2007;247(2):197-203. doi:10.1016/j.canlet.2006.04.004
330. Rozek LS, Lipkin SM, Fearon ER, et al. CDX2 polymorphisms, RNA expression, and risk of colorectal cancer. *Cancer Res*. 2005;65(13):5488-5492. doi:10.1158/0008-5472.CAN-04-3645
331. Hinoi T, Loda M, Fearon ER. Silencing of CDX2 expression in colon cancer via a dominant repression pathway. *J Biol Chem*. 2003;278(45):44608-44616. doi:10.1074/jbc.M307435200
332. Kawai H, Tomii K, Toyooka S, et al. Promoter methylation downregulates CDX2 expression in colorectal carcinomas. *Oncol Rep*. 2005;13(3):547-551.
333. Wang Y, Li Z, Li W, Liu S, Han B. Methylation of promoter region of CDX2 gene in colorectal cancer. *Oncol Lett*. 2016;12(5):3229-3233. doi:10.3892/ol.2016.5109
334. Graule J, Uth K, Fischer E, et al. CDX2 in colorectal cancer is an independent prognostic factor and regulated by promoter methylation and histone deacetylation in tumors of the serrated pathway. *Clin Epigenetics*. 2018;10:120. doi:10.1186/s13148-018-0548-2
335. Shakya A, Kang J, Chumley J, Williams MA, Tantin D. Oct1 Is a Switchable, Bipotential Stabilizer of Repressed and Inducible Transcriptional States. *J Biol Chem*. 2011;286(1):450-459. doi:10.1074/jbc.M110.174045
336. Enomoto H, Watanabe H, Nishikura K, Umezawa H, Asakura H. Topographic distribution of *Helicobacter pylori* in the resected stomach. *European Journal of Gastroenterology & Hepatology*. 1998;10(6):473-478.
337. Chen HN, Wang Z, Li X, Zhou ZG. *Helicobacter pylori* eradication cannot reduce the risk of gastric cancer in patients with intestinal metaplasia and dysplasia: evidence from a meta-analysis. *Gastric Cancer*. 2016;19(1):166-175. doi:10.1007/s10120-015-0462-7
338. Shiotani A, Uedo N, Iishi H, et al. Re-expression of sonic hedgehog and reduction of CDX2 after *Helicobacter pylori* eradication prior to incomplete intestinal metaplasia. *Int J Cancer*. 2007;121(6):1182-1189. doi:10.1002/ijc.22835
339. Vauhkonen M, Vauhkonen H, Sipponen P. *Helicobacter pylori* infection induces a reversible expression of the CDX2 transcription factor protein in human gastric epithelium. *Scandinavian Journal of Gastroenterology*. 2008;43(8):915-921. doi:10.1080/00365520802014841

340. Camilo V, Barros R, Sousa S, et al. Helicobacter pylori and the BMP pathway regulate CDX2 and SOX2 expression in gastric cells. *Carcinogenesis*. 2012;33(10):1985-1992. doi:10.1093/carcin/bgs233
341. Bleuming SA, Kodach LL, Garcia Leon MJ, et al. Altered bone morphogenetic protein signalling in the Helicobacter pylori-infected stomach. *J Pathol*. 2006;209(2):190-197. doi:10.1002/path.1976
342. Chen HY, Hu Y, Xu XB, et al. Upregulation of oncogene Activin A receptor type I by Helicobacter pylori infection promotes gastric intestinal metaplasia via regulating CDX2. *Helicobacter*. Published online September 7, 2021:e12849. doi:10.1111/hel.12849
343. Asano N, Imatani A, Watanabe T, et al. Cdx2 Expression and Intestinal Metaplasia Induced by H. pylori Infection of Gastric Cells Is Regulated by NOD1-Mediated Innate Immune Responses. *Cancer Res*. 2016;76(5):1135-1145. doi:10.1158/0008-5472.CAN-15-2272
344. Barros R, Pereira B, Duluc I, et al. Key elements of the BMP/SMAD pathway co-localize with CDX2 in intestinal metaplasia and regulate CDX2 expression in human gastric cell lines. *J Pathol*. 2008;215(4):411-420. doi:10.1002/path.2369
345. Xu Y, Watanabe T, Tanigawa T, et al. Bile Acids Induce Cdx2 Expression Through the Farnesoid X Receptor in Gastric Epithelial Cells. *J Clin Biochem Nutr*. 2010;46(1):81-86. doi:10.3164/jcbrn.09-71
346. Yuan T, Ni Z, Han C, et al. SOX2 interferes with the function of CDX2 in bile acid-induced gastric intestinal metaplasia. *Cancer Cell International*. 2019;19(1):24. doi:10.1186/s12935-019-0739-8
347. Debruyne PR, Witek M, Gong L, et al. Bile acids induce ectopic expression of intestinal guanylyl cyclase C Through nuclear factor-kappaB and Cdx2 in human esophageal cells. *Gastroenterology*. 2006;130(4):1191-1206. doi:10.1053/j.gastro.2005.12.032
348. Li T, Guo H, Li H, et al. MicroRNA-92a-1-5p increases CDX2 by targeting FOXD1 in bile acids-induced gastric intestinal metaplasia. *Gut*. 2019;68(10):1751-1763. doi:10.1136/gutjnl-2017-315318
349. Sue S, Shibata W, Kameta E, et al. Intestine-specific homeobox (ISX) induces intestinal metaplasia and cell proliferation to contribute to gastric carcinogenesis. *J Gastroenterol*. 2016;51(10):949-960. doi:10.1007/s00535-016-1176-2
350. Samadani AA, Nikbakhsh N, Pilehchian M, Fattahi S, Akhavan-Niaki H. Epigenetic changes of CDX2 in gastric adenocarcinoma. *J Cell Commun Signal*. 2016;10(4):267-272. doi:10.1007/s12079-016-0327-9
351. Samadani AA, Nikbakhsh N, Taheri H, et al. CDX1/2 and KLF5 Expression and Epigenetic Modulation of Sonic Hedgehog Signaling in Gastric Adenocarcinoma. *Pathol Oncol Res*. 2019;25(3):1215-1222. doi:10.1007/s12253-019-00594-4
352. Tsukamoto T, Inada K, Tanaka H, et al. Down-regulation of a gastric transcription factor, Sox2, and ectopic expression of intestinal homeobox genes, Cdx1 and Cdx2: inverse correlation during progression from gastric/intestinal-mixed to complete intestinal metaplasia. *J Cancer Res Clin Oncol*. 2004;130(3):135-145. doi:10.1007/s00432-003-0519-6
353. Niu H, Jia Y, Li T, Su B. SOX2 Inhibition Promotes Promoter Demethylation of CDX2 to Facilitate Gastric Intestinal Metaplasia. *Dig Dis Sci*. 2017;62(1):124-132. doi:10.1007/s10620-016-4361-5
354. ZHANG JF, ZHANG JG, KUAI XL, et al. Reactivation of the homeotic tumor suppressor gene CDX2 by 5-aza-2'-deoxycytidine-induced demethylation inhibits cell proliferation and induces caspase-independent apoptosis in gastric cancer cells. *Exp Ther Med*. 2013;5(3):735-741. doi:10.3892/etm.2013.901
355. Chase A, Reiter A, Burci L, et al. Fusion of ETV6 to the caudal-related homeobox gene CDX2 in acute myeloid leukemia with the t(12;13)(p13;q12). *Blood*. 1999;93(3):1025-1031.

356. Rawat VPS, Cusan M, Deshpande A, et al. Ectopic expression of the homeobox gene Cdx2 is the transforming event in a mouse model of t(12;13)(p13;q12) acute myeloid leukemia. *Proc Natl Acad Sci U S A*. 2004;101(3):817-822. doi:10.1073/pnas.0305555101
357. Jafek JL, Shakya A, Tai PY, et al. Transcription factor Oct1 protects against hematopoietic stress and promotes acute myeloid leukemia. *Exp Hematol*. 2019;76:38-48.e2. doi:10.1016/j.exphem.2019.07.002
358. Rose NR, Klose RJ. Understanding the relationship between DNA methylation and histone lysine methylation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 2014;1839(12):1362-1372. doi:10.1016/j.bbagr.2014.02.007
359. Van Duyne GD. A structural view of cre-loxp site-specific recombination. *Annu Rev Biophys Biomol Struct*. 2001;30:87-104. doi:10.1146/annurev.biophys.30.1.87
360. Kühn R, Schwenk F, Aguet M, Rajewsky K. Inducible gene targeting in mice. *Science*. 1995;269(5229):1427-1429.
361. Mari L, Milano F, Parikh K, et al. A pSMAD/CDX2 Complex Is Essential for the Intestinalization of Epithelial Metaplasia. *Cell Reports*. 2014;7(4):1197-1210. doi:10.1016/j.celrep.2014.03.074
362. Skayneh H, Jishi B, Hleihel R, et al. A Critical Review of Animal Models Used in Acute Myeloid Leukemia Pathophysiology. *Genes (Basel)*. 2019;10(8):614. doi:10.3390/genes10080614
363. Mashima K, Azuma M, Fujiwara K, et al. Differential Localization and Invasion of Tumor Cells in Mouse Models of Human and Murine Leukemias. *Acta Histochem Cytochem*. 2020;53(3):43-53. doi:10.1267/ahc.19035
364. Khoud MB, Ingegnere T, Quesnel B, Mitra S, Brinster C. Acute Myeloid Leukemia: Is It T Time? *Cancers (Basel)*. 2021;13(10):2385. doi:10.3390/cancers13102385
365. Driss V, Quesnel B, Brinster C. Monocyte chemoattractant protein 1 (MCP-1/CCL2) contributes to thymus atrophy in acute myeloid leukemia. *European Journal of Immunology*. 2015;45(2):396-406. doi:10.1002/eji.201444736
366. Vu T, Straube J, Porter AH, et al. Hematopoietic stem and progenitor cell-restricted Cdx2 expression induces transformation to myelodysplasia and acute leukemia. *Nat Commun*. 2020;11(1):3021. doi:10.1038/s41467-020-16840-2
367. Paczulla A, Mbarga MN, Quintanilla-Martinez L, Lengerke C. CDX2 promotes leukemogenesis by modulating leukemic cell - bm niche interactions. *Experimental Hematology*. 2017;53:S73. doi:10.1016/j.exphem.2017.06.148
368. McKinney-Freeman SL, Lengerke C, Jang IH, et al. Modulation of murine embryonic stem cell-derived CD41+c-kit+ hematopoietic progenitors by ectopic expression of Cdx genes. *Blood*. 2008;111(10):4944-4953. doi:10.1182/blood-2007-11-124644
369. Rangatia J, Vangala RK, Singh SM, et al. Elevated c-Jun expression in acute myeloid leukemias inhibits C/EBPalpha DNA binding via leucine zipper domain interaction. *Oncogene*. 2003;22(30):4760-4764. doi:10.1038/sj.onc.1206664
370. Trivedi AK, Pal P, Behre G, Singh SM. Multiple ways of C/EBPalpha inhibition in myeloid leukaemia. *Eur J Cancer*. 2008;44(11):1516-1523. doi:10.1016/j.ejca.2008.04.022
371. Randrianarison-Huetz V, Laurent B, Bardet V, Blobe GC, Huetz F, Duménil D. Gfi-1B controls human erythroid and megakaryocytic differentiation by regulating TGF-β signaling at the bipotent erythro-megakaryocytic progenitor stage. *Blood*. 2010;115(14):2784-2795. doi:10.1182/blood-2009-09-241752

372. Kinross KM, Clark AJ, Iazzolino RM, Humbert PO. E2f4 regulates fetal erythropoiesis through the promotion of cellular proliferation. *Blood*. 2006;108(3):886-895. doi:10.1182/blood-2005-09-008656
373. Lee WY, Weinberg OK, Pinkus GS. GATA1 Is a Sensitive and Specific Nuclear Marker for Erythroid and Megakaryocytic Lineages. *Am J Clin Pathol*. 2017;147(4):420-426. doi:10.1093/ajcp/axq018
374. Zeddies S, Jansen SBG, di Summa F, et al. MEIS1 regulates early erythroid and megakaryocytic cell fate. *Haematologica*. 2014;99(10):1555-1564. doi:10.3324/haematol.2014.106567
375. Miller ME, Rosten P, Lemieux ME, Lai C, Humphries RK. Meis1 Is Required for Adult Mouse Erythropoiesis, Megakaryopoiesis and Hematopoietic Stem Cell Expansion. *PLoS One*. 2016;11(3):e0151584. doi:10.1371/journal.pone.0151584
376. Dumitriu B, Patrick MR, Petschek JP, et al. Sox6 cell-autonomously stimulates erythroid cell survival, proliferation, and terminal maturation and is thereby an important enhancer of definitive erythropoiesis during mouse development. *Blood*. 2006;108(4):1198-1207. doi:10.1182/blood-2006-02-004184
377. Wang X, Crispino JD, Letting DL, Nakazawa M, Poncz M, Blobel GA. Control of megakaryocyte-specific gene expression by GATA-1 and FOG-1: role of Ets transcription factors. *EMBO J*. 2002;21(19):5225-5234. doi:10.1093/emboj/cdf527
378. Ko M, Huang Y, Jankowska AM, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*. 2010;468(7325):839-843. doi:10.1038/nature09586
379. Sardina JL, Collombet S, Tian TV, et al. Transcription Factors Drive Tet2-Mediated Enhancer Demethylation to Reprogram Cell Fate. *Cell Stem Cell*. 2018;23(5):727-741.e9. doi:10.1016/j.stem.2018.08.016
380. Kallin EM, Rodríguez-Ubrea J, Christensen J, et al. Tet2 facilitates the derepression of myeloid target genes during CEBP α -induced transdifferentiation of pre-B cells. *Mol Cell*. 2012;48(2):266-276. doi:10.1016/j.molcel.2012.08.007
381. Mendes K, Schmidhofer S, Minderjahn J, et al. The epigenetic pioneer EGR2 initiates DNA demethylation in differentiating monocytes at both stable and transient binding sites. *Nat Commun*. 2021;12(1):1556. doi:10.1038/s41467-021-21661-y
382. Wang K, Sengupta S, Magnani L, Wilson CA, Henry RW, Knott JG. Brg1 Is Required for Cdx2-Mediated Repression of Oct4 Expression in Mouse Blastocysts. *PLoS One*. 2010;5(5):e10622. doi:10.1371/journal.pone.0010622
383. Bossen C, Murre CS, Chang AN, Mansson R, Rodewald HR, Murre C. The chromatin remodeler Brg1 activates enhancer repertoires to establish B cell identity and modulate cell growth. *Nat Immunol*. 2015;16(7):775-784. doi:10.1038/ni.3170
384. Gebuhr TC, Kovalev GI, Bultman S, Godfrey V, Su L, Magnuson T. The role of Brg1, a catalytic subunit of mammalian chromatin-remodeling complexes, in T cell development. *J Exp Med*. 2003;198(12):1937-1949. doi:10.1084/jem.20030714
385. Vradii D, Wagner S, Doan DN, et al. Brg1, the ATPase subunit of the SWI/SNF chromatin remodeling complex, is required for myeloid differentiation to granulocytes. *J Cell Physiol*. 2006;206(1):112-118. doi:10.1002/jcp.20432
386. Griffin CT, Brennan J, Magnuson T. The chromatin-remodeling enzyme BRG1 plays an essential role in primitive erythropoiesis and vascular development. *Development*. 2008;135(3):493-500. doi:10.1242/dev.010090

387. Enos ME, Bancos SA, Bushnell T, Crispe IN. E2F4 Modulates Differentiation and Gene Expression in Hematopoietic Progenitor Cells during Commitment to the Lymphoid Lineage. *The Journal of Immunology*. 2008;180(6):3699-3707. doi:10.4049/jimmunol.180.6.3699
388. Ficara F, Crisafulli L, Lin C, et al. Pbx1 restrains myeloid maturation while preserving lymphoid potential in hematopoietic progenitors. *J Cell Sci*. 2013;126(Pt 14):3181-3191. doi:10.1242/jcs.125435
389. Bröske AM, Vockentanz L, Kharazi S, et al. DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nat Genet*. 2009;41(11):1207-1215. doi:10.1038/ng.463
390. Skvarova Kramarzova K, Fiser K, Mejstrikova E, et al. Homeobox gene expression in acute myeloid leukemia is linked to typical underlying molecular aberrations. *Journal of Hematology & Oncology*. 2014;7(1):94. doi:10.1186/s13045-014-0094-0
391. Collins EM, Thompson A. HOX genes in normal, engineered and malignant hematopoiesis. *Int J Dev Biol*. 2018;62(11-12):847-856. doi:10.1387/ijdb.180206at
392. Collins CT, Hess JL. Role of HOXA9 in leukemia: dysregulation, cofactors and essential targets. *Oncogene*. 2016;35(9):1090-1098. doi:10.1038/onc.2015.174
393. Gwin KA, Shapiro MB, Dolence JJ, Huang ZL, Medina KL. Hoxa9 and Flt3 signaling synergistically regulate an early checkpoint in lymphopoiesis. *J Immunol*. 2013;191(2):745-754. doi:10.4049/jimmunol.1203294
394. Gwin K, Dolence JJ, Shapiro MB, Medina KL. Differential requirement for Hoxa9 in the development and differentiation of B, NK, and DC-lineage cells from Flt3+ multipotential progenitors. *BMC Immunol*. 2013;14:5. doi:10.1186/1471-2172-14-5
395. Zriwil A, Böiers C, Kristiansen TA, et al. Direct role of FLT3 in regulation of early lymphoid progenitors. *Br J Haematol*. 2018;183(4):588-600. doi:10.1111/bjh.15578
396. So CW, Karsunky H, Wong P, Weissman IL, Cleary ML. Leukemic transformation of hematopoietic progenitors by MLL-GAS7 in the absence of Hoxa7 or Hoxa9. *Blood*. 2004;103(8):3192-3199. doi:10.1182/blood-2003-10-3722
397. Wurm M, Kowalski J, Heckl D, et al. Ectopic expression of HOXC6 blocks myeloid differentiation and predisposes to malignant transformation. *Exp Hematol*. 2014;42(2):114-25.e4. doi:10.1016/j.exphem.2013.10.004
398. Arnaout HH, Mokhtar DA, Samy RM, Khames SA, Omar SA. CDX2 gene expression in acute lymphoblastic leukemia. *Journal of the Egyptian National Cancer Institute*. 2014;26(2):55-59. doi:10.1016/j.jnci.2013.12.002
399. Dong M. Role of transforming growth factor-beta in hematologic malignancies. *Blood*. 2006;107(12):4589-4596. doi:10.1182/blood-2005-10-4169
400. Lin Z, Gao C, Ning Y, He X, Wu W, Chen YG. The Pseudoreceptor BMP and Activin Membrane-bound Inhibitor Positively Modulates Wnt/ β -Catenin Signaling. *Journal of Biological Chemistry*. 2008;283(48):33053-33058. doi:10.1074/jbc.M804039200
401. Onichtchouk D, Chen YG, Dosch R, et al. Silencing of TGF-beta signalling by the pseudoreceptor BAMBI. *Nature*. 1999;401(6752):480-485. doi:10.1038/46794
402. Yan X, Lin Z, Chen F, et al. Human BAMBI cooperates with Smad7 to inhibit transforming growth factor-beta signaling. *J Biol Chem*. 2009;284(44):30097-30104. doi:10.1074/jbc.M109.049304
403. Sekiya T, Adachi S, Kohu K, et al. Identification of BMP and activin membrane-bound inhibitor (BAMBI), an inhibitor of transforming growth factor-beta signaling, as a target of the beta-catenin pathway in colorectal tumor cells. *J Biol Chem*. 2004;279(8):6840-6846. doi:10.1074/jbc.M310876200

404. Blank U, Karlsson S. TGF- signaling in the control of hematopoietic stem cells. *Blood*. 2015;125(23):3542-3550. doi:10.1182/blood-2014-12-618090
405. Challen GA, Boles NC, Chambers SM, Goodell MA. Distinct Hematopoietic Stem Cell Subtypes Are Differentially Regulated by TGFβ1. *Cell Stem Cell*. 2010;6(3):265-278. doi:10.1016/j.stem.2010.02.002
406. Bombara C, Ignatz RA. TGF-beta inhibits proliferation of and promotes differentiation of human promonocytic leukemia cells. *J Cell Physiol*. 1992;153(1):30-37. doi:10.1002/jcp.1041530106
407. Gonzalez-Junca A, Driscoll KE, Pellicciotta I, et al. Autocrine TGFβ Is a Survival Factor for Monocytes and Drives Immunosuppressive Lineage Commitment. *Cancer Immunol Res*. 2019;7(2):306-320. doi:10.1158/2326-6066.CIR-18-0310
408. Karlsson G, Blank U, Moody JL, et al. Smad4 is critical for self-renewal of hematopoietic stem cells. *J Exp Med*. 2007;204(3):467-474. doi:10.1084/jem.20060465
409. Blank U, Karlsson G, Moody JL, et al. Smad7 promotes self-renewal of hematopoietic stem cells. *Blood*. 2006;108(13):4246-4254. doi:10.1182/blood-2006-02-005611
410. Jiang L, Han X, Wang J, et al. SHP-1 regulates hematopoietic stem cell quiescence by coordinating TGF-β signaling. *J Exp Med*. 2018;215(5):1337-1347. doi:10.1084/jem.20171477
411. Hu M, Lu Y, Wang S, et al. CD63 acts as a functional marker in maintaining hematopoietic stem cell quiescence through supporting TGFβ signaling in mice. *Cell Death Differ*. Published online August 6, 2021. doi:10.1038/s41418-021-00848-2
412. Park SM, Deering RP, Lu Y, et al. Musashi-2 controls cell fate, lineage bias, and TGF-β signaling in HSCs. *J Exp Med*. 2014;211(1):71-87. doi:10.1084/jem.20130736
413. Hinge A, Filippi MD. Deconstructing the Complexity of TGFβ Signaling in Hematopoietic Stem Cells: Quiescence and Beyond. *Curr Stem Cell Rep*. 2016;2(4):388-397. doi:10.1007/s40778-016-0069-x
414. Chadwick K, Shojaei F, Gallacher L, Bhatia M. Smad7 alters cell fate decisions of human hematopoietic repopulating cells. *Blood*. 2005;105(5):1905-1915. doi:10.1182/blood-2004-03-0881
415. Hestdal K, Jacobsen SE, Ruscetti FW, Longo DL, Boone TC, Keller JR. Increased granulopoiesis after sequential administration of transforming growth factor-beta 1 and granulocyte-macrophage colony-stimulating factor. *Exp Hematol*. 1993;21(6):799-805.
416. Keller JR, Jacobsen SE, Sill KT, Ellingsworth LR, Ruscetti FW. Stimulation of granulopoiesis by transforming growth factor beta: synergy with granulocyte/macrophage-colony-stimulating factor. *Proc Natl Acad Sci U S A*. 1991;88(16):7190-7194. doi:10.1073/pnas.88.16.7190
417. Zermati Y, Fichelson S, Valensi F, et al. Transforming growth factor inhibits erythropoiesis by blocking proliferation and accelerating differentiation of erythroid progenitors. *Exp Hematol*. 2000;28(8):885-894. doi:10.1016/s0301-472x(00)00488-4
418. Fuchs O, Simakova O, Klener P, et al. Inhibition of Smad5 in human hematopoietic progenitors blocks erythroid differentiation induced by BMP4. *Blood Cells Mol Dis*. 2002;28(2):221-233. doi:10.1006/bcmd.2002.0487
419. Raghuvanshi S, Dahariya S, Sharma DS, Kovuru N, Sahu I, Gutti RK. RUNX1 and TGF-β signaling cross talk regulates Ca²⁺ ion channels expression and activity during megakaryocyte development. *FEBS J*. 2020;287(24):5411-5438. doi:10.1111/febs.15329
420. Jeanpierre S, Nicolini FE, Kaniewski B, et al. BMP4 regulation of human megakaryocytic differentiation is involved in thrombopoietin signaling. *Blood*. 2008;112(8):3154-3163. doi:10.1182/blood-2008-03-145326

421. Yasuda T, Wirtz T, Zhang B, et al. Studying Epstein-Barr virus pathologies and immune surveillance by reconstructing EBV infection in mice. *Cold Spring Harb Symp Quant Biol.* 2013;78:259-263. doi:10.1101/sqb.2013.78.020222
422. Sewastianik T, Jiang M, Sukhdeo K, et al. Constitutive Ras signaling and Ink4a/Arf inactivation cooperate during the development of B-ALL in mice. *Blood Adv.* 2017;1(25):2361-2374. doi:10.1182/bloodadvances.2017012211
423. Hobeika E, Dautzenberg M, Levit-Zerdoun E, Pelanda R, Reth M. Conditional Selection of B Cells in Mice With an Inducible B Cell Development. *Front Immunol.* 2018;9:1806. doi:10.3389/fimmu.2018.01806
424. Rumfelt LL, Zhou Y, Rowley BM, Shinton SA, Hardy RR. Lineage specification and plasticity in CD19- early B cell precursors. *J Exp Med.* 2006;203(3):675-687. doi:10.1084/jem.20052444
425. Hu T, Murdaugh R, Nakada D. Transcriptional and Microenvironmental Regulation of Lineage Ambiguity in Leukemia. *Front Oncol.* 2017;7:268. doi:10.3389/fonc.2017.00268
426. Gröschel S, Sanders MA, Hoogenboezem R, et al. A Single Oncogenic Enhancer Rearrangement Causes Concomitant EVI1 and GATA2 Deregulation in Leukemia. *Cell.* 2014;157(2):369-381. doi:10.1016/j.cell.2014.02.019
427. Lu X, Freund JN, Muller M, et al. Differential regulation of CDX1 and CDX2 gene expression by deficiency in methyl group donors. *Biochimie.* 2008;90(5):697-704. doi:10.1016/j.biochi.2007.12.002
428. Benahmed F, Gross I, Guenot D, et al. The Microenvironment Controls CDX2 Homeobox Gene Expression in Colorectal Cancer Cells. *Am J Pathol.* 2007;170(2):733-744. doi:10.2353/ajpath.2007.060696
429. Waldum HL, Fossmark R. Types of Gastric Carcinomas. *Int J Mol Sci.* 2018;19(12):4109. doi:10.3390/ijms19124109
430. Park DY, Srivastava A, Kim GH, et al. CDX2 expression in the intestinal-type gastric epithelial neoplasia: frequency and significance. *Mod Pathol.* 2010;23(1):54-61. doi:10.1038/modpathol.2009.135
431. Guo L, Lee YT, Zhou Y, Huang Y. Targeting epigenetic regulatory machinery to overcome cancer therapy resistance. *Semin Cancer Biol.* Published online January 6, 2021:S1044-579X(20)30282-0. doi:10.1016/j.semcancer.2020.12.022
432. Segert JA, Gisselbrecht SS, Bulyk ML. Transcriptional Silencers: Driving Gene Expression with the Brakes On. *Trends Genet.* 2021;37(6):514-527. doi:10.1016/j.tig.2021.02.002
433. Gisselbrecht SS, Palagi A, Kurland JV, et al. Transcriptional Silencers in Drosophila Serve a Dual Role as Transcriptional Enhancers in Alternate Cellular Contexts. *Mol Cell.* 2020;77(2):324-337.e8. doi:10.1016/j.molcel.2019.10.004
434. Huang D, Petrykowska HM, Miller BF, Elnitski L, Ovcharenko I. Identification of human silencers by correlating cross-tissue epigenetic profiles and gene expression. *Genome Res.* 2019;29(4):657-667. doi:10.1101/gr.247007.118
435. Cheng CK, Wong THY, Wan TSK, et al. RUNX1 upregulation via disruption of long-range transcriptional control by a novel t(5;21)(q13;q22) translocation in acute myeloid leukemia. *Molecular Cancer.* 2018;17(1):133. doi:10.1186/s12943-018-0881-2
436. Heintzman ND, Stuart RK, Hon G, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet.* 2007;39(3):311-318. doi:10.1038/ng1966

437. Krijger PHL, Geeven G, Bianchi V, Hilvering CRE, de Laat W. 4C-seq from beginning to end: A detailed protocol for sample preparation and data analysis. *Methods*. 2020;170:17-32. doi:10.1016/j.ymeth.2019.07.014
438. Barros R, Pereira D, Callé C, et al. Dynamics of SOX2 and CDX2 Expression in Barrett's Mucosa. *Dis Markers*. 2016;2016:1532791. doi:10.1155/2016/1532791
439. Watts JA, Zhang C, Klein-Szanto AJ, et al. Study of FoxA pioneer factor at silent genes reveals Rfx-repressed enhancer at Cdx2 and a potential indicator of esophageal adenocarcinoma development. *PLoS Genet*. 2011;7(9):e1002277. doi:10.1371/journal.pgen.1002277

onction et régulation du gène homéotique Cdx2 dans les leucémies aiguës

Résumé

CDX2 est gène homéotique exerçant de nombreuses fonctions au cours du développement embryonnaire. Son expression est restreinte au niveau de l'épithélium intestinal chez l'adulte, où il participe au maintien de l'identité et de l'homéostasie intestinale, et exerce une fonction de suppresseur de tumeur. Une expression ectopique de CDX2 est observée dans la moelle osseuse de 80% des patients atteints de leucémie aiguë, où il est oncogénique. Les mécanismes menant à cette activation aberrante et à l'oncogenèse induite sont inconnus.

Au cours de ma thèse, je me suis intéressée à sa fonction oncogénique dans le compartiment hématopoïétique en y induisant son expression dans un modèle murin. Celui-ci développe des leucémies aiguës myéloïdes, caractérisées par une expansion de monocytes immatures. Le blocage de leur maturation fait notamment intervenir une diminution de la signalisation par la voie du TGF- β . J'ai d'autre part initié une étude épigénétique globale visant à étudier les mécanismes de régulation du gène Cdx2, via l'identification de enhancers responsables de son expression dans 3 contextes cellulaires différents.

Mots-clefs : CDX2, Oncogène, Leucémie aiguë, TGF- β , Hématopoïèse, Epigénétique.

Abstract

CDX2 is a homeotic gene with many functions during the embryonic development. Its expression is restricted to the intestinal epithelium in adults, where it participates in the maintenance of intestinal identity and homeostasis, and has a tumor suppressor function. Ectopic expression of CDX2 is observed in the bone marrow of 80% of patients with acute leukemia, where it is oncogenic. The mechanisms leading to this aberrant activation and induced oncogenesis are unknown.

During my thesis, I investigated its oncogenic function in the hematopoietic compartment by inducing its expression in a mouse model. This model develops acute myeloid leukemia, characterized by an expansion of immature monocytes. The blockage of their maturation involves a decrease of the TGF- β signaling pathway. I also initiated a global epigenetic study aiming at studying the regulation mechanisms of the Cdx2 gene, via the identification of enhancers responsible for its expression in 3 different cellular contexts.

Keywords: CDX2, Oncogene, Acute leukemia, TGF- β , Hematopoiesis, Epigenetic.